

RL-TR-93-119, Vol II (of two)
In-House Report
June 1993

AD-A266 916



2

PROCEEDINGS OF THE 1992 ANTENNA APPLICATIONS SYMPOSIUM

Paul Mayes, et al

DTIC
ELECTE
JUL 20 1993
S E D

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York

93-16335



93

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-93-119, Vol II, (of two) has been reviewed and is approved for publication.

APPROVED:



DANIEL J. JACAVANCO, Chief
Antennas and Components Division
Electromagnetics and Reliability Directorate

FOR THE COMMANDER:



JOHN K. SCHINDLER
Director of Electromagnetics and Reliability

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (ERAS) Hanscom AFB MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 1993		3. REPORT TYPE AND DATES COVERED In-house Volume II
4. TITLE AND SUBTITLE Proceedings of the 1992 Antenna Applications Symposium			5. FUNDING NUMBERS PE 62702F PR 4600 TA 14 WU PH	
6. AUTHOR(S) Paul Mayes, et al				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Rome Laboratory/ERAS 31 Grenier Street Hanscom AFB, MA 01731-3010 Project Engineer: John Antonucci/ERAS			8. PERFORMING ORGANIZATION REPORT NUMBER RL-TR-93-119 (II)	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Volume I consists of pages 1 through 336 Volume II consists of pages 337 through 644				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The Proceedings of the 1992 Antenna Applications Symposium is a collection of state-of-the-art papers relating to phased array antennas, multibeam antennas, satellite antennas, microstrip antennas, reflector antennas, HF, VHF, UHF and various other antennas.				
14. SUBJECT TERMS Antennas Satellite antennas Broadband antennas		Microstrip Reflector HF, VHF, UHF		Multibeam antennas Array antennas
17. SECURITY CLASSIFICATION OF REPORT Unclassified		18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified		19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified
15. NUMBER OF PAGES 314			16. PRICE CODE	
20. LIMITATION OF ABSTRACT SAR				

CONTENTS

WEDNESDAY. 23 SEPTEMBER 1992

OPTOELECTRONIC ARRAYS

Keynote: Optically Fed and controlled Phased Array Antennas - A Dream or a Real Possibility?	1
An Ultrafast Optoelectronic THz Beam System	16
Antenna True-Time-Delay Beamsteering Utilizing Fiber Optics	46
An Optically Controlled Ka-Band Phased Array Antenna	63
* Monolithic Photonic Receiver for Phased Array Signal Distribution	
THz Electromagnetic Radiation From Planar Photoconducting Structures	77

PHASED ARRAYS

Neural Beamforming for Phased Array Antennas	106
Digital Beamforming	116
Failure Monitoring and Correction in Phased Arrays	124
Waveguide Distribution Networks for EHF Phased Array Antennas	154
Wide Angle Impedance Matching Surfaces for Circular Waveguide Phased Array Antennas With 70 Degree Scan Capability	176
Broadband Array Antennas of Dual Flared Notch Elements	198
A Multiband Phased Array Antenna	218
Design of an Integral IFF Array for a Slotted Antenna	246
* NOT INCLUDED IN THIS VOLUME	

THURSDAY, 24 SEPTEMBER 1992

BROADBAND AND/OR LOW PROFILE ELEMENTS

A Wideband Sub-Array Radiator for Advanced Avionics Applications	255
Waveguide Excited Microstrip Patch Antenna: Theory and Experiment	285
A DFT Synthesis Method for Finite Arrays of Dipoles on Layered Media	307
Pentagonal Microstrip Patch Antenna for Circular Polarization - Revisited	337
Planar Spiral, A Microstrip Antenna?	363
* Half Space Log Periodic Antenna for VHF	
Analysis of Log-Periodic Folded Dipole Array	395
A Broadband Panel Antenna for HDTV Applications	403

COMPUTATIONAL ELECTROMAGNETICS

Necessary and Sufficient Conditions for Portable, Reliable, and Useful CEM Software	417
Advances in Modeling and Simulation of Complex Radiating Structures and Platforms: The Role of Computational Electromagnetics	433
Benchmark Problems for Electromagnetics Software Validation: An Overview of the Joint ACES/IEE-AP Effort	457
Computational Electromagnetics - Considerations for the Next Generation of Supercodes	465
* Microstrip Antenna Computation Research Using Massively Parallel Processors	
Fast Algorithms for Wave Scattering Developed at the Electromagnetics Laboratory, University of Illinois	495
* NOT INCLUDED IN THIS VOLUME	

FRIDAY, 25 SEPTEMBER 1992

BROADBAND AND OTHERS

Evaluation of Frequency Selective Reflector Antenna Systems	534
* E-3A Reflectionless Manifold Performance in the Presence of Mismatched Radiating Elements	
Productivity Gains Using Multiple-Channel, Multiple- Frequency Measurement Techniques for Testing the E2C Antenna	552
Calibration of Mismatch Errors in Antenna Gain Measurements	558
A Novel X-Band, Circularly Polarized Feed for the I-30 Radar Antenna System	564
Invisible U-Shaped Slot Communications Antenna	594
Design Considerations and Analysis of a Cross Interferometer Antenna for Airborne Applications	607
Degree of Freedom Requirements for Angular Sector Nulling	637

* NOT INCLUDED IN THIS VOLUME

DTIC QUALITY INSPECTED 5

Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

PENTAGONAL MICROSTRIP PATCH ANTENNA FOR CIRCULAR POLARIZATION - REVISITED

K. A. Michalski and C.-I. G. Hsu

**Electromagnetics & Microwave Laboratory
Department of Electrical Engineering
Texas A&M University
College Station, Texas 77843-3128**

Abstract

A coax-fed pentagonal microstrip patch antenna is analyzed using the mixed-potential integral equation approach. The substrate (and superstrate, if present) may be electrically thick and uniaxially anisotropic. Vector basis functions defined over triangular finite elements are used to approximate the current distribution on the conducting patch. The coax pin current is expanded in terms of piecewise linear subdomain basis functions and the coax aperture is modeled in terms of a magnetic current frill. A simple probe-to-patch junction attachment mode, compatible with the triangular element model of the microstrip patch, is implemented. Sample computed results are presented for a circularly polarized pentagonal antenna and are compared with measured data.

1 Introduction

The pentagonal microstrip patch antenna was first proposed and experimentally investigated by Weinschel [1], who demonstrated that it can radiate a

circularly polarized field when excited from a single, properly placed feed. The antenna geometry is shown in Fig. 1, where the feed point is denoted

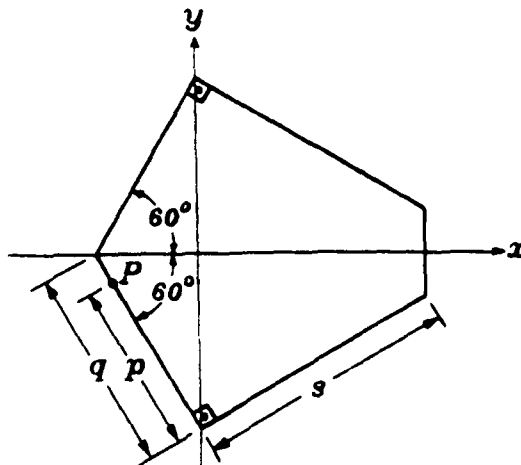


Figure 1: Pentagonal microstrip patch antenna.

by 'P.' This patch, which is symmetric with respect to the x axis, may be viewed as one third of an equilateral triangular patch with the tip cut off. The first theoretical analysis of this antenna was done by Carver and Coffey [2], who used a cavity model in conjunction with the finite element method. As was explained by Richards et al. [3], the circular polarization is obtained by exciting two orthogonal modes of the antenna, which have close resonant frequencies. An improved theory, which utilizes the concept of 'effective dimensions' to account for the fringing effect, was presented by Suzuki and Chiba [4], and by Haneishi and Suzuki [5]. Until recently, the pentagonal patch was not amenable to any other existing analysis technique.

In this paper, we revisit the pentagonal microstrip patch antenna problem using a space domain approach based on the mixed-potential integral equation formulation in conjunction with a triangular element model of the patch. In this method, which was pioneered by Pichon et al. [6] and adopted

by others [7], [8], [9], [10], [11], [12], [13], [14], the effects of the substrate and superstrate (if present) are rigorously taken into account by means of the vector and scalar potential Green's functions for the layered medium, which is assumed to be of infinite extent. The current distribution on the microstrip patch, which may have arbitrary shape, is approximated in terms of vector basis functions defined over triangular subdomains [15], and the current expansion coefficients are found by the method of moments (MOM) [16].

In the approach presented here, the substrate and the cover layer (if present) may be electrically thick and uniaxially anisotropic. The current on the coax pin is expanded in terms of piecewise linear subdomain basis functions and the coax aperture is modeled by a magnetic current frill [17], [10]. A simple probe-to-patch junction attachment mode [9], compatible with the triangular element model of the microstrip patch, is implemented. The corresponding eigenvalue problem, where there is no excitation [12], is also considered. Sample computed results are presented for both the source-free and coax-excited pentagonal microstrip patch antenna and are compared with measured data.

2 Formulation

A. Problem statement and assumptions

The geometry of the problem under consideration is illustrated in Fig. 2, where the substrate and the cover layer, which are assumed to be nonmagnetic, are characterized by their transverse and longitudinal dielectric constants, ϵ_{tn} and ϵ_{zn} , respectively, relative to free space, and anisotropy ratios $\nu_n = \epsilon_{zn}/\epsilon_{tn}$. If dielectric losses are present, ϵ_{tn} and ϵ_{zn} are multiplied by $(1 - j \tan \delta_n)$, where $\tan \delta_n$ is the loss tangent of the n th layer. (Here and throughout this paper, the $e^{j\omega t}$ time convention is implied.) The coax probe,

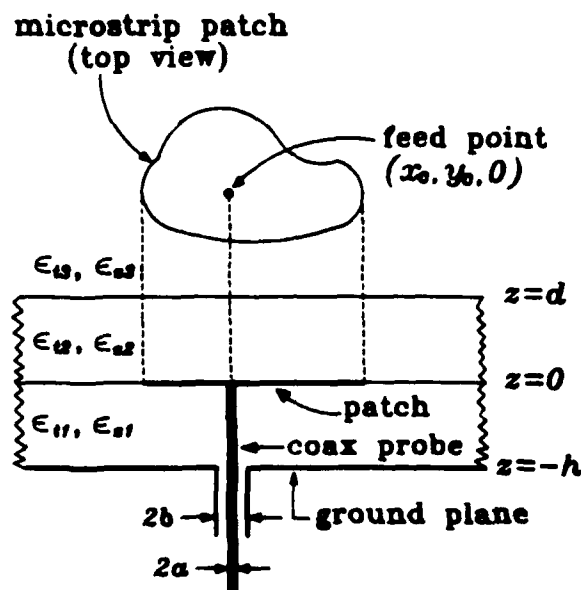


Figure 2: Geometry of a coax-fed microstrip patch antenna.

which has the inner and outer radii a and b , respectively, is centered at (x_c, y_c) . The probe and the patch are assumed to be perfectly conducting.

By invoking the equivalence principle [18], the original problem of Fig. 2 may be replaced by its equivalent, shown in Fig. 3. In the latter, the conducting patch and the coax pin are replaced by equivalent electric currents. Also, an equivalent magnetic current frill, M_s , is placed over the coax aperture, which is shorted. To simplify the analysis, we assume that the distribution of the aperture electric field is that of the TEM coax mode, which also specifies M_s . The latter may then be regarded as the source of the 'incident' field exciting the antenna. We further assume that the coax probe current is z -directed and is azimuthally invariant on its circumference. As a result, the patch current, J_s , which resides on the surface S and represents the vector sum of the currents that exist on the bottom and top sides of the microstrip patch, and the total coax pin current, I , uniformly distributed on

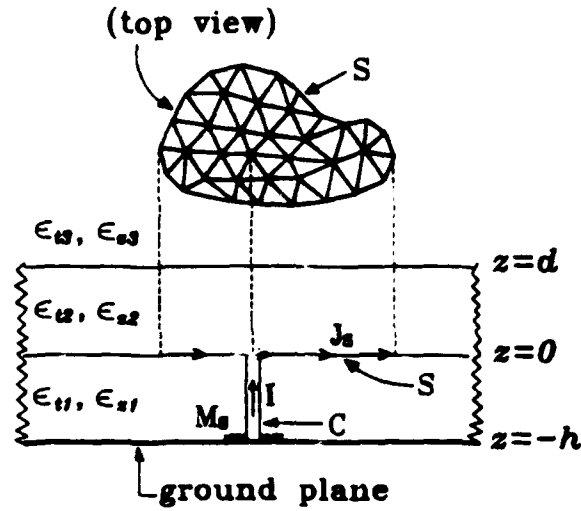


Figure 3: Problem equivalent to that in Fig. 2.

the cylindrical surface C , remain the basic unknowns of the problem.

Although this paper concerns a pentagonal patch antenna, the formulation is general and is applicable to any shape that is amenable to a triangular mesh approximation.

B. Integral equation formulation

Upon using the Green's functions for the layered medium of Fig. 3, and enforcing the condition of vanishing tangential electric fields on the patch and the pin, we may derive a coupled set of integral equations for J_s and I , as

$$\begin{aligned} \int_S G_{zz}(\mathbf{r}|\mathbf{r}') J_s(\rho') dS' + \nabla_s \int_S G_\phi(\mathbf{r}|\mathbf{r}') \nabla_s' \cdot J_s(\rho') dS' \\ + \int_{-h}^0 [\hat{x} K_{xz}(\mathbf{r}|z') + \hat{y} K_{yz}(\mathbf{r}|z')] I(z') dz' \\ + \nabla_s \int_{-h}^0 K_\phi(\mathbf{r}|z') \frac{dI(z')}{dz'} dz' = \mathbf{E}_s^i(\mathbf{r}), \quad \mathbf{r} \in S \end{aligned} \quad (1)$$

$$\begin{aligned}
& \int_S [\mathcal{G}_{xx}(z|\mathbf{r}')\hat{x} + \mathcal{G}_{xy}(z|\mathbf{r}')\hat{y}] \cdot \mathbf{J}_s(\boldsymbol{\rho}') dS' + \frac{d}{dz} \int_S \mathcal{G}_\phi(z|\mathbf{r}') \nabla'_s \cdot \mathbf{J}_s(\boldsymbol{\rho}') dS' \\
& + \int_{-h}^0 \mathcal{K}_{xx}(z|z') I(z') dz' + \frac{d}{dz} \int_{-h}^0 \mathcal{K}_\phi(z|z') \frac{dI(z')}{dz'} dz' \\
& = \mathcal{E}_x^i(z), \quad -h < z < 0
\end{aligned} \quad (2)$$

where $\mathbf{r}$ is the vector from the coordinate origin to a point (x, y, z) , $\boldsymbol{\rho}$ is the projection of $\mathbf{r}$ on the (x, y) plane, unit vectors are distinguished by '·', the subscript 's' denotes components tangential to S , and a prime signifies that the respective quantity is evaluated at the source coordinates. In (2), we only enforce the vanishing of the z component of the electric field in the average on the circumference of the coax probe.

Using the notation [19]

$$S_n \{f(k_p)\} = \frac{1}{2\pi} \int_0^\infty f(k_p) J_n(k_p \xi) k_p^{n+1} dk_p, \quad n=0, 1 \quad (3)$$

$$\xi = \sqrt{(x-x')^2 + (y-y')^2}, \quad \zeta = \arctan\left(\frac{y-y'}{x-x'}\right) \quad (4)$$

where J_n is the Bessel function of order n , the kernel functions appearing in (1)-(2) can be expressed as

$$G_{xx}(\mathbf{r}|\mathbf{r}') = S_0 \{V_i^h(k_p; z|z')\} \quad (5)$$

$$G_{xz}(\mathbf{r}|\mathbf{r}') = -jk_0\eta_0 \cos \zeta S_1 \left\{ \frac{I_i^h(k_p; z|z') - I_i^e(k_p; z|z')}{k_p^2} \right\} \quad (6)$$

$$G_{zz}(\mathbf{r}|\mathbf{r}') = -jk_0\eta_0 \cos \zeta S_1 \left\{ \frac{V_v^h(k_p; z|z') - V_v^e(k_p; z|z')}{k_p^2} \right\} \quad (7)$$

$$G_{zz}(\mathbf{r}|\mathbf{r}') = \eta_0^2 S_0 \left\{ \left(\frac{k_0}{k_p}\right)^2 I_v^h(k_p; z|z') + \left[\frac{\epsilon_z + \epsilon'_z}{\epsilon_z \epsilon'_z} - \left(\frac{k_0}{k_p}\right)^2 \right] I_v^e(k_p; z|z') \right\} \quad (8)$$

$$G_\phi(\mathbf{r}|\mathbf{r}') = S_0 \left\{ \frac{V_i^h(k_p; z|z') - V_i^e(k_p; z|z')}{k_p^2} \right\} \quad (9)$$

where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ and $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ denote, respectively, the intrinsic impedance and wavenumber of free space, and where we follow the convention that primed quantities are evaluated in the source layer, while the unprimed parameters pertain to the layer in which the field is being evaluated. The remaining kernels, G_{xy} and G_{yx} , are given by (6) and (7), respectively, with $\cos \zeta$ replaced by $\sin \zeta$. In (1)–(2), we use the notation

$$K_\phi(\mathbf{r}|z') = \frac{1}{2\pi a} \oint_C G_\phi(\mathbf{r}|\mathbf{r}') d\ell', \quad \mathbf{r}' \in C \quad (10)$$

where the integral, which is along the circumference of the coax probe, may be introduced inside the spectral integral in (9) and evaluated in closed form. As a result, (x', y') are replaced by (x_c, y_c) and the original integrand in (3) is multiplied by $J_0(k_p a)$. Similar remarks apply to K_{xx} , K_{yy} , and K_{zz} . The script symbols used in (2) denote the average values of the respective kernels on the circumference of the probe. Hence,

$$\mathcal{G}_\phi(z|\mathbf{r}') = \frac{1}{2\pi a} \oint_C C_\phi(\mathbf{r}|\mathbf{r}') d\ell, \quad \mathbf{r} \in C, \mathbf{r}' \in S \quad (11)$$

and similarly for $\mathcal{G}_{xx}$ and $\mathcal{G}_{yy}$. The effect of this integration is to multiply the original integrand in (3) by $J_0(k_p a)$ and to replace (x, y) by (x_c, y_c) . The coax kernels K_ϕ and K_{zz} are azimuthally invariant when $\mathbf{r} \in C$, hence we put $\mathcal{K}_\phi(z|z') = K_\phi(\mathbf{r}|z')$, where in the latter $\xi = a$, and similarly for $\mathcal{K}_{zz}$.

Finally, it is important to note that the integral equations (1)–(2) are in the mixed-potential form [20] amenable to the existing numerical solution procedures developed for scatterers of arbitrary shape residing in a free space [15].

C. Transmission-line network analog of the layered medium

As an aid in deriving (5)–(9), we have employed a transmission line network analog of the layered medium, in which each layer is represented by a transmission line section. This analog comprises two networks, which arise

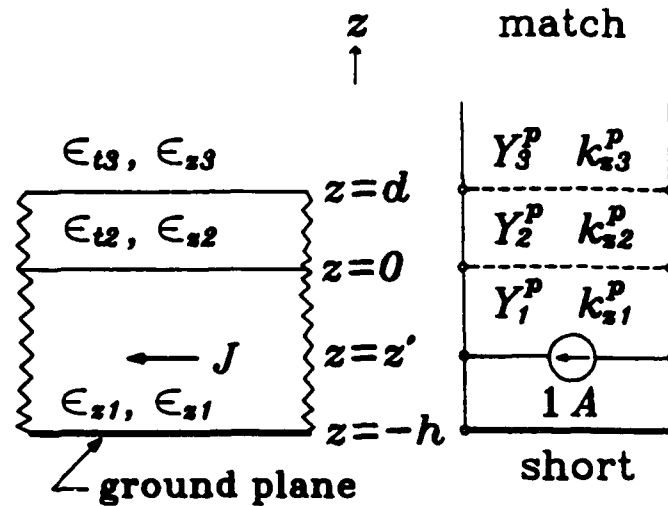


Figure 4: Transmission-line network analog of the layered medium.

from the decomposition of the electromagnetic field into partial fields that are transverse-magnetic (TM) and transverse-electric (TE) to z [21, pp. 185–217]. The quantities corresponding to the TM and TE networks are distinguished by the superscripts 'e' and 'h', respectively. The voltage and current transmission-line Green's functions, $V_i^p(k_p; z|z')$ and $I_i^p(k_p; z|z')$, where 'p' stands for 'e' or 'h', represent the voltage and current, respectively, at a point z , excited by a unit-strength current source located at z' . This situation is illustrated in Fig. 4. The other Green's functions, $V_v^p(k_p; z|z')$ and $I_v^p(k_p; z|z')$, are generated by a unit-strength voltage source. The characteristic admittance and propagation constant of the n th section of the TE and TM transmission lines are given, respectively, as

$$Y_n^h = \frac{k_{zn}^h}{k_0 \eta_0}, \quad k_{zn}^h = \sqrt{k_0^2 \epsilon_{tn} - k_p^2} \quad (12)$$

and

$$Y_n^e = \frac{k_0 \epsilon_{tn}}{\eta_0 k_{zn}^e}, \quad k_{zn}^e = \sqrt{k_0^2 \epsilon_{tn} - k_p^2 / \nu_n} \quad (13)$$

It should be noted that the integral equations (1)–(2) are valid for a medium with an arbitrary number of layers, provided that the appropriate transmission-line Green's functions are used in (5)–(9). These Green's functions, which are easily determined by the network theory methods, are not listed here due to lack of space.

It can be shown that the kernel functions which appear in (1)–(2) exhibit only mild singularities when the source and observation points coincide on S and C , making these equations particularly amenable to numerical solution procedures.

D. Excitation model

We assume that the aperture field in the problem of Fig. 2 is that of a TEM coax mode with a voltage V_{DV} . Consequently, with respect to a polar coordinate system (ρ, ϑ) centered at (x_c, y_c) , the magnetic surface current M_s in Fig. 3 has only a ϑ component, given as [18, p. 112]

$$M_\vartheta(\rho) = -\frac{K}{\rho}, \quad K = \frac{V_{DV}}{\ln(b/a)} \quad (14)$$

in which $a < \rho < b$. The 'incident' electric field E^i in (1)–(2), due to this magnetic current frill embedded at $z = -h$ in the layered medium can be expressed as

$$E_z^i(\mathbf{r}) = 2\pi K \cos \zeta S_1 \left\{ [J_0(k_p a) - J_0(k_p b)] \frac{V_v^e(k_p; z| - h)}{k_p^2} \right\} \quad (15)$$

$$E_r^i(\mathbf{r}) = \frac{2\pi\eta_0 K}{jk_0\epsilon_r} S_0 \{ [J_0(k_p a) - J_0(k_p b)] I_v^e(k_p; z| - h) \} \quad (16)$$

where $(x', y') = (x_c, y_c)$. $E_y^i(\mathbf{r})$ is given by (15) with $\cos \zeta$ replaced by $\sin \zeta$. Since $E_z^i(\mathbf{r})$ is azimuthally invariant when $\mathbf{r} \in C$, we set in (2) $\mathcal{E}_z^i(z) = E_z^i(\mathbf{r})$, where the latter is given by (16) with $\xi = a$. We note that (15)–(16) are in a general form applicable to a medium with an arbitrary number of layers.

E. Input impedance and far fields

Once the integral equations (1)–(2) are solved for the currents J_s and I , the antenna's input impedance may be found from

$$Z_{IN} = \frac{V_{IN}}{I_{IN}} \quad (17)$$

where I_{IN} is the current at the base of the coax probe. This simple model of the coax feed is only accurate for electrically small radii a and b . More elaborate models [22], [23], which take into account the higher-order modes in the coax, could also be used in the framework of the present formulation.

The far field radiated by J_s and I can be determined by the stationary-phase method [24]. It is then found that the contribution of the patch current to the far field in the direction (θ, φ) is

$$E_\theta \sim V_i^e(k_0 \sin \theta; d|0) \int_S \hat{\rho} \cdot J_s(\rho') e^{jk_0 \rho' \sin \theta \cos(\varphi - \varphi')} dS' \quad (18)$$

$$E_\varphi \sim \cos \theta V_i^h(k_0 \sin \theta; d|0) \int_S \hat{\varphi} \cdot J_s(\rho') e^{jk_0 \rho' \sin \theta \cos(\varphi - \varphi')} dS' \quad (19)$$

and that the coax probe contributes

$$E_\theta \sim \sin \theta \cos \theta \frac{\eta_0^2}{\epsilon_z'} \int_{-h}^0 I_v^e(k_0 \sin \theta; d|z') I(z') dz' \quad (20)$$

We have omitted in the above a common spherical wave factor and some phase terms that do not affect the field magnitude. It is also assumed here that the medium of the upper half-space is air, and that the coax probe radius is electrically small ($k_0 a \ll 1$). Similar expressions for the far field of the magnetic frill current can also be given, but are omitted here because the radiation from the coax aperture, which is assumed to be electrically small, has been found to be negligible in comparison to the radiation by the patch and the coax probe. We note that (18)–(20) are valid for multilayer media, provided that the appropriate transmission-line Green's functions are employed.

3 Numerical Method

A. Weak form of the integral equations

The MOM is applied to the weak forms of (1) and (2), which are obtained by 'testing' them with suitably selected weight functions $\{A_k\}$ and $\{N_k\}$, respectively. As a result, upon integrating by parts, we obtain

$$\begin{aligned} & \left\langle A_k; \left\langle G_{xx}, J_s \right\rangle'_S \right\rangle_S - \left\langle \nabla_s \cdot A_k; \left\langle G_\phi, \nabla'_s \cdot J_s \right\rangle'_S \right\rangle_S \\ & + \left\langle A_k; \left\langle \hat{x} K_{xx} + \hat{y} K_{yy}, I \right\rangle'_C \right\rangle_C - \left\langle \nabla_s \cdot A_k; \left\langle K_\phi, \frac{dI}{dz'} \right\rangle'_C \right\rangle_C \\ & = \left\langle A_k; E_s^i \right\rangle_S \end{aligned} \quad (21)$$

$$\begin{aligned} & \left\langle N_k; \left\langle G_{xx} \hat{x} + G_{yy} \hat{y}; J_s \right\rangle'_S \right\rangle_C - \left\langle \frac{dN_k}{dz}; \left\langle G_\phi, \nabla'_s \cdot J_s \right\rangle'_S \right\rangle_C \\ & + \left\langle N_k; \left\langle K_{xx}, I \right\rangle'_C \right\rangle_C - \left\langle \frac{dN_k}{dz}; \left\langle K_\phi, \frac{dI}{dz'} \right\rangle'_C \right\rangle_C \\ & = \left\langle N_k; \mathcal{E}_z^i \right\rangle_C \end{aligned} \quad (22)$$

where we have introduced the notation ' $\langle ; \rangle$ ' for an integral of a product of two functions separated by the comma. The dot over the comma signifies a 'dot product' of vector arguments. The prime over ' $\rangle$ ' indicates that the integration is over the primed coordinates, while the subscript 'S' or 'C' indicates the domain of integration (see Fig. 3).

B. Patch current expansion

The microstrip patch is modeled by triangular elements, as indicated in Fig. 3. The probe-to-patch junction can be located anywhere on the patch, including edges and corners, but it must coincide with a node of the triangular element mesh. The nodes of each triangular element are assigned the indices i , j , and k in the counterclockwise direction, as illustrated in Fig. 5.

We adopt here a local indexing scheme, in which these indices assume the values 1, 2, or 3 in a cyclic manner. The sides of the n th triangle, S_n , with

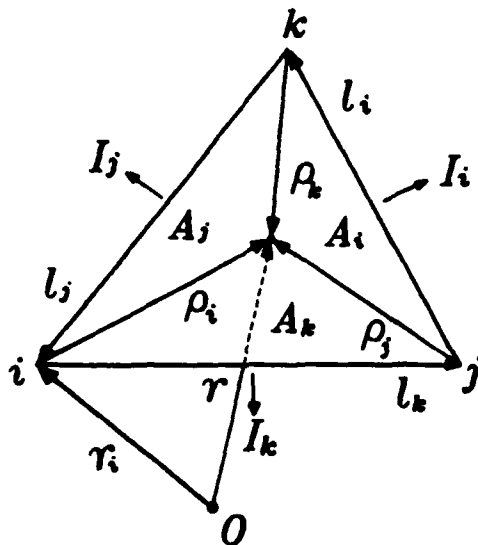


Figure 5: Local coordinates associated with a triangular element.

an area $A^{(n)}$, are formed by three edge vectors $\ell_i^{(n)}$, with $\ell_i^{(n)}$ oriented from node j to node k . The i th node of S_n is located by the position vector $r_i^{(n)}$. The patch current density and its divergence on S_n are represented as

$$J_s^{(n)}(\rho) = \sum_{i=1}^3 I_i^{(n)} A_i^{(n)}, \quad \nabla \cdot J_s^{(n)}(\rho) = \sum_{i=1}^3 \frac{I_i^{(n)}}{A^{(n)}} \quad (23)$$

where $I_i^{(n)}$ is the total current leaving the i th edge, and $A_i^{(n)}$ is a vector basis function defined as [15]

$$A_i^{(n)} = \frac{\rho_i^{(n)}}{2A^{(n)}}, \quad \rho_i^{(n)} = \ell_k^{(n)} L_j - \ell_j^{(n)} L_k \quad (24)$$

Here, L_i is the area coordinate associated with the i th node [25, p. 110], and is given as

$$L_i = \frac{A_i^{(n)}}{A^{(n)}}, \quad \sum_{i=1}^3 L_i = 1 \quad (25)$$

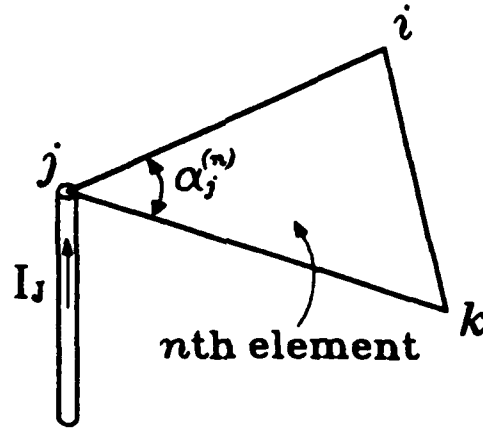


Figure 6: Geometry of the probe-to-patch junction.

where $A_i^{(n)}$ is the area of the triangle formed by the observation point within the n th element and the nodes j and k .

C. Attachment mode

The current expansion (23) is not suitable for an element having the probe-to-patch junction point as one of its nodes. To represent the current distribution on such 'junction elements,' a special 'attachment mode' is required, which is superposed on the non-junction part of the current, represented by (23). Let there be N_J elements attached to the junction node, where the current I_J enters the patch. As illustrated in Fig. 6, the junction node of each element is assigned a local index j , and the angle between the two edges of the n th junction element that meet at this node is designated $\alpha_j^{(n)}$. The sum of the junction vertex angles is denoted by α_J . The current density and its divergence on the n th junction element are expressed as [10]

$$\mathbf{J}_s^{(n)}(\rho) = -\frac{I_J \alpha_j^{(n)}}{2\alpha_J} [A_i^{(n)} + A_k^{(n)}], \quad \nabla \cdot \mathbf{J}_s^{(n)}(\rho) = -\frac{I_J \alpha_j^{(n)}}{\alpha_J A^{(n)}} \quad (26)$$

We note that this current is 'injected' into the patch from the element edges that meet at the junction node, and does not model the diverging current

behavior near the junction. However, the current continuity at the junction is satisfied in a global sense, i.e., the total current entering the N_J junction elements is equal to I_J . The attachment mode (26) is easily implemented in the MOM procedure, because it is a linear combination of two regular basis functions (24) already used to expand the non-junction part of the current.

D. Probe current expansion

The coaxial probe is modeled by a number of line segments. The end points of the n th probe segment, C_n , are assigned the local indices i and j , which assume the values 1 and 2 in a cyclic manner. The coordinates of these points are given by $z_i^{(n)}$, where $i = 1, 2$, and the length of C_n is denoted by $h^{(n)}$. The axial current and its derivative on C_n are represented as

$$I^{(n)}(z) = I_i^{(n)} N_i + I_j^{(n)} N_j, \quad \frac{d}{dz} I^{(n)}(z) = \frac{I_j^{(n)} - I_i^{(n)}}{h^{(n)}} \quad (27)$$

where $I_i^{(n)}$ is the total z -directed current at the i th node of the element, and N_i and N_j are linear shape functions [26, p. 99] given as

$$N_i(s) = \frac{1-s}{2}, \quad N_j(s) = N_i(-s), \quad -1 < s < +1 \quad (28)$$

Here, s is a local coordinate, which varies from -1 to $+1$ on a segment, as illustrated in Fig. 7. An arbitrary point within C_n can be defined by $z = z_i^{(n)} + h^{(n)} N_j(s)$.

E. Global MOM matrix assembly and solution procedure

When the expansions (23) and (27) are substituted into the integral equations (21)–(22), the coefficients $I_i^{(n)}$ are constrained by the boundary conditions, which require the continuity of the normal components of $\mathbf{J}_s$ across the edges shared by adjacent elements, or their vanishing at the boundary edges of S , as well as the continuity of the coax probe current between adjacent segments. On the uppermost probe segment, $I_j^{(n)}$ is set to I_J , which explicitly

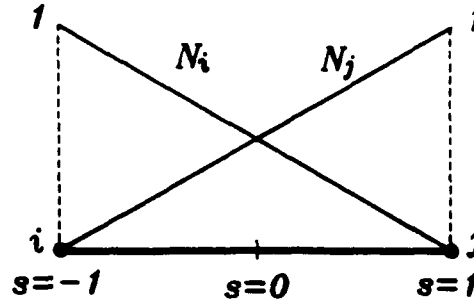


Figure 7: Linear shape functions on a line segment element.

enforces the continuity of the attachment mode current at the probe-to-patch junction. Of course, I_j is not known and must be computed together with the other current expansion coefficients. On the probe segment adjacent to the ground plane, $I_i^{(n)}$ is equal to the input current I_{IN} used in (17).

As a result of this procedure, the integral equations (21)–(22) are converted into a matrix equation for the current expansion coefficients, which may be solved by standard procedures. Once the coefficients $I_i^{(n)}$ are known, Z_{IN} follows from (17) and the far fields are computed from (18)–(20). The complex resonant frequencies, which we express as

$$f = f_r \left(1 + \frac{j}{2Q} \right) \quad (29)$$

where f_r is the real part of f and Q is the quality factor (or Q -factor), may be found as zeros of the matrix determinant, by means of the Müller's method [27, p. 136].

For the solution procedure described above to be practical, the spectral integrals (3) which occur in (5)–(9) and in (15)–(16) must be efficiently evaluated. Hence, these integrals are accelerated by asymptotic integrand subtraction and the method of averages [28], [19]. In addition to these techniques, an interpolation and table look-up scheme is implemented to further reduce the computation time [28], [19]. The integration path must be prop-

erly deformed to avoid the integrand singularities which occur on or near the real axis in the k_p plane [12].

The matrix elements comprise integrals over the triangular patch subdomains and over the probe segments. The former are evaluated by a numerical quadrature especially developed for triangular domains [25, p. 113], and the latter by a Gauss quadrature. The singular parts of the kernels are extracted and integrated analytically [29].

When the current expansions (23) and (26) are substituted for J_s in the far field expressions (18)–(19), the resulting integrals may be evaluated in closed form [30], [31]. Similarly, when I in (20) is represented by (27), the resulting integrals may be evaluated analytically.

4 Sample Results

In this section we present sample computed and measured results for a pentagonal patch antenna (see Fig. 1) with the dimensions $s = 57.3$ mm, $q = 43.9$ mm, and $p = 38.6$ mm ($q/s = 0.766$, $p/q = 0.879$), which have been scaled down from an antenna investigated by Carver and Coffey [2]. The substrate is isotropic, with $\epsilon_t = \epsilon_z = 2.43$, loss tangent $\tan \delta = 6 \times 10^{-4}$, and thickness $h = 1.516$ mm (see Fig. 2). The cover layer is absent. The coax radii are $a = 0.635$ mm and $b = 2.095$ mm.

In Figs. 8 and 9 are shown vector plots of the computed patch currents of the first two source-free modes of the antenna, which resonate at the frequencies 1.615 GHz and 1.721 GHz with the Q -factors 84.2 and 72.1, respectively. These figures also illustrate the triangular mesh used to model the antenna in the numerical procedure. This mesh consists of 95 nodes, 153 elements, and 247 edges, 35 of which are on the boundary of the patch. Thus, the size of the MOM matrix is 212×212 . Only $\Re\{J_s\}$ is plotted in Figs. 8 and

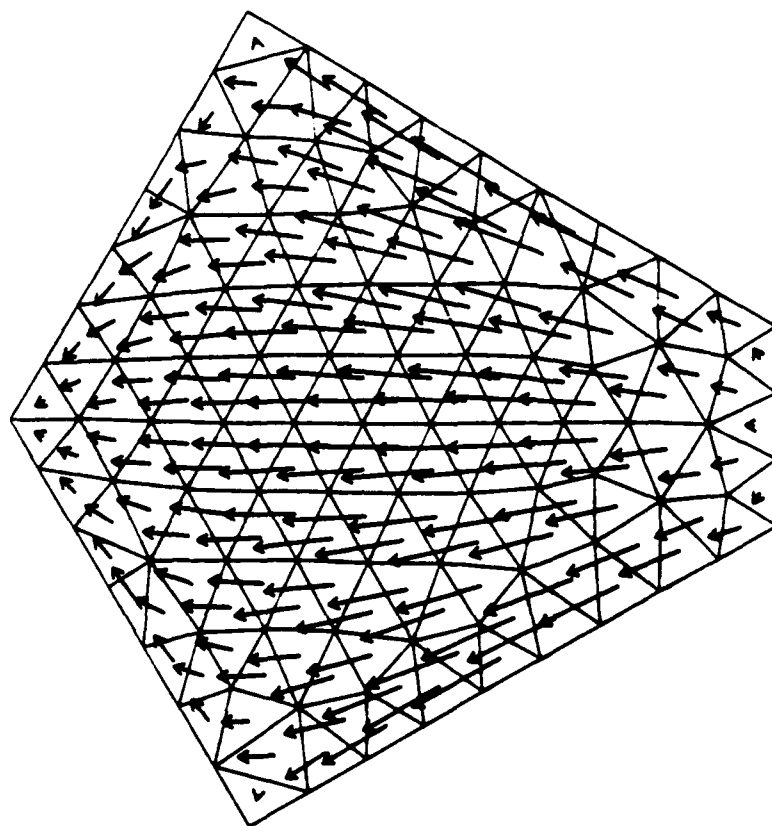


Figure 8: Vector plot of the patch current associated with the first resonant mode ($f_r = 1.615$ GHz, $Q = 84.2$) of the pentagonal microstrip antenna.

9. because $\Im\{J_s\}$ was found to be negligible. We note that the two modal currents appear to be orthogonal to one another.

In Fig. 10 we compare on a Smith chart the plots of the computed and measured input impedance vs. frequency for the coax-fed pentagonal antenna. The triangle element model of the patch was that shown in Figs. 8 and 9. Because the substrate was electrically thin in the frequency range considered, only two segments were used to model the coax probe. The agreement between the numerical and experimental results in Fig. 10 is only fair. We conjecture that the observed discrepancy may be attributed to the simplifying assumptions made in the coax feed modeling, which may be violated in

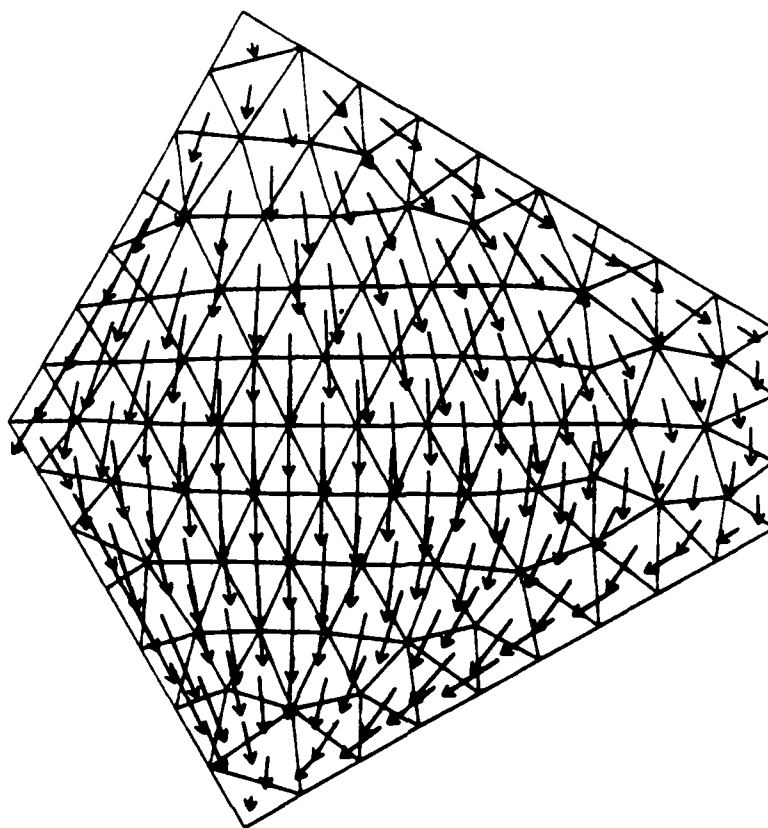


Figure 9: Vector plot of the patch current associated with the second resonant mode ($f_r = 1.721$ GHz, $Q = 72.1$) of the pentagonal microstrip antenna.

the case of an edge-driven patch. We note the presence of the characteristic cusp in these plots, which is indicative of a coupling of two modes with close resonant frequencies. In Figs. 11a and 11b we show vector plots of $\Re\{J_s\}$ and $\Im\{J_s\}$ of this antenna at the 'cusp frequency' $f = 1.71$ GHz. We note that near the coax probe the imaginary part of the patch current density is directed radially away from the feed point, as expected. On the remainder of the patch, the real and imaginary parts of the current are nearly orthogonal to each other and, with maximum current densities of 5.532 A/m and 5.812 A/m, respectively, are close in magnitude. We therefore expect this antenna to radiate a circularly polarized field at this frequency. This is indeed

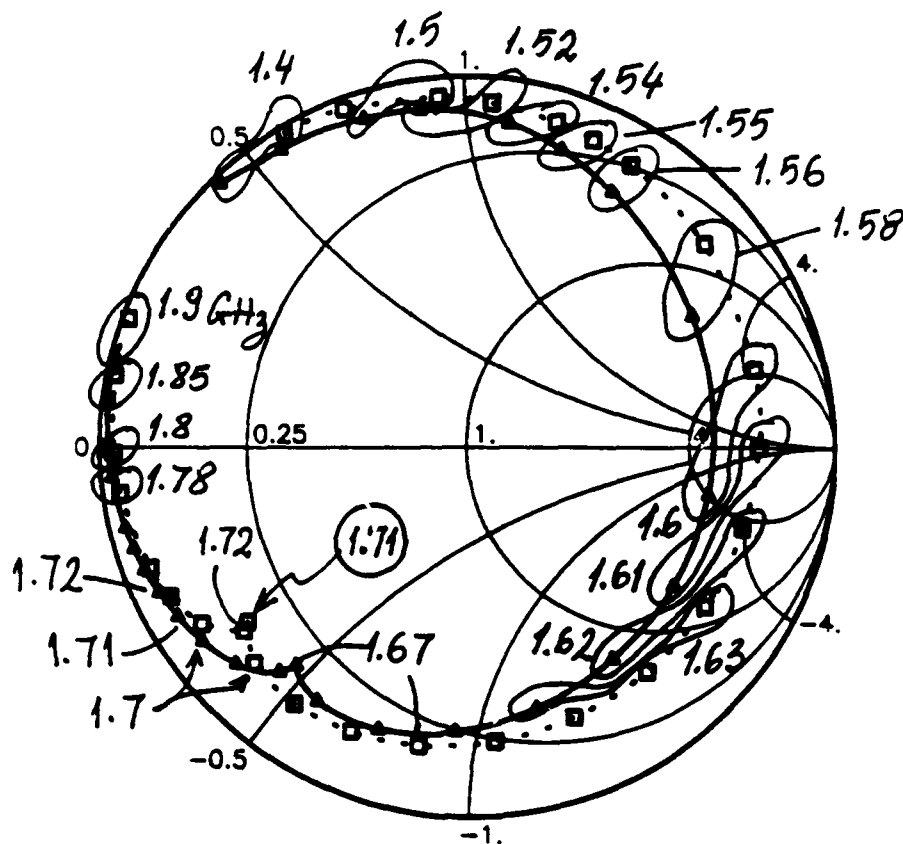


Figure 10: Smith chart plot of the computed (squares) and measured (triangles) input impedance of the coax-fed pentagonal microstrip patch antenna.

observed in Figs. 12a and 12b, which show the corresponding computed far field patterns in the xz ($\varphi = 0^\circ$) and yz ($\varphi = 90^\circ$) planes, respectively. The circular polarization can only be achieved in a narrow frequency band, and the pattern deteriorates rapidly away from $f = 1.71$ GHz.

Finally, in Fig. 13 we show the contribution of the coax probe to the far field pattern in the $\varphi = 0^\circ$ plane at the frequency $f = 1.71$ GHz. The probe radiates only the θ -component of the electric field, which is φ -independent. We note that its magnitude is at least 20 dB below that of the patch field,

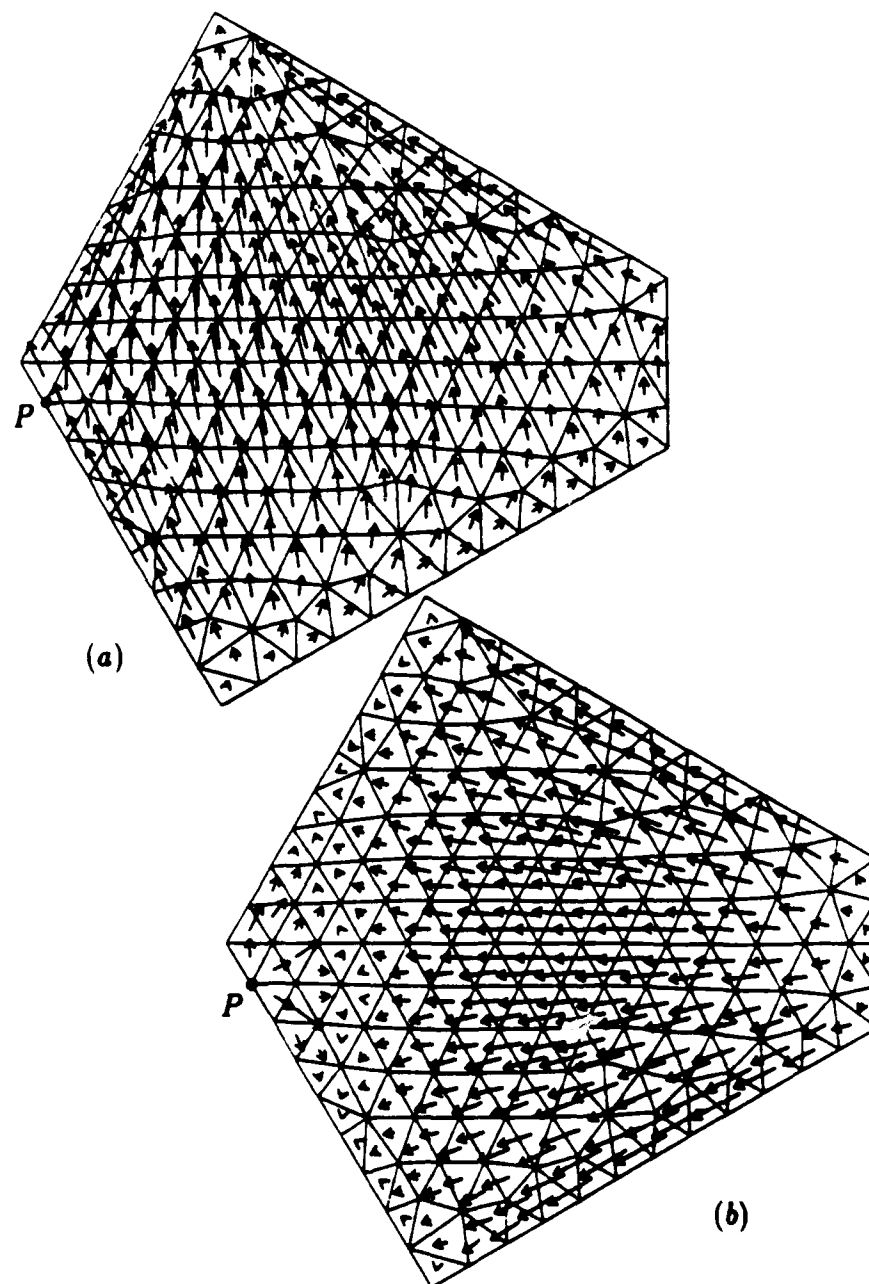
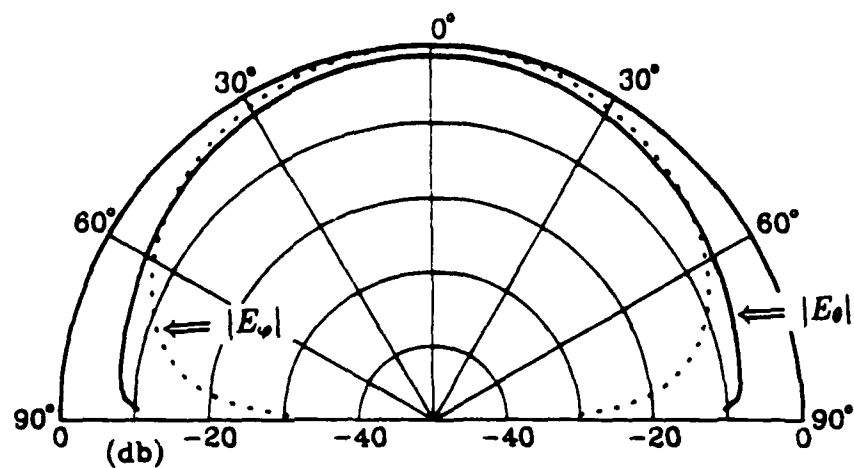
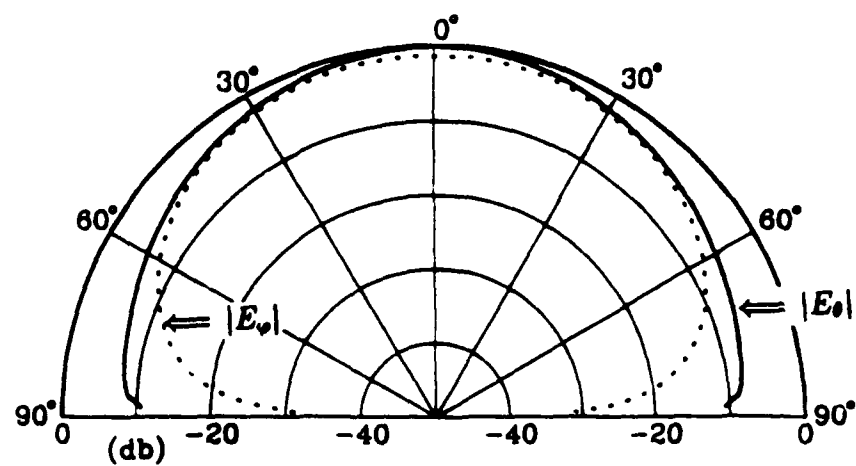


Figure 11: Vector plots of (a) the real part and (b) the imaginary part of the patch current on the coax-fed pentagonal microstrip antenna at $f = 1.71$ GHz.



(a)



(b)

Figure 12: Far field patterns in (a) the $\varphi = 0^\circ$ plane and (b) the $\varphi = 90^\circ$ plane of the coax-fed pentagonal microstrip patch antenna at $f = 1.71$ GHz.

and its contribution to the total field is negligible.

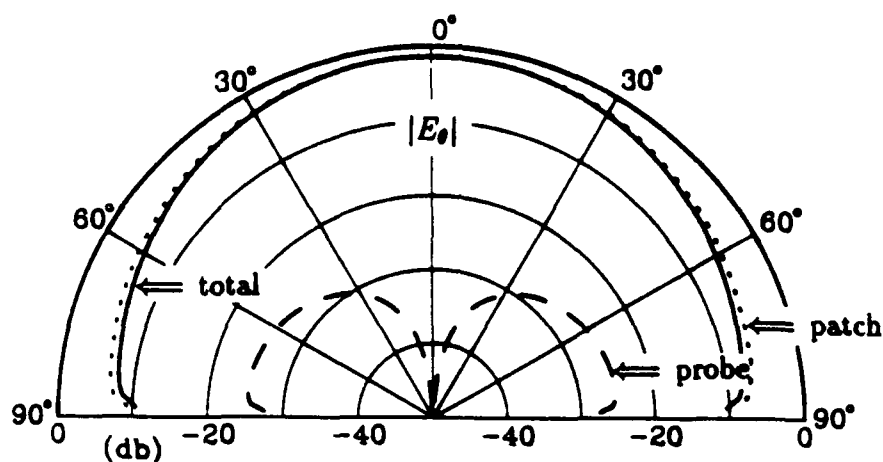


Figure 13: Far field pattern in the $\varphi = 0^\circ$ plane of the coax-fed pentagonal microstrip patch antenna at $f = 1.71$ GHz showing the contribution of the coax probe (long dash line).

5 Conclusion

We have developed an elegant and efficient integral equation approach for the analysis of coax-fed microstrip patch antennas of arbitrary shape. The method has been applied to investigate a pentagonal patch antenna, designed to radiate a circularly polarized field when excited by a single coax feed. This antenna configuration is not amenable to any other presently available rigorous analysis technique. With the approach presented here, microstrip patch antennas of various, possibly irregular shapes may be analyzed within a single theoretical framework, using the same computer code.

References

- [1] H. D. Weinschel, "A cylindrical array of circularly polarized microstrip antenna," in *Digest IEEE AP-S Int. Symp.*, pp. 177-180, 1975.
- [2] K. R. Carver and E. L. Coffey, "Theoretical investigation of the microstrip antenna," Tech. Rept. PT-00929, Phys. Sci. Lab., New Mexico State Univ., Jan. 1979.
- [3] W. F. Richards, Y. T. Lo, and P. Simon, "Design and theory of circularly polarized microstrip antennas," in *Digest IEEE AP-S Int. Symp.*, pp. 117-120, 1979.
- [4] Y. Suzuki and T. Chiba, "Computer analysis method for arbitrarily shaped microstrip antenna with multiterminals," *IEEE Trans. Antennas Propagat.*, vol. AP-32, pp. 585-590, June 1984.
- [5] M. Haneishi and Y. Suzuki, "Circular polarization and bandwidth," in *Handbook of Microstrip Antennas* (J. R. James and P. S. Hall, eds.), pp. 219-274, London: Peter Peregrinus, 1989.
- [6] P. Pichon, J. R. Mosig, and A. Papiernik, "Input impedance of arbitrarily shaped microstrip antennas," *Electron. Lett.*, vol. 24, pp. 1214-1215, Sept. 1988.
- [7] M. J. Notter and C. G. Parini, "Modelling of arbitrary shaped M/S patch antennas in multiple substrate media," *Electron. Lett.*, vol. 26, pp. 50-52, 1990.
- [8] D. Zheng and K. A. Michalski, "Analysis of arbitrarily shaped coax-fed microstrip antennas—A hybrid mixed-potential integral equation approach," *Microwave & Opt. Technol. Lett.*, vol. 3, pp. 200-203, June 1990.

- [9] D. Zheng and K. A. Michalski, "Analysis of arbitrarily shaped coax-fed microstrip antennas with thick substrates," *Electron. Lett.*, vol. 26, no. 12, pp. 794-795, 1990.
- [10] D. Zheng and K. A. Michalski, "Analysis of coaxially fed microstrip antennas of arbitrary shape with thick substrates," *J. Electromagn. Waves Appl.*, vol. 5, no. 12, pp. 1303-1327, 1991.
- [11] K.-L. Wu, J. Litva, R. Fralich, and C. Wu, "Full-wave analysis of arbitrarily shaped line-fed microstrip antennas using triangular finite-element method," *IEE Proc., Pt. H*, vol. 138, pp. 421-428, Oct. 1991.
- [12] K. A. Michalski and D. Zheng, "Analysis of microstrip resonators of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 112-119, Jan. 1992.
- [13] T. K. Sarkar, P. Midya, Z. A. Maricevic, M. Kahrizi, S. M. Rao, and A. R. Djordjevic, "Analysis of arbitrarily shaped microstrip patch antennas using the Sommerfeld formulation," *Int. J. Microwave Millimeter-Wave Computer-Aided Eng.*, vol. 2, no. 3, pp. 168-178, 1992.
- [14] F. A. Monferrer, A. A. Kishk, and A. W. Glisson, "Green's function analysis of planar circuits in a two-layer grounded medium," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 690-696, June 1992.
- [15] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 409-418, May 1982.
- [16] R. F. Harrington, *Field Computation by Moment Methods*. New York: Macmillan, 1968. Reprinted by Krieger Publishing Co., Melbourne, FL, 1982.

- [17] R. C. Hall and J. R. Mosig, "The analysis of coaxially fed microstrip antennas with electrically thick substrates," *Electromagn.*, vol. 9, no. 4, pp. 367-384, 1989.
- [18] R. F. Harrington, *Time-Harmonic Electromagnetic Field*. New York: McGraw-Hill, 1961.
- [19] J. R. Mosig, R. C. Hall, and F. E. Gardiol, "Numerical analysis of microstrip patch antennas," in *Handbook of Microstrip Antennas* (J. R. James and P. S. Hall, eds.), pp. 391-453, London: Peter Peregrinus, 1989.
- [20] K. A. Michalski, "The mixed-potential electric field integral equation for objects in layered media," *Arch. Elek. Übertragung.*, vol. 39, pp. 317-322, Sept.-Oct. 1985.
- [21] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, N.J.: Prentice Hall, 1973.
- [22] M. Davidovitz and Y. T. Lo, "Input impedance of a probe-fed circular microstrip antenna with thick substrate," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 905-911, July 1986.
- [23] W. C. Chew, Z. Nie, and Y. T. Lo, "The effect of feed on the input impedance of a microstrip antenna," *Microwave & Opt. Technol. Lett.*, vol. 3, pp. 79-83, Mar. 1990.
- [24] W. C. Chew, "A quick way to approximate a Sommerfeld-Weyl-type integral," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 1654-1657, Nov. 1988.
- [25] L. Lapidus and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*. New York: Wiley, 1982.

- [26] O. C. Zienkiewicz and K. Morgan, *Finite Elements and Approximation*. New York: Wiley, 1983.
- [27] D. M. Young and R. T. Gregory, *A Survey of Numerical Mathematics*. New York: Dover, 1988.
- [28] J. R. Mosig, "Integral equation technique," in *Numerical Techniques for Microwave and Millimeter-Wave Passive Structures* (T. Itoh, ed.), pp. 133-213, New York: Wiley, 1989.
- [29] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, vol. AP-32, pp. 276-281, Mar. 1984.
- [30] B. Houshmand, W. C. Chew, and S. W. Lee, "Fourier transform of a linear distribution with triangular support and its applications in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 252-254, Feb. 1991.
- [31] K. McInturf and P. S. Simon, "The Fourier transform of linearly varying functions with polygonal support," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 1441-1443, Sept. 1991.

PLANAR SPIRAL, A MICROSTRIP ANTENNA?

Sam C. Kuo

GTE Government Systems, ED

Mountain View, CA 94039

1. INTRODUCTION

The planar spiral antenna is a good candidate for applications where extremely broadband, light weight, relatively small size, and circular polarization are required. The two basic types of planar spirals are the Archimedian and the equiangular spirals. The Archimedian spiral is a broadband antenna that has been successfully used to cover frequency bands of more than three octaves. The equiangular spiral is a frequency-independent antenna, which theoretically has an unlimited bandwidth. In practice, the highest operating frequency of a planar equiangular spiral is limited by the details of the feed point, and the lowest operating frequency is limited by the size of the aperture. Planar spiral antennas radiate bidirectionally; and, when operating in this manner, they provide relatively frequency-independent performance over a large frequency range. Unfortunately, the bi-directional radiation is undesirable for most applications. For this reason, the most common configuration of the planar spiral antenna is the cavity-backed spiral, where the spiral radiator is mounted at the opening of a circular cavity for unidirectional radiation. Unidirectional radiation can also be obtained by simply placing the spiral radiator over a ground plane and supporting it at the edge. The presence of the cavity or the ground plane, however, greatly reduces the useful bandwidth of planar spiral antennas for most applications because of the degradation of the radiation

patterns at higher operating frequencies. This degradation is caused primarily by the excitation of higher modes. The introduction of absorber or lossy materials between the planar spiral and the metal surface reduces the magnitudes of these higher-order modes. If the cavity is made very deep and filled completely with absorber, the excitation of the higher-mode radiation can be eliminated almost entirely. Unfortunately, the addition of absorber or lossy material decreases the efficiency of the cavity-backed spiral. Over the years, antenna engineers have been searching for a cavity design that will preserve the broadband performance of the planar spiral without introducing excessive loss.

With the advent of microstrip antennas in the 1970s, the interest in low-profile antennas thrived. The disadvantage of the microstrip antenna has always been its limited bandwidth coverage. The planar spiral antenna, being flat and broadband, appears to be a good candidate for low-profile applications. If the planar spiral could be made to operate in the low-profile configuration, it would make an ideal broadband microstrip antenna.

2. DISCUSSION

When a two-arm planar spiral antenna is excited in the conventional manner, equal amplitude and out-of-phase currents are used to feed the spiral from the inside terminals. Near the feed point, the currents on adjacent spiral arms are out of phase; therefore, no radiation takes place. The excitation currents travel along the spiral arms until the currents on adjacent arms are almost in-phase. This is where radiation takes place. For a properly designed Archimedian or equiangular spiral, the region where radiation starts to take place is very close to a ring on the spiral aperture whose circumference is slightly less than one wavelength at the operating

frequency. As the currents continue to travel along the spiral, radiation continues until they have all been radiated. When a planar spiral is operating in free space or with a lossy cavity, most of the excitation currents are radiated in this region, which is called the "one-wavelength mode" active region. Radiation from this region has the maximum power on the antenna axis, and the far-field phase has a phase variation of 360 degrees with a 360 degree rotation about the antenna axis. The radiation patterns are axisymmetric. As long as the spiral aperture is large enough for most of the currents to be radiated efficiently, the gain and beamwidth of the radiation patterns are very constant as a function of operating frequency. If the spiral aperture is small in terms of wavelength, not all of the excitation currents will be radiated before they reach the end of spiral arms. These residual currents are reflected back by the spiral arm ends and radiate in the opposite sense of circular polarization. This usually occurs at the lower frequencies, where the radiating patterns are broader and have higher axial ratios and lower gain because of the smaller active region.

When a spiral is placed over a ground plane or mounted at the opening of a lossless cavity, the presence of the ground plane or the cavity bottom decreases the radiation efficiency of the spiral radiator. Consequently, not all of the excitation currents are radiated in the "one-wavelength mode" active region and the residual currents travel along the spiral arms until the conditions are favorable for radiation to occur again. This usually takes place near a ring on the spiral aperture whose circumference is approximately three times the wavelength at the operating frequency. This radiation is referred to as the "three-wavelength mode." The "three-wavelength mode" radiation patterns are also axisymmetric; however, the

far-field phase has a 1080-degree variation with a 360-degree rotation about the antenna axis. Although both the "one-wavelength mode" and "three-wavelength mode" radiation patterns are symmetrical about the antenna axis, the composite patterns are no longer axisymmetric. Because the far-field phase of the "three-wavelength mode" radiation patterns have a phase variation three times that of the one-wavelength mode, the amplitudes of the two patterns will add in the plane where the two patterns are in phase. In the orthogonal plane, however, the two patterns are out of phase, and the amplitudes will subtract. The composite pattern is, thus, a fan-shaped beam. The ratio of the maximum to the minimum width of the fan-shaped beam depends on the magnitude of the "three-wavelength mode" radiation, which is inversely proportional to the spacing between the spiral and the ground plane. Since the distance between the two active regions varies as a function of frequency, the fan-shaped beam rotates about the antenna axis as the frequency changes. This rotation causes the azimuthal beamwidth to vary as a function of frequency. Since the biggest market for the cavity-backed spiral is airborne early warning systems, the variation of the beamwidth as a function of the frequency is very undesirable.

When this fan shaped beam was first discovered by the spiral-cavity designers, the degradation was thought to be caused by the cavity wall. However, it was found that removing the cavity wall (by simply placing the planar spiral over a finite ground plane, or covering the cavity wall with absorbent material) failed to eliminate the pattern degradation at the higher frequencies. Therefore, it could be concluded that the major contributor to the pattern degradation had to be the bottom of the cavity, or the ground plane. The performance of a planar spiral over a finite

flat, circular ground plane was studied by many companies in the industry in late 1960s and early 70s. Unfortunately the results were rarely published in the open literature because no one wanted to share information with their competitors.

Although the spiral antenna provides better performance without a cavity wall, most antenna designers still prefer the cavity configuration because the cavity, spiral radiator, and the balun are a mechanically integral unit. Designing a broadband spiral cavity was a tedious task involving endless cutting and trying. The earlier cavities were shallow because the designers wanted to use the bottom of the cavity as a reflector. However, because of the need for a transformer between the 50-ohm unbalanced input to the balanced, high-impedance output at the spiral input terminals, cavities were made deeper. This allowed a longer transformer to be used for better impedance transformation. The deeper cavity allows for more absorber to be used in the cavity to preserve the free-space-pattern performance of the planar spirals. Unfortunately the gains of these antennas are generally very low. If the absorber is kept at a small distance from the spiral radiator, the loss in gain can be kept at approximately 3 dB. Thus, the design of the spiral cavity is generally a trade-off of gain versus pattern quality (pattern symmetry and axial ratio).

Since a planar spiral using a shallow cavity or a small spacing above a ground plane will have this pattern degradation at higher frequencies, the microstrip-mode spiral antenna must have the same problem.

3. VERIFICATION

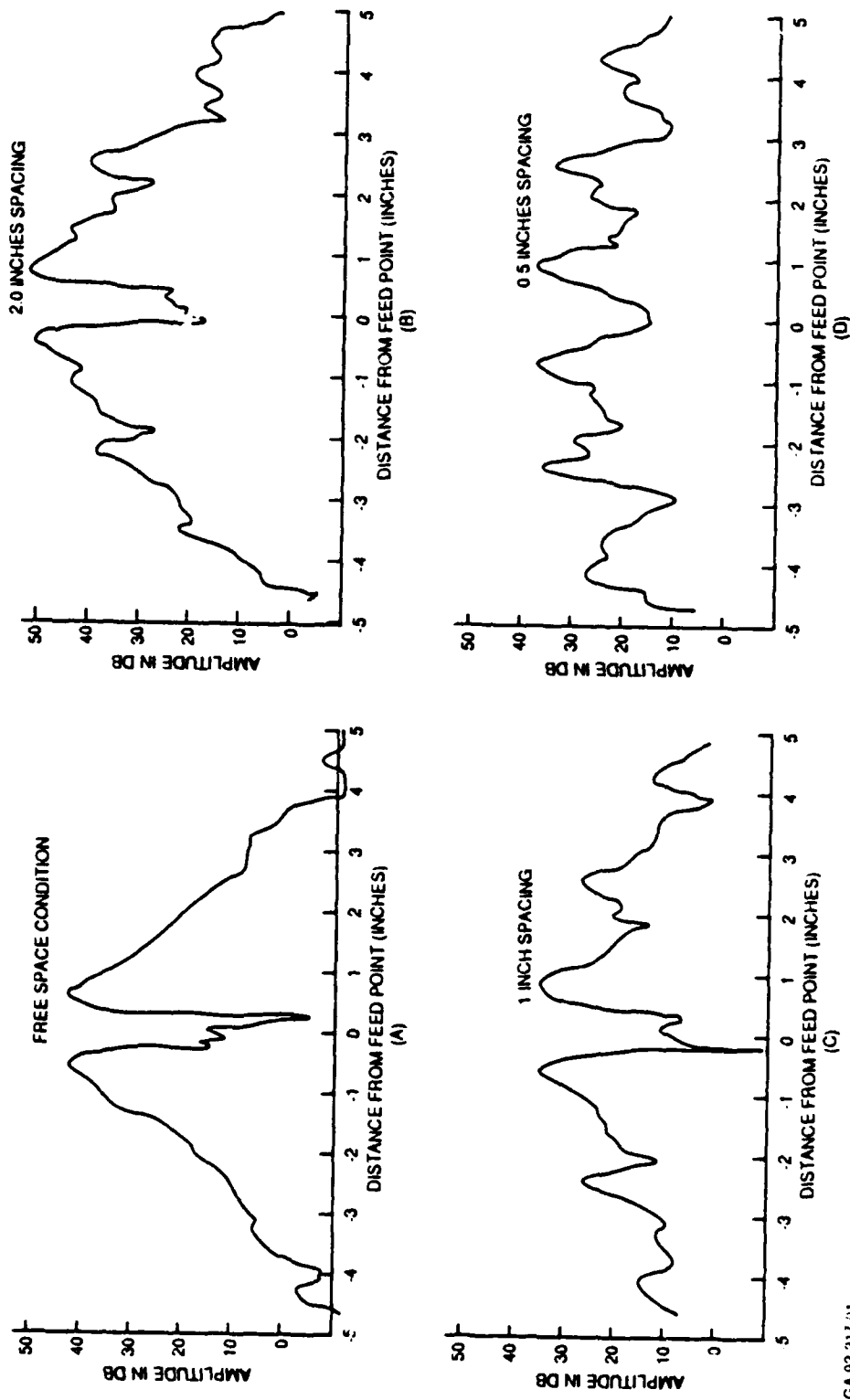
In order to demonstrate that the microstrip-mode spiral antenna will have pattern degradation at higher operating frequencies, three tests were performed:

near-field current probing, transmission-line test, and conical-cut pattern measurements. The following are the results of these tests.

3.1 Near-Field Current Probing

In order to demonstrate the performance of the two-arm planar spiral over the ground plane at small spacings as an approximation to "microstrip," the near field currents of a 10-inch planar spiral were probed. This Archimedian spiral was etched on a 0.060-inch G-10 fiberglass PC board. A fiberglass PC board was used because it would provide a rigid, flat surface to maintain a constant spacing between the current probe and the spiral surface. The spiral was supported by several absorber rings (Emerson Cumming LS-24 material) with an inside diameter of 9.5 inches and an outside diameter of 10.5 inches. the spacing recorded on the near field plots was measured from the bottom of the spiral. Thus, the spiral radiator was actually 0.060 inch further above the ground plane.

When the near-field currents were measured, the probe moved on a straight line. It started from one edge, passed over the feed point, and stopped at the other edge of the spiral antenna. A frequency of 2.0 GHz was selected because a 10-inch spiral can support two higher modes at this frequency. The plots of the near-field currents measurement are shown in Figure 1. The x-axis of the plot illustrates the probe position on the spiral (distance from the feed point), and the y-axis shows the relative amplitudes of the near-field currents. The near field of this antenna was first probed with the spiral operating in the free-space condition; this plot is shown in Figure 1-a. It can be observed that the amplitudes of currents in the near field start to build as the currents on the spiral travel from the feed point toward the edge and reach maximum amplitudes at approximately 0.65 inch from



GA 92 217 01

Figure 1. Near Field Amplitude of a 10-Inch Archimedian Spiral at 2.0 GHz

the feed point. This corresponds to a ring of approximately 0.7λ in circumference at 2.0 GHz. It should be pointed out that the maximum radiation does not take place at this point but, rather, where the near field currents start to decay. Radiation continues as the currents travel along the spiral arms until most of the excitation currents are radiated. The continuously decreasing amplitudes suggest that no higher modes were excited.

It was discussed previously that, when a planar spiral is placed over a ground plane, the radiation efficiency of the antenna will decrease as the spacing is decreased. Hence, the smaller the spacing, the more excitation currents will penetrate the "one-wavelength mode" active region. This effect is demonstrated in Figures 1-b through 1-d, which show the near-field currents of the 10-inch spiral as the ground-to-spiral spacing is varied.

Figure 1-b shows the behavior of the near-field currents when the spacing is 2 inches. In this plot, a second set of amplitude peaks can be observed. These peaks are approximately 2.4 inches from the feed point, which corresponds to a 2.6λ ring at 2.0 GHz. The second peak demonstrates a second radiation region, which is located around a ring whose circumference is approximately 3λ at 2.0 GHz. The amplitude of the second peak is about 10 units below the first. (It should be realized that all the amplitudes measured here are plotted on a logarithmic scale but represent relative values.) This near-field current plot suggests the presence of the "three-wavelength mode" radiation.

The spacing was decreased to 1 inch, and the near-field currents were again probed. The measurements are shown in Figure 1-c. It can be seen that the position of the second peak is the same as Figure 1-b, but the corresponding

amplitude has increased to approximately 6 units below the first peak. This demonstrates the increase of the level of the " 3λ mode" radiation. When the spacing was further reduced to 0.5 inch, the near field currents were again probed (see Figure 1-d). Here the second peak is only 2 units below the first peak; and, furthermore, a third set of peaks is also observed. The third peak is approximately 4.2 inches from the feed point and corresponds to a 4.5λ circle. This third peak is the evidence of the excitation of the "5-wavelength mode" radiation.

When the test frequency was increased to 3.0 GHz, the near field currents were probed again with a 0.5 inch spacing. The results are shown in Figure 2. A fourth set of peaks can be observed here because a 10-inch spiral will support the seven-wavelength mode at 3.0 GHz.

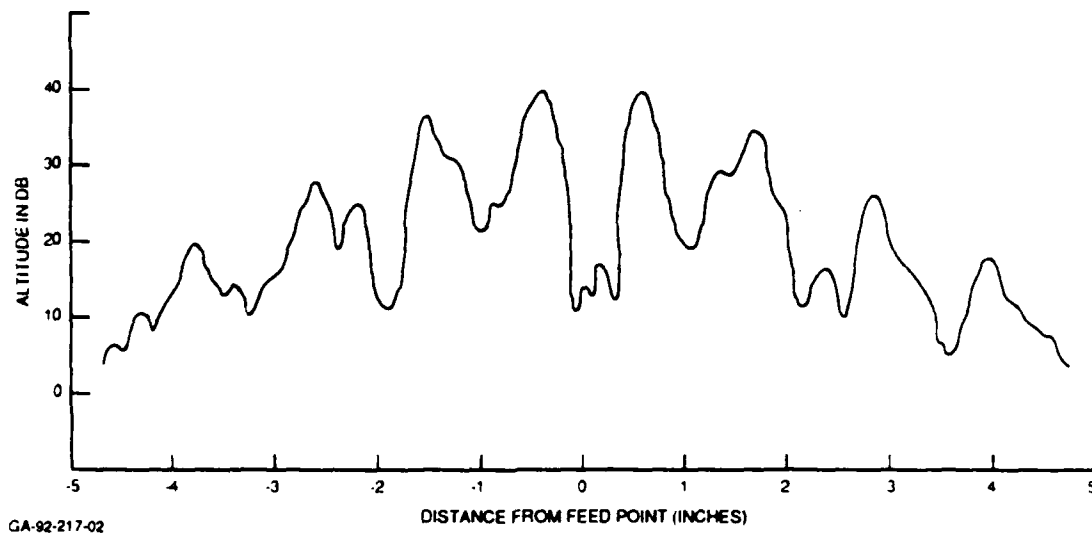


Figure 2. Near Field Amplitude of a 10-Inch Archimedian Spiral at 3.0 GHz

Probing the near-field currents and plotting the resulting data clearly demonstrates the excitation of the higher modes when the spiral antenna is placed over a ground plane. The relative magnitudes of the higher-mode radiations are inversely proportional to the spacings.

3.2 Transmission Line Test

The spiral arms play the roles of both radiator and transmission line in operation. As discussed earlier, the excitation currents will travel on the spiral arms until conditions are favorable for them to radiate. If the spiral aperture is large in terms of the operating wavelength and the external conditions allow the currents to radiate efficiently, a very small amount of current will reach the ends of the spiral arms. In a conventional spiral-antenna design, if the aperture is too small in terms of the operating wavelength, the edge of the spiral is usually terminated in a lossy material to minimize the reflection of the excitation currents, thereby reducing the axial ratio at the lower frequencies. If the spiral-arm ends are terminated into a receiver, the amount of the unradiated current should be measurable.

The test set-up consists of a 3-inch Archimedian spiral etched on 0.030-inch teflon fiberglass PC board, which is excited from both the center and outer terminals. A modified Marchand balun is used at the center terminals and a magic tee at the outer terminals. The VSWR of the spiral antenna measured at the input to the Marchand balun is less than 1.6:1 and, at the outer terminals, 3.5:1. No attempt was made to reduce the VSWR at the outer terminals because this was an unfunded study. A 3-inch circular flat ground plane was placed under the spiral, and various thicknesses of rings were made from LS-24 for changing of the spacings. (The absorber rings have an i.d. and o.d. of 2.8 and 3.2 inches, respectively.) The

experiments were carried out by sweeping the frequency across the frequency range of 2.25 to 3.0 GHz because a 3-inch spiral cannot support higher modes at these frequencies. Thus, the unradiated current will go into the receivers at the outer terminals. The test set-up is shown in Figure 3.

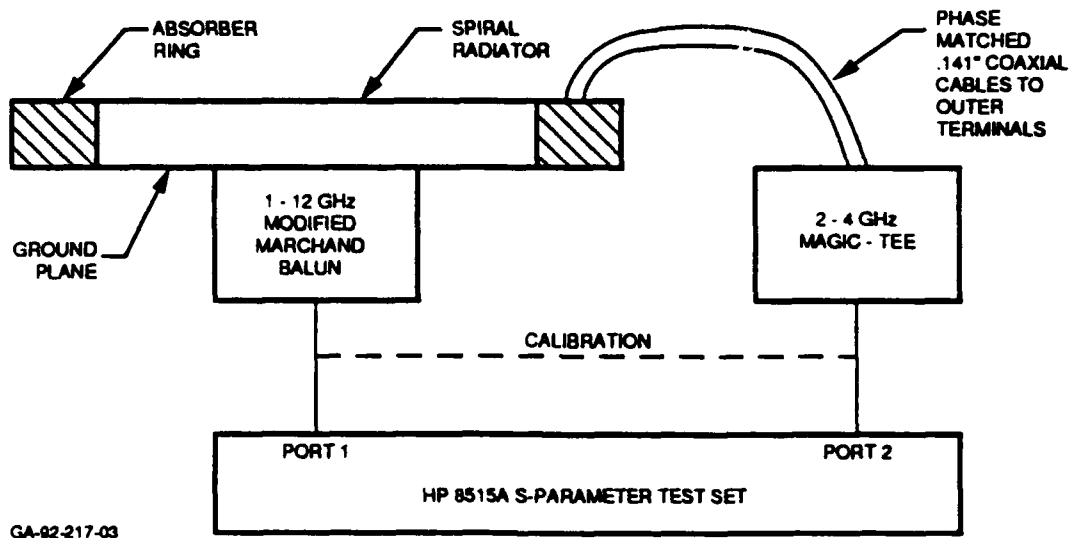


Figure 3. Test Set-Up for the Transmission Test

In this test, the amplitude of the residual current at the output terminals was measured by using the input current as the reference. The system was calibrated without the spiral apparatus in line. Unfortunately, the loss through the balun, the spiral arms, the coaxial cables, and the magic tee were not known. Thus, it was difficult to estimate the amount of the current that actually radiated. The relative magnitude of the residual current measured at the outer terminals, however, provides some idea of the effect of the spacing between the spiral and the ground plane to the radiation efficiency of the spiral antenna.

The output current was first measured with a 3 inch circular disk of Emerson Cummin AN-75 absorber between the spiral and the ground plane. This was to simulate the free space conditions and a 0.1-inch foam spacer was placed between the spiral and the absorber in order to prevent direct contact between the two. When the absorber disk was removed, the output current was again measured at three different spacings, 0.7, 0.3, and 0.1 inch. Each time the absorber ring was used as a spacer. The measured data, shown in Figure 4, shows two things. The first is that less current is radiating at 2.25 GHz than at higher frequencies. This makes sense because the spiral aperture has a smaller active region at the lower frequencies and, therefore, less current is radiating. The second is that the level of the output current will increase when the spacing between the spiral and the ground plane is reduced. If the spiral aperture is large enough to support higher modes, the level of the higher modes will increase as the spacing is decreased. These test results agree with those in the previous near-field current-probing test.

3.3 Conical Cut Radiation Patterns

The axisymmetry of a radiation pattern can be best verified by taking conical-cut measurements. This can be accomplished by pointing the test antenna at some off axis angle, rotating the test antenna and recording the amplitude variations. If the radiation pattern is symmetrical about the antenna axis, a circle is obtained when the the antenna is rotated 360 degrees. For a properly designed spiral antenna excited by a well behaved balun, this occurs when only the one-wavelength mode is excited. When the three-wavelength mode is also excited, an elliptically shaped conical cut pattern will be measured, representing a fan-shaped beam. The amount of ellipticity of the conical-cut pattern indicates the strength of

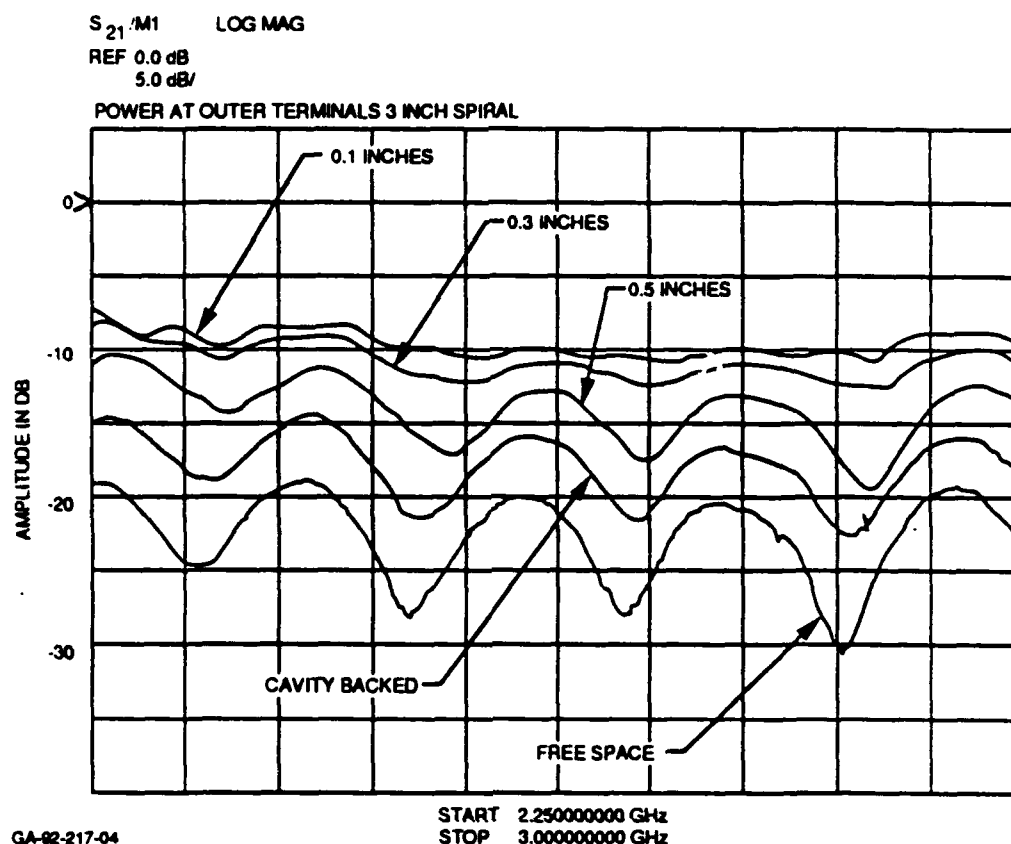


Figure 4. Power Level at the Outer Terminals of a 3-Inch Spiral Versus Spacings

the three-wavelength mode being excited. The ellipticity is also a function of the off-axis angle at which the conical cut is taken. Generally, the ellipticity increases as the off-axis angle is increased. For instance, a conical-cut elliptical pattern measured at 30 degrees off the antenna axis will be more circular than one measured at 80 degrees from the antenna axis. However, a conical-cut pattern measured at 60 degrees off the antenna axis usually provides all the information one needs to evaluate the performance of the spiral. When angles larger than 60 degrees are used, the axial ratio is higher and the signal level is very low for pattern

measurements. For this reason, the conical cut patterns are measured at 60 degrees off the antenna axis for this paper. Archimedian spirals were used for this test. Two spirals, of identical design, with 3 inch diameters were used; one was etched on G10 fiberglass PC board, the other on Teflon fiberglass board. A test frequency of 6 GHz was selected because a 3-inch spiral will support higher mode radiation at this frequency. Conical-cut patterns were measured with the spiral radiator spaced at various distances from a 3-inch circular ground plane. In an attempt to eliminate the ground effect, a 4 inch disk made from AN-75 was used. Absorber rings of the same material as the ones used for the transmission test were also used here to support the spiral radiator. The i.d. of the rings was reduced to 3.6 inches. The pattern was measured for the G-10 spiral with the absorber disk first and is shown in Figure 5. After the disk was removed, the patterns were measured with the spiral at 0.5-, 0.3-, and 0.1-inch spacings. These patterns are shown in Figures 5b through 5d, respectively.

Figure 6 shows the patterns of the Teflon spiral which was measured only at 0.3 inches spacing. The conical pattern of a 3 inch cavity-backed spiral, which has a 0.7-inch deep cavity, was also measured at 60 degrees off the antenna axis and is shown in Figure 7. After the conical patterns were obtained, the azimuthal patterns in two orthogonal planes were measured with the spiral oriented in the manner where the patterns represent the maximum and the minimum beamwidths of the fan-shaped beam. The azimuthal patterns of the various configurations are shown in Figures 8 through 13.

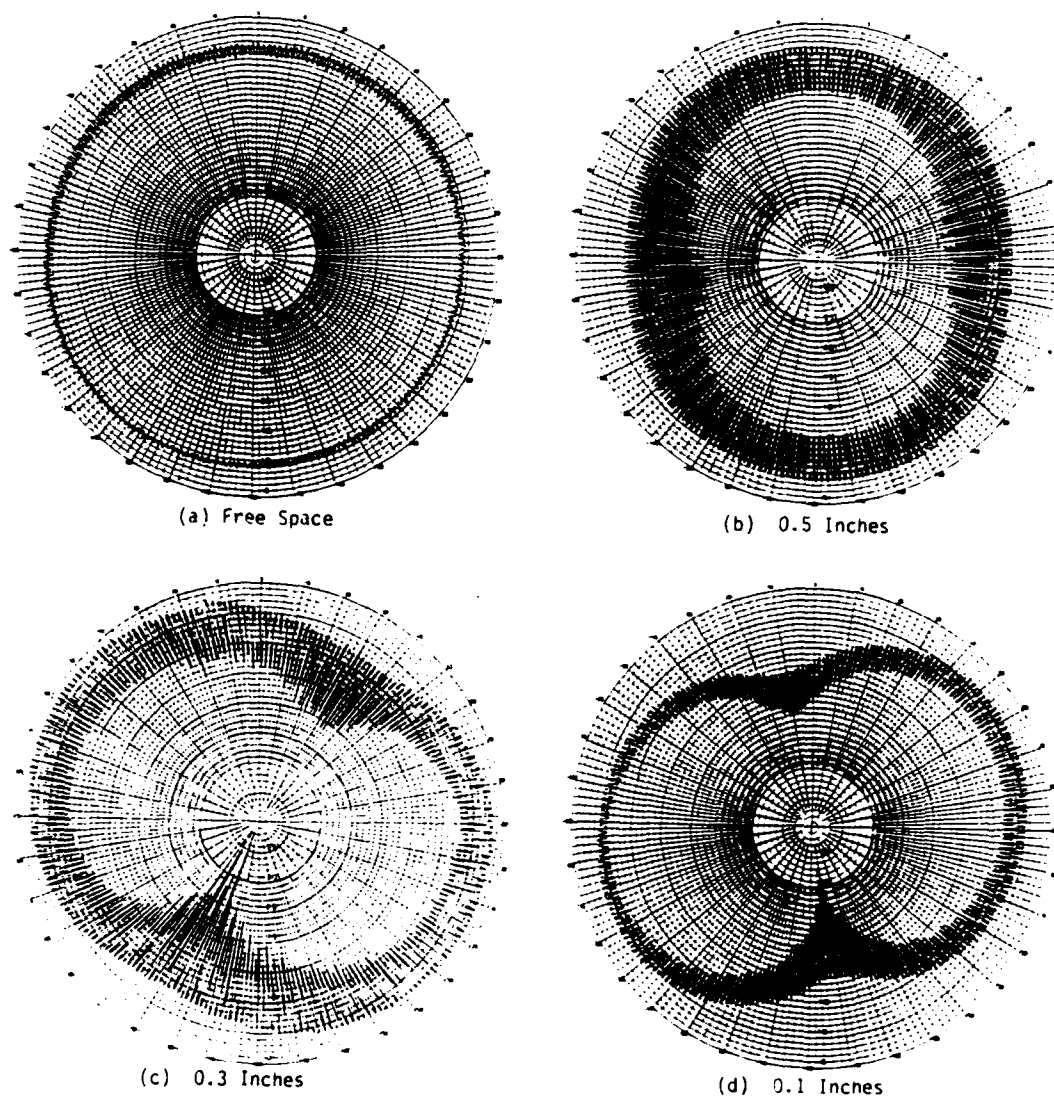


Figure 5. 6 GHz Conical Cut Patterns, 60° Off Antenna Axis. 3-Inch Archimedian Spiral Over a 3-Inch Circular Ground Plane at Various Spacings

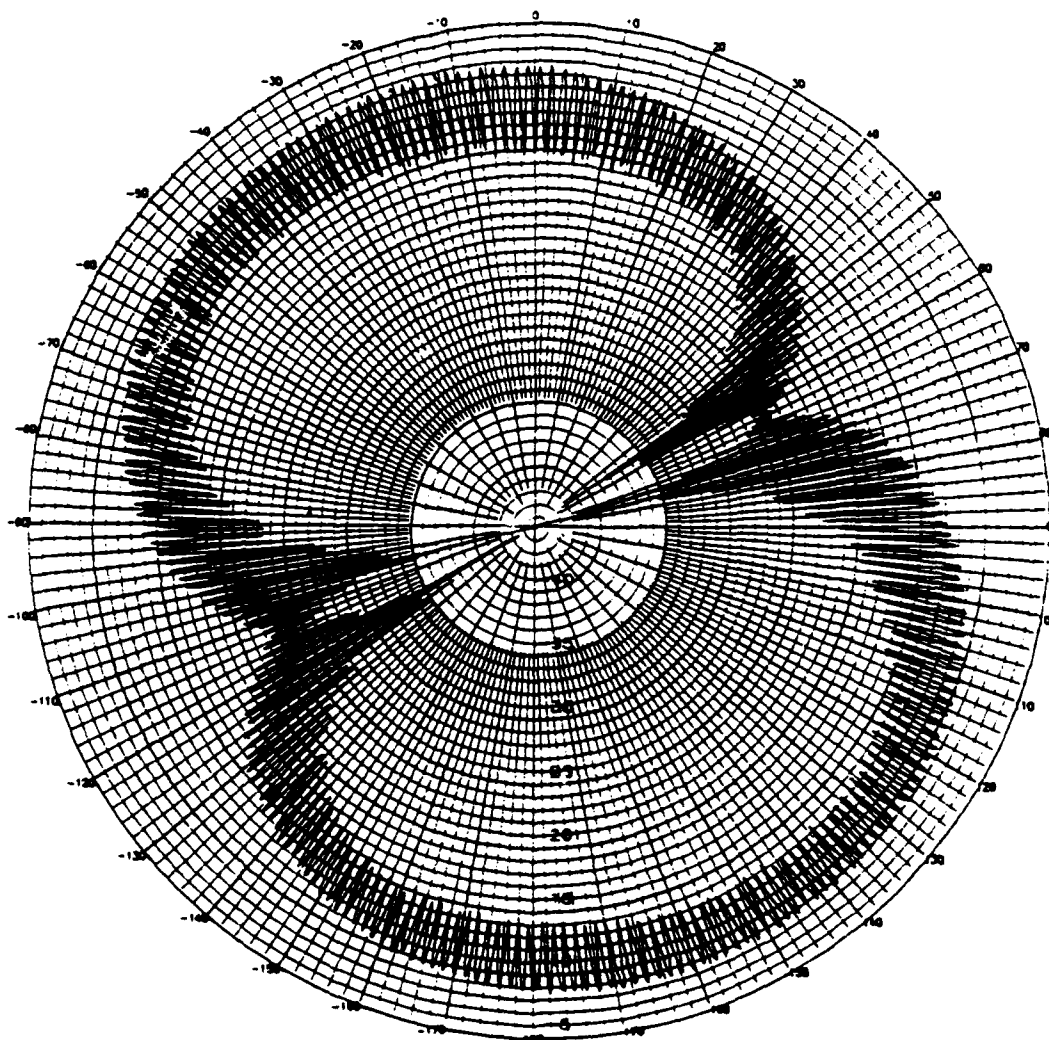


Figure 6. 6 GHz Conical Cut Pattern, 60° Off Antenna Axis. 3-Inch Teflon Spiral Over a 3-Inch Circular Ground Plane at 0.3 Inches Spacing

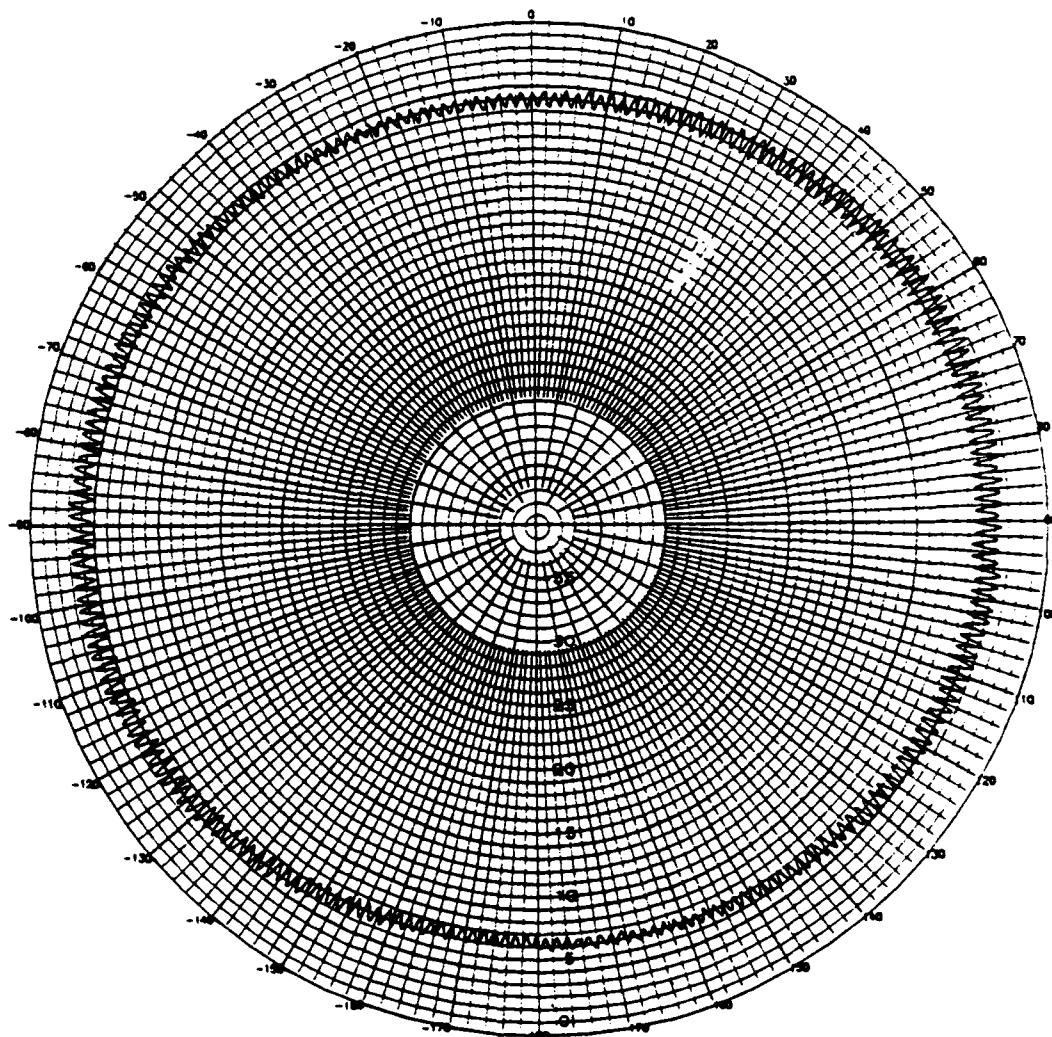
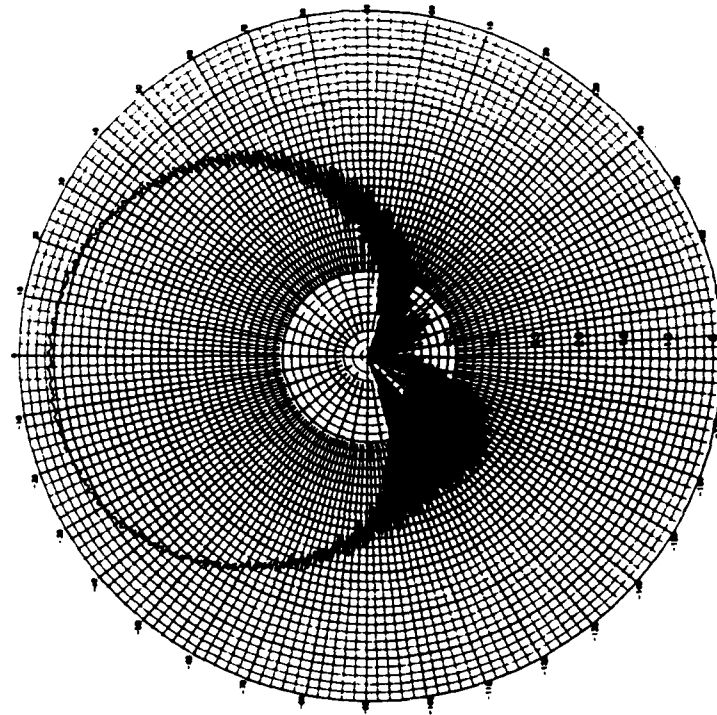
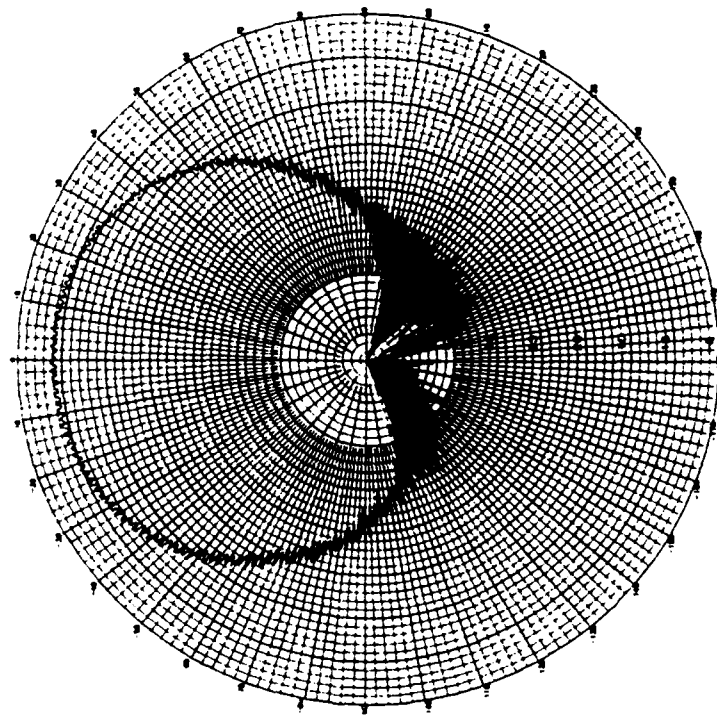


Figure 7. 6 GHz Conical Cut Pattern, 3-Inch Cavity Backed Spiral



Elevation



Azimuth

Figure 8. Radiation Patterns of a 3-Inch Archimedian Spiral in Free Space at 6.0 GHz

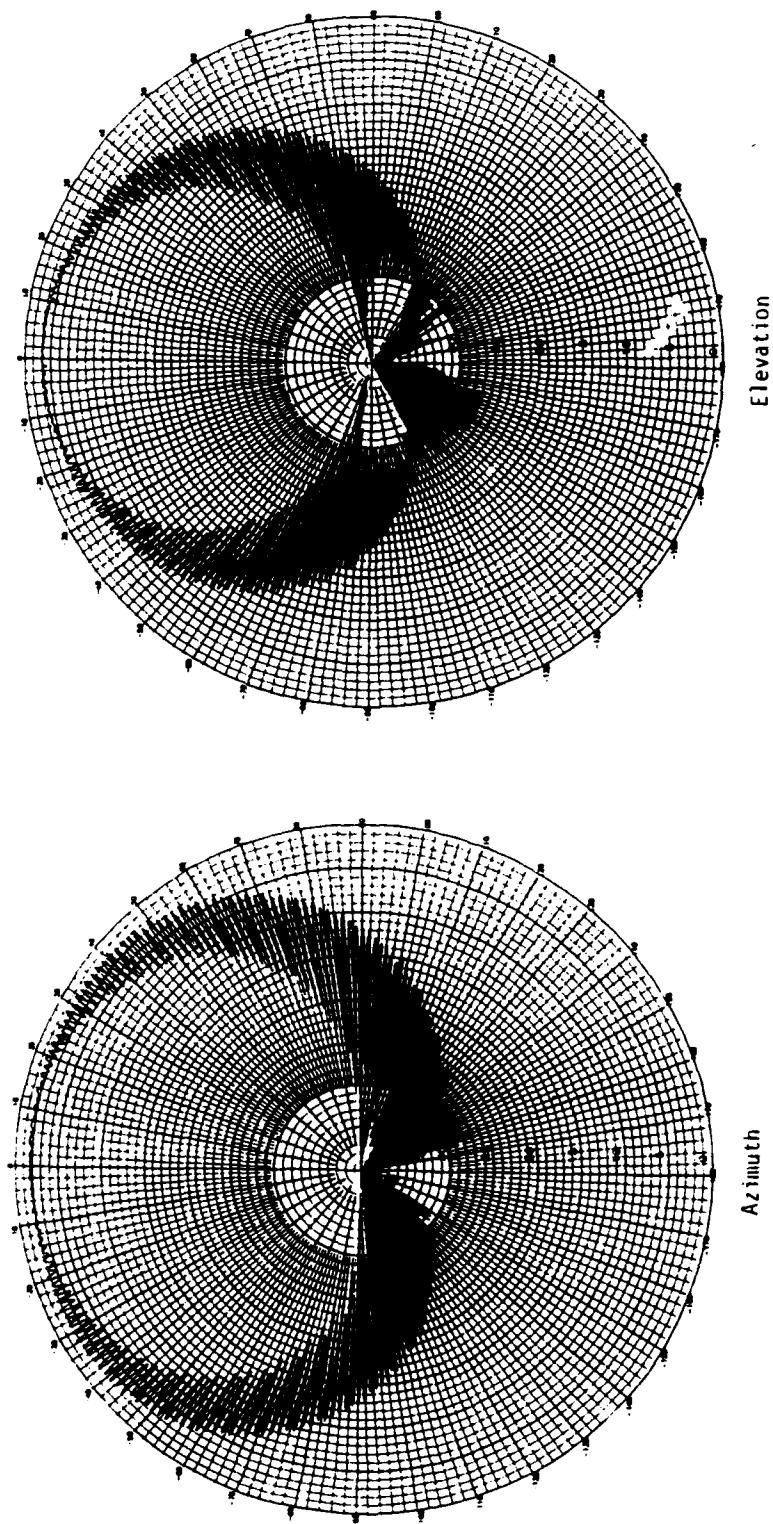
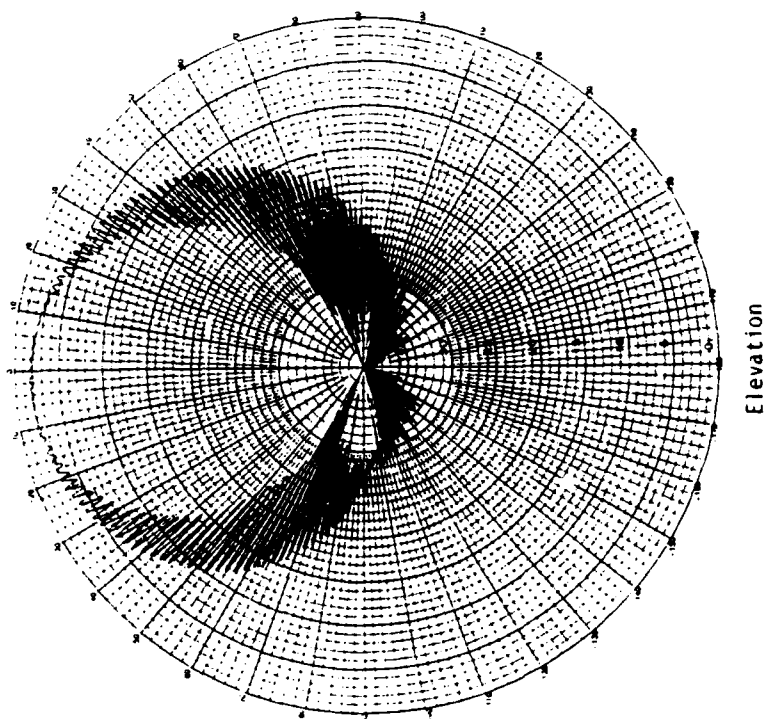
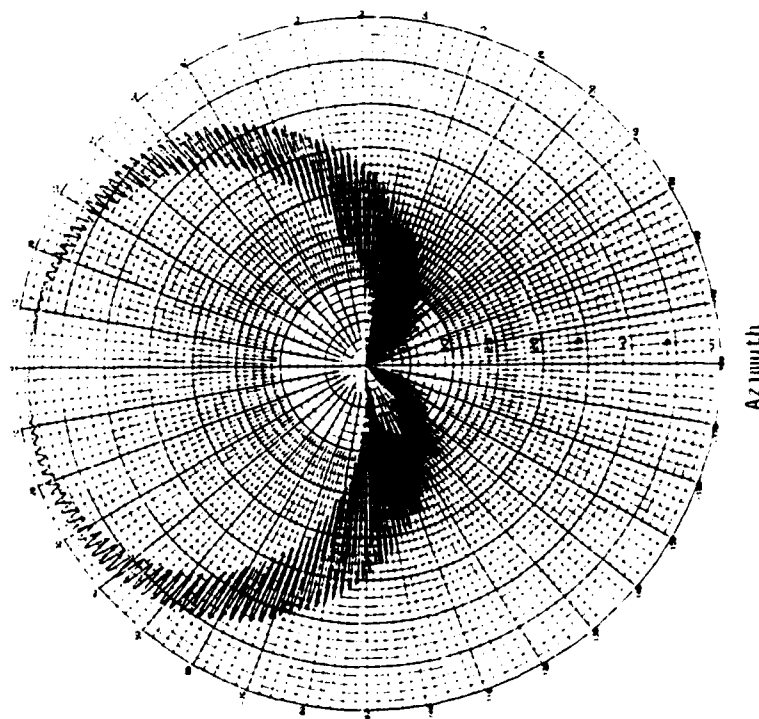


Figure 9. 6.0 GHz Radiation Patterns of a 3-Inch Archimedian Spiral at 0.5 Inches Over a 3-Inch Circular Ground Plane



Elevation



Azimuth

Figure 10. 6.0 GHz Radiation Patterns of a 3-Inch Archimedian Spiral at 0.3 Inches Over a 3-Inch Circular Ground Plane

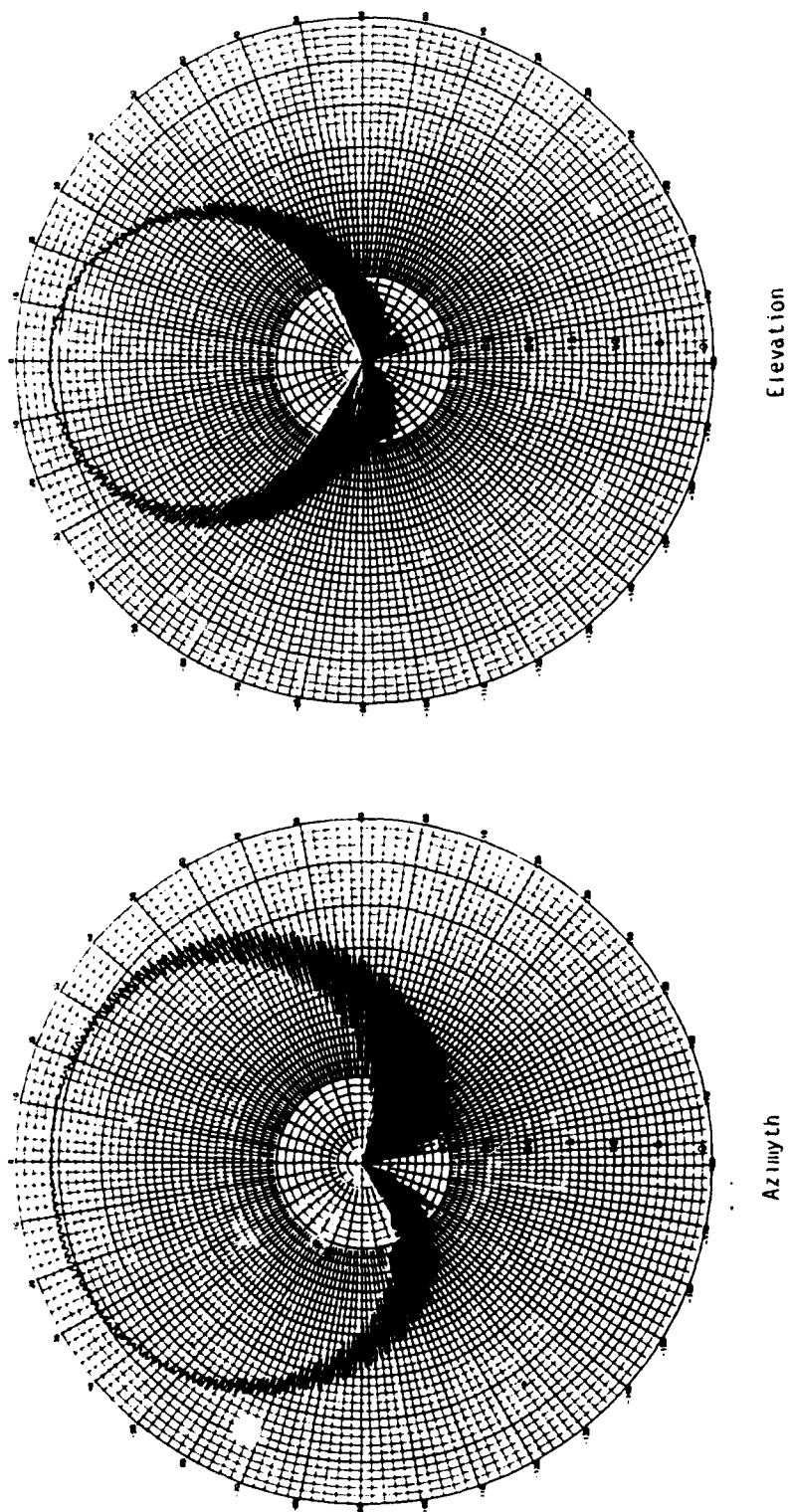
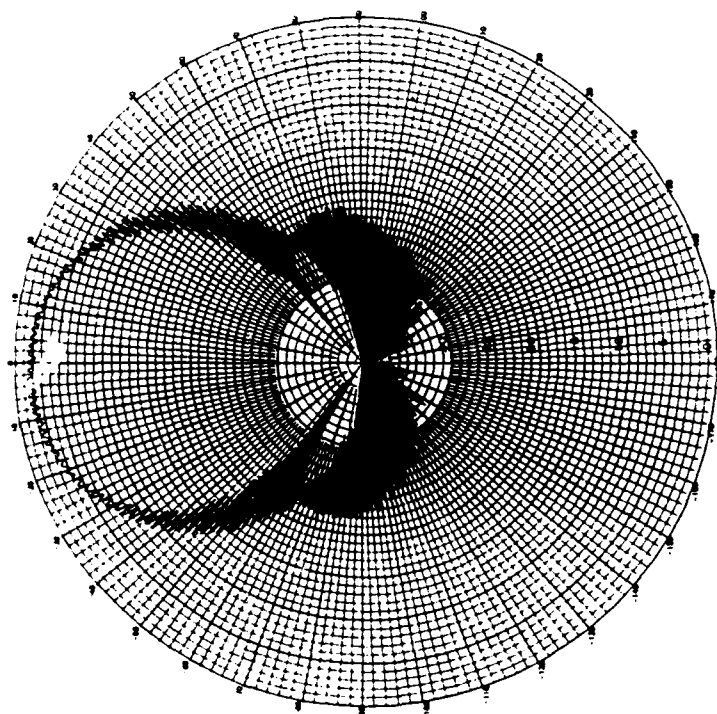
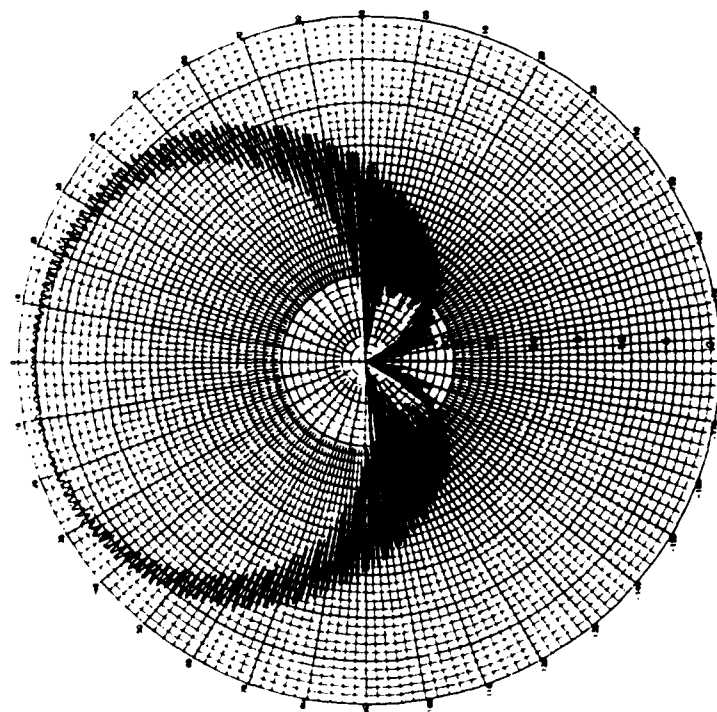


Figure 11. 6.0 GHz Radiation Patterns of a 3-Inch Archimedian Spiral at 0.1 Inches Over a 3-Inch Circular Ground Plane



Elevation



Azimuth

Figure 12. Radiation Patterns of a 3-Inch Archimedian Spiral (Teflon) at 0.3 Inches Over a 3-Inch Circular Ground Plane

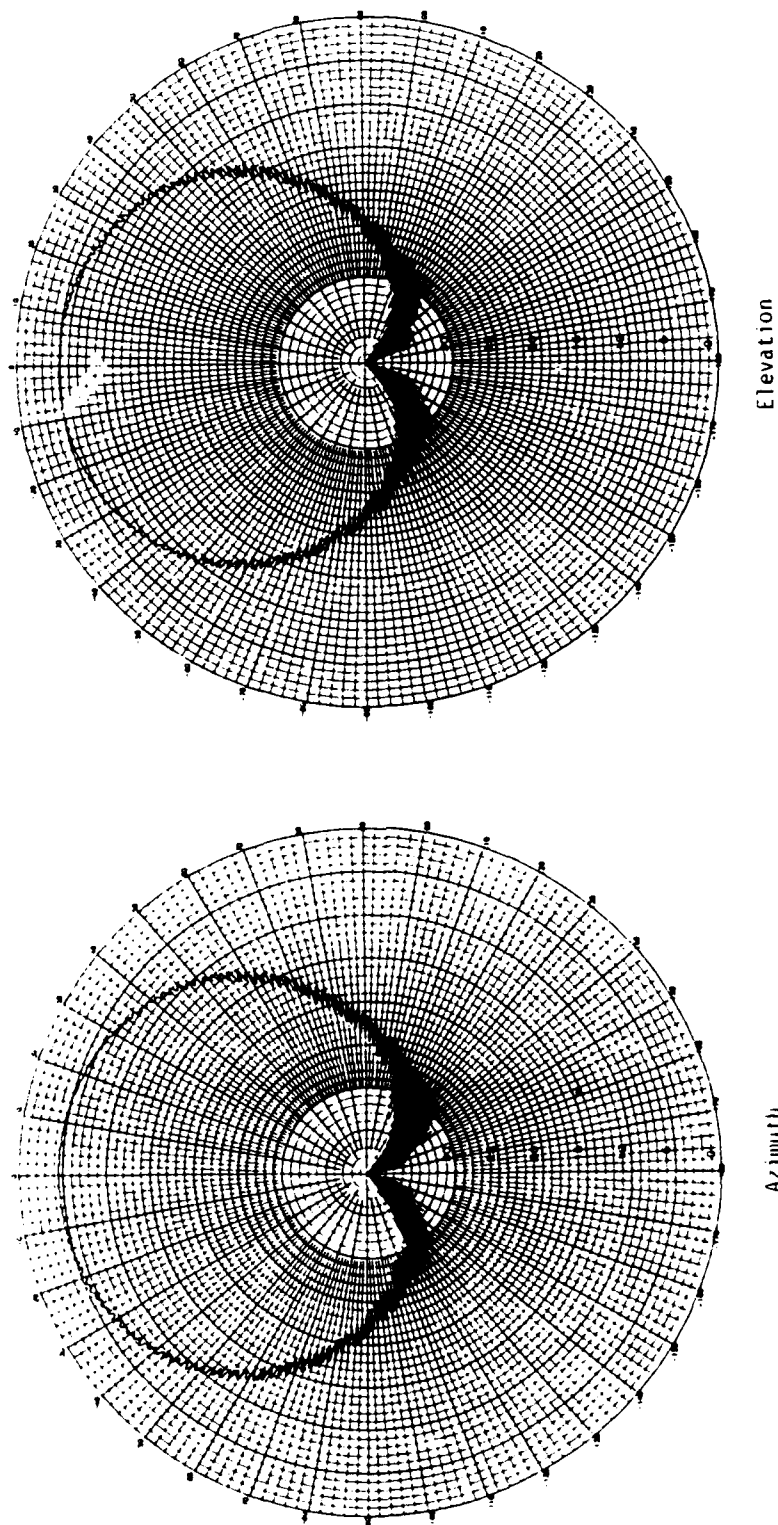


Figure 13. Radiation Patterns of a 3-Inch Cavity Backed Archimedian Spiral (0.7 Inch Depth)

The near field and the transmission tests only suggest the presence of higher-mode radiation. In the conical-cut pattern tests, however, the radiation patterns definitely show the existence of higher-mode excitation. The conical-cut pattern shown in Figure 5-a was obtained by simulating a spiral in free space by using the absorber disk. Here, only the one-wavelength mode radiation is excited, and the conical-cut pattern is nearly circular. When the conical cut pattern is like this, the two orthogonal azimuth patterns are almost identical as demonstrated by the two patterns shown in Figure 8. These patterns have fairly well behaved axial ratios, but the spiral antenna has a low gain. the absorber disk was later removed and replaced with absorber rings, and the conical-cut patterns are no longer circular. The patterns shown in Figures 5-b and 5-c, for 0.5-and 0.3-inch spacings, respectively, have an elliptical shape that demonstrates the excitation of the three-wavelength mode radiation. The elliptical conical-cut patterns indicate that the radiation patterns of the spiral are fan shaped. The two orthogonal azimuth patterns, shown in Figures 9 and 10 for 0.5- and 0.3-inch spacings, respectively, were measured at the highest- and the lowest-amplitude points on the conical-cut patterns and demonstrate the maximum and the minimum beamwidths of the fan-shaped beams. When the spacing is 0.5 inch, the 3-dB beamwidths of fan-shaped beam are 95° and 60° at the widest and narrowest parts, respectively, for a ratio of approximately 1.5:1. When the spacing is reduced to 0.3 inch, the 3 dB beamwidths of the wide and narrow part of the beam were 78° and 47° for a ratio of 1.65. When the spacing is further reduced to 0.1 inch, the conical-cut patterns are no longer elliptical in shape, but a figure 8, as shown in Figure 5-d. Two orthogonal patterns were measured and are shown in Figure 11. The 3-dB

beamwidths of the widest and the narrowest part of the fan-shaped beam are 84° and 43° , respectively, giving a ratio of almost 2:1.

It is well known that a spiral antenna etched on Teflon fiberglass PC board provides higher gain, especially at microwave frequencies, than that of one etched on G10 board because of the lower Loss tangent of the Teflon fiberglass material. The lower loss of the Teflon board is a disadvantage for spirals operating over a ground plane because the residual currents are attenuated less between the active regions of the two modes. This was demonstrated by replacing the G-10 spiral with the Teflon spiral and repeating the same tests. The conical-cut pattern measured at 6.0 GHz is shown in Figure 6. The pattern no longer elliptical but rather, figure 8 shaped; it has developed four lobes and four nulls. Figure 12 shows that the azimuthal pattern, measured in the narrow plane, also has two sidelobe peaks at 60° and two nulls at approximately 52° off the antenna axis. Clearly, some other higher modes have been excited. If an equiangular spiral were used instead of the Archimedian, even more high-order modes would be excited. This is because the spiral arms of an equiangular spiral are much shorter and much wider than those of the Archimedian spiral. Therefore, the attenuation between the one- and three-wavelength-mode active regions is much lower and, consequently, there will be more higher-mode radiation with higher magnitude.

The conical cut pattern of a cavity-backed spiral with a relatively shallow cavity (0.7 inch for the 3-inch spiral) was measured; it shows almost the same performance (circularity and axial ratio) as that of the spiral with the absorber disk backing. The two orthogonal patterns are shown in Figure 12 for comparison.

When the 3 inch circular ground plane was replaced with a 10-inch one and the tests were repeated, the test data showed no significant change.

Two phenomena are seen from these radiation patterns: First, the verification of the fact that, when the spacing between the spiral and the ground plane is reduced, more higher-mode radiation is excited. When the spiral aperture is large enough to support the three-wavelength mode, the radiation patterns are no longer axisymmetrical. The fan-shaped beam will rotate about the antenna axis as the frequency changes. This results in the cycling of the beamwidth as a function of frequency through any great-circle plane of the spiral antenna.

Secondly, microstrip-mode spirals have high off-axis axial ratios that decrease when the ground-plane spacing is reduced. The axial ratio of the 3 inch spiral measured with a 0.3 inch spacing from the ground plane is unacceptable for many applications. When this spacing was reduced to 0.1 inch, the axial ratio was reduced by about half. Unfortunately, the gain of this spiral, spaced at 0.1 inch, is comparable to that of a cavity backed spiral, which has much better pattern characteristics.

4. EXCITATION CONSIDERATIONS

The two-arm spiral antenna is a balanced antenna requiring an excitation having two equal amplitude and 180° out-of-phase currents. Deviations in amplitude or phase balance of these excitation currents degrades the performance. A common device to excite the two-arm spiral is called a "balun" (a device with an unbalanced input and a balanced output).

In the infant days of the spiral antenna, a coaxial cable was soldered to the spiral arm and connected at the feed point in the infinite-balun manner. This

method is very lossy and is also frequency limited. Some people even attempted to connect the inner and outer conductor of a coaxial cable directly to the two input terminals of the spiral antenna. In the late 1950s, several-octave magictees were used to evaluate the performance of the broadband spiral antenna. Two broadband baluns were introduced in 1960. The Duncan and Minerva Balun¹, a tapered coaxial balun, has a relatively low loss, but phase balance is not very desirable, and it tends to radiate at microwave frequencies, degrading the spiral patterns. The Bawer and Wolfe Balun² is a stripline version of the Roberts Balun. Even though it is a stripline structure, it is installed perpendicular to the spiral radiator. Many versions of modified Marchand baluns, developed later by the cavity-backed-spiral designers, also provide excellent phase and amplitude balance over very broad frequency bands.

All of these baluns add to the depth of the spiral antenna. The balun itself increases the depth of the spiral antenna considerably. Also, since the input to the balun is 50 ohms and the spiral antenna usually has a higher input impedance, an additional transformer is required to match the impedances. The length of this transformer depends on the required bandwidth and the lowest operating frequency. Broadband transformers are generally relatively long. For this reason, the depth of the cavity-backed spiral was increased to accommodate this feed structure.

To an antenna designer, the feed mechanism is an integral part of the spiral antenna, and its depth is considered part of the antenna.

5. CONCLUSION

The performance of the planar spiral antenna over a ground plane has been revisited. The radiation patterns of this antenna over a finite ground plane, at close

spacings (microstrip mode), have been examined in detail. When the spacing between the ground plane and the radiating element is kept small for a 3-inch spiral (between 0.3 and 0.5 inch), the radiation patterns are characterized by very high off-axis axial ratios. As long as the spiral aperture is large enough to support the 3λ mode radiation, the radiation patterns will be fan-shaped. Although these microstrip-mode spiral antennas have approximately 2 to 3 dB more gain than a cavity-backed spiral of the same aperture, the qualities of their radiation patterns is so poor that they are undesirable for most applications.

The largest markets for low-profile antennas are air-borne direction-finding systems, and, as can be seen from the foregoing measurements, microstrip-mode spiral antennas would provide poor performance because of their high axial ratios and beamwidth variations.

The earlier airborne DF systems used four spiral antennas to provide azimuthal coverage of 360 degrees. Therefore, only the portion of the radiation pattern between $\pm 90^\circ$ from the antenna axis are utilized. The newer systems use six antennas, where only the portion of pattern between $\pm 60^\circ$ from the antenna axis is used. Even with the added number of antennas, the performance of the microstrip-mode spiral is inadequate. For instance, Figure 14 illustrates the overlay of the two orthogonal-cut patterns of the fan-shaped beam shown in Figure 10. At 60° from the antenna axis, the DF error introduced by the antenna alone exceeds 20° (with 0.4 dB/deg as the error gradient).

For a monitor antenna, which is required to receive any linearly polarized signal within a 120° conical sector, the cavity-backed spiral can provide much better coverage than a microstrip spiral. This is because of the symmetrical beam and

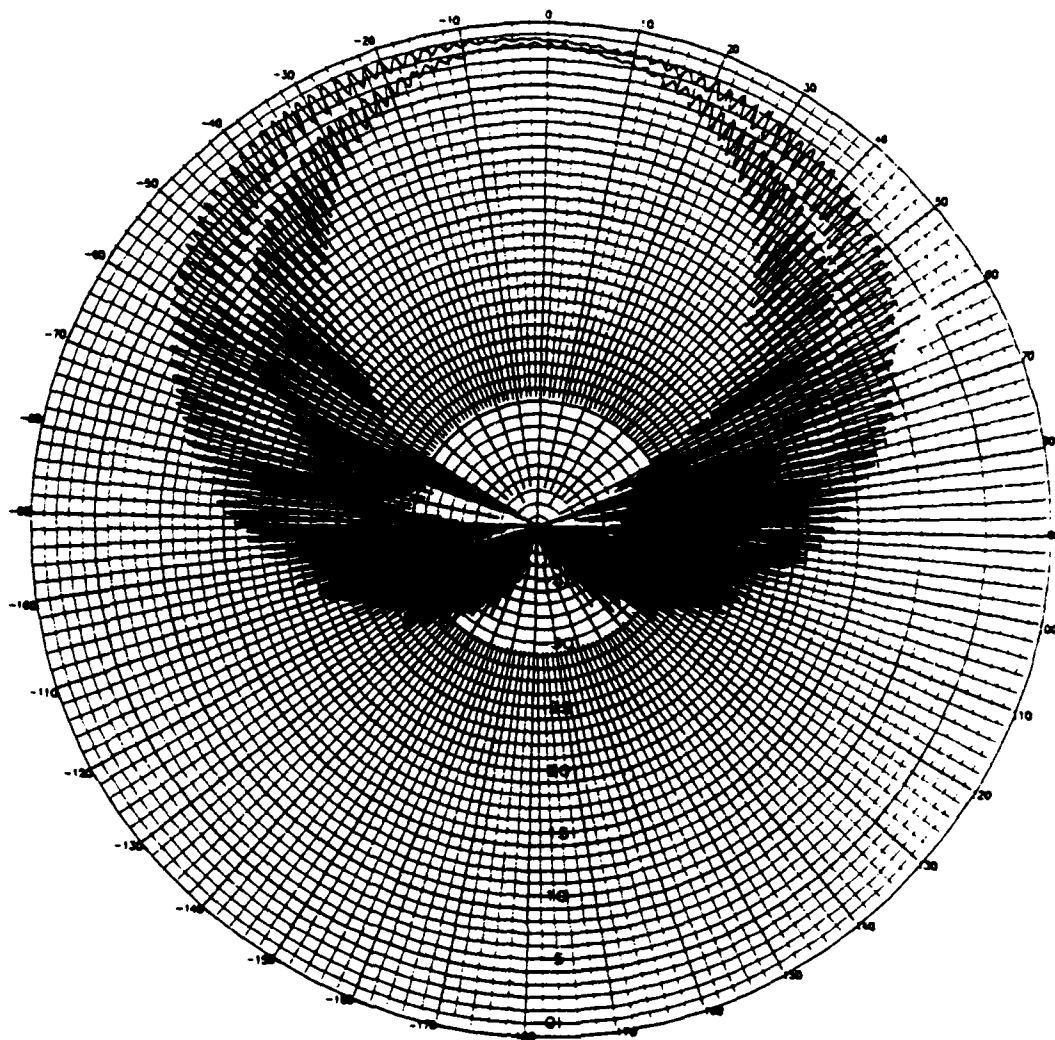


Figure 14. 6 GHz Overlay Azimuth and Elevation Patterns of a 3-Inch Archimedian Spiral over a 3-Inch Ground Plane at 0.3 Inches Spacing

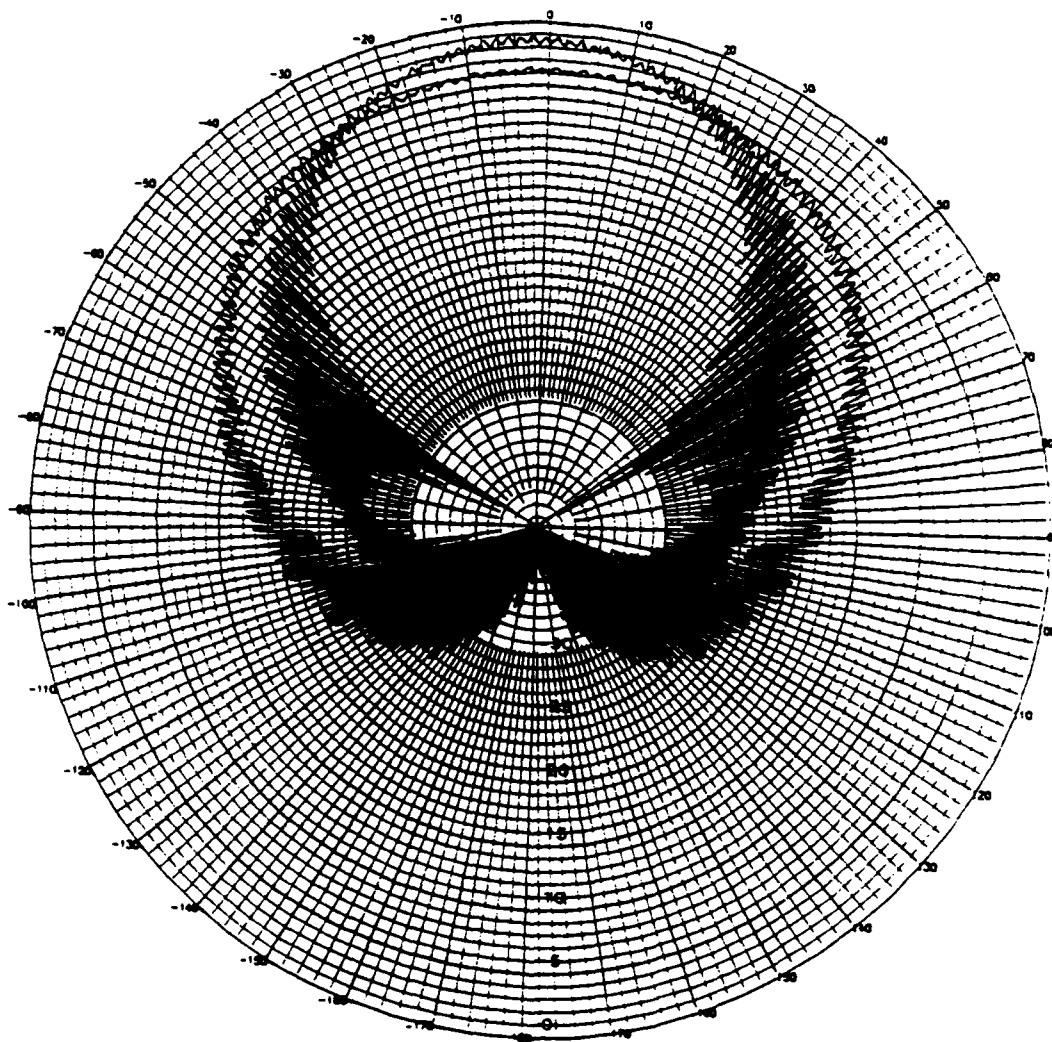


Figure 15. Overlay Patterns of a Cavity Backed Spiral and a Microstrip Mode Spiral of the Same Design

good axial ratio even though the peak gain of the cavity-backed spiral may be 2 to 3 dB less. This can be shown by comparing the overlaid patterns of a cavity-backed spiral versus a microstrip spiral of the same design, as shown in Figure 15.

Spiral antenna designers prefer the cavity configuration because of its superior pattern performance but also because of its mechanical integrity. The spiral element, cavity, balun, and the impedance transformer are packaged into a single unit that has been qualified for many difficult environmental requirements. It is not clear how the balun structures can be packaged into the microstrip-mode spirals so that they can be "pasted" on to automobiles or aircraft.

The sinuous antenna, introduced recently, has a lower Q than the spiral antenna and is not so sensitive to the position of the ground plane. This means that fewer higher modes will be excited, and, therefore, it may be a better choice for microstrip-mode operations. However, a high off-axis axial ratio is still a problem that cannot be eliminated by using the sinuous elements. The sinuous antenna, unlike a two-arm spiral, requires a complicated four-port device for its excitation that tends to prevent it from realization as a low profile structure.

Unless a new technology is developed that will allow a low-profile antenna to be fed efficiently without increasing its profile, the microstrip planar spiral, with its less-than-desirable performance, does not offer any advantages over a standard cavity-backed spiral. Both have parts that need to be recessed below the surface of interest.

6. REFERENCES

1. Duncan, J.W. and Minerva, V.P. (1960): 100 : 1 balun transformer. Proc. IRE 48 (No. 2): 156-165.
2. Bawer, R. and Wolfe, J.J. (1960): A printed circuit balun for use with spiral antennas, Proc. IRE-Trans.-MTT 8 (No. 3): 319-325.

**ANALYSIS OF LOG-PERIODIC
FOLDED DIPOLE ARRAY**

X00169

H. Shnitkin

**Norden Systems, Inc.
P.O. Box 5300
Norwalk, CT 06856**

1992 Antenna Applications Symposium

September 1992



ANALYSIS OF LOG-PERIODIC FOLDED DIPOLE ARRAY

**H. Shnitkin
Norden Systems, Inc.
P.O. Box 5300
Norwalk, CT 06856**

Abstract

About 25 years ago John J. Greiser invented a log-periodic antenna array (reference 1) consisting of folded dipole or monopole elements at the University of Illinois. His results, as described in his U.S. Patent 3,369,243, led to an extremely simple log-periodic backfire dipole configuration that permits printing of the entire log-periodic array on one side of a circuit board. The usual cumbersome twist of the feeder line between two adjacent dipoles has thus been avoided by providing the required 180-degree phase shift through the folded dipole configuration. Greiser's simple configuration can easily be bifurcated, without sacrifice of performance, making log-periodic monopole arrays easily realizable. Thus, readily manufacturable low-profile, log-period arrays became practical. To supplement Greiser's experimental data, a mathematical model of a reentrant folded dipole is derived in this paper. With such a model, a complete antenna excitation function and far field radiation pattern could be derived. This model would also allow parametric analysis and optimization of the antenna dimensions, leading to higher gain and lower sidelobe antennas with improved impedance match.

In the hope that a better theoretical understanding of the behavior of the folded dipole element accompanied by a mathematical model would encourage

wider use of the log-periodic folded dipole arrays (LPFDA), or monopole arrays, a detailed theoretical discussion of the folded element used in Greiser's experiments is presented here.

1. Theoretical Approach

First, it must be realized that the basic radiating element of an LPFDA is a two-port network, appropriately named a "reentrant folded dipole," as shown in Figure 1. Second, these dipoles are connected in series, rather than in shunt, to the feeder transmission line. From the circuit of Figure 1, a four-terminal network admittance matrix can be derived, leading to an equivalent circuit shown in Figure 2. Subsequent examination of this equivalent circuit yields an understanding of how a 180-degree phase shift is obtained in an LPFDA without twisting the feeder line. In addition, the transmission phase and amplitude passed to the remaining dipoles of the array has been computed.

2. Circuit Analysis

The reentrant folded dipole of Figure 1 is an extension of the single-port folded dipole, used commercially, and treated theoretically in references 2 and 3. Essentially, the two-port shown in Figure 1 operates in two simultaneous modes, namely, the transmission line stub mode (denoted by subscript "s") and the dipole radiation mode (denoted by subscript "r"). Superposition of the two modes will be applied in the process of developing a two-port admittance matrix.

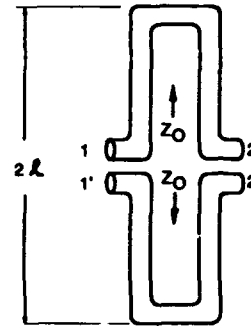
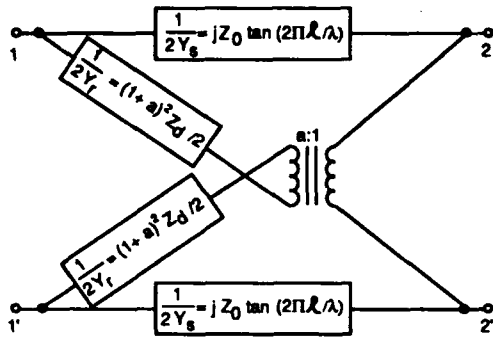
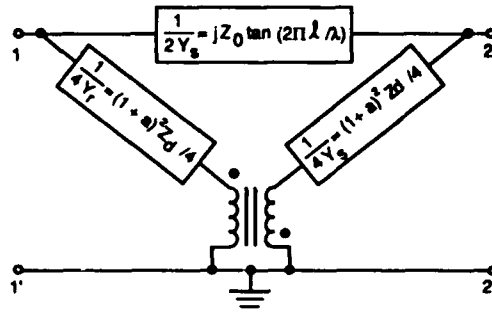


Figure 1. Configuration of Reentrant Folded Dipole



NOTE: EXPRESSIONS REPRESENT IMPEDANCES !

Figure 2. Equivalent Circuit of Reentrant Folded Dipole



NOTE: EXPRESSIONS REPRESENT IMPEDANCES !

Figure 3. Equivalent Circuit of Reentrant Folded Monopole

If terminals 2 and 2' are short circuited and if the left and right dipole arms are of the same diameter, the well-known single-port folded dipole results, with its input admittance of $Y_r + Y_s$, where, from reference 2,

$$Y_r = 1/(4 \cdot Z_{\text{equiv. cyl. dipole}}) \text{ and } Y_s = -j / [(2Z_{\text{o-stub}} \cdot \tan(2\pi l / \lambda))] \quad (1)$$

Consequently, the Y_{11} and Y_{22} terms of the Y-matrix for the two-port folded dipole are also equal to $Y_r + Y_s$, so that:

$$I_1 = (Y_r + Y_s) V_1 + Y_{12} V_2 \text{ and } I_2 = Y_{21} V_1 + (Y_r + Y_s) V_2 \quad (2)$$

To compute $Y_{12} = Y_{21}$, set $V_1 = V_2$ to excite the even mode and form an equivalent cylindrical dipole. Now $I_1 = I_2$ and

$$Y_{in} = (I_1 + I_2)/V_1 = 2I_1/V_1 = 2(Y_r + Y_s + Y_{12}) \quad (3)$$

from equation (2) left.

Since Y_{in} also equals $1/Z_{equiv. \text{ cyl. dipole}}$, then from equation (1) left

$$4Y_r = 2(Y_r + Y_s + Y_{12}) \text{ or } Y_{12} = Y_{21} = Y_r - Y_s \quad (4)$$

Consequently,

$$I_1 = (Y_r + Y_s) V_1 + (Y_r - Y_s) V_2 \text{ and } I_2 = (Y_r - Y_s) V_1 + (Y_r + Y_s) V_2 \quad (5) \& (6)$$

To confirm the above results, simply set $V_1 = -V_2$, which results in $I_1 = -I_2$ and therefore no dipole radiation. This agrees with $Y_{in} = I_1/2V_1 = Y_s$, from equations 5 and 6.

A refinement of the above includes different diameters for the left and right dipole arms. As per reference 2, an asymmetry constant "a" is defined, yielding $Y_r = 1/[Z_{dipole} \cdot (1 + a)^2]$ and:

$$I_1 = (Y_r + Y_s) V_1 + (a Y_r - Y_s) V_2 \quad (7)$$

$$I_2 = (a Y_r - Y_s) V_1 + (a^2 Y_r + Y_s) V_2 \quad (8)$$



Figure 4. Ten-Element LP Folded Monopole Array

The equivalent circuits resulting from these equations are shown in Figures 2 and 3, for dipole and monopole, respectively.

Now the transfer ratio, V_2/V_1 , can be obtained from equation 8 by setting $I_2 = -V_2 Y_L$ yielding:

$$V_2 / V_1 = -(aY_r - Y_s) / (a^2 Y_r + Y_s + Y_L) \quad (9)$$

At dipole resonance, $Y_s = 0$, making V_2/V_1 negative or 180° out of phase. Away from dipole resonance Y_r is much smaller than Y_s , so that: $V_2/V_1 = Y_s/(Y_s + Y_L) = Z_L/(Z_s + Z_L)$, indicating that the shorted transmission line stub is simply connected in series.

To use equations 7 and 8 for LPFDA dipole excitation and radiation analysis, the equivalent ABCD - matrix parameters should be evaluated and subsequently cascaded, via matrix multiplication, with the interconnecting transmission lines between adjacent folded dipoles.

3. Experimental and Computational Results

To illustrate the potential of the LPFDA, measured patterns of two typical LP folded monopole arrays on large ground planes have been obtained. The first is a ten-element array, shown in Figure 4, for a broad-band IFF array antenna. Its apex angle is 58 degrees and its proportionality factor 0.8. Measured radiation pattern is given in Figure 5. The second is a 21-element array, shown in Figure 6, for use as a flush-mounted direction finding interferometer. Its apex angle is 15 degrees and its proportionality factor 0.91. Measured radiation patterns at $f_s=3$ GHz, $f_c = 6$ GHz, and $f_x = 10$ GHz are given in Figure 7.

The above examples preceded the folded dipole analysis presented here and therefore display performance deficiencies, e.g., gain, sidelobe, impedance, and front/back ratio. Application of the folded dipole theory would undoubtedly improve their performance.

4. Conclusion

The admittance matrix of a reentrant folded dipole has been derived, its equivalent circuit evaluated and a mathematical model presented. These results when combined with mutual coupling theory and an iterative network computation allow the evaluation of all element excitation coefficients. From these, far-field radiation patterns of a log-periodic folded dipole backfire array can be predicted and tolerance analysis as well as optimization of antenna dimensions for best gain and sidelobes can be performed.

5. References

1. Greiser, J.W., U.S. Patent No. 3,369,243 "Log-periodic Antenna Structure," Feb. 13, 1968.
2. Jasik, H., 1961, *Antenna Engineering Handbook*, pp. 3-13.
3. Mushiake, Y., Oct. 1954, "An Exact Step-up Impedance Ratio Chart of a Folded Antenna," IEEE Trans, Vol. AP-3, No. 4.

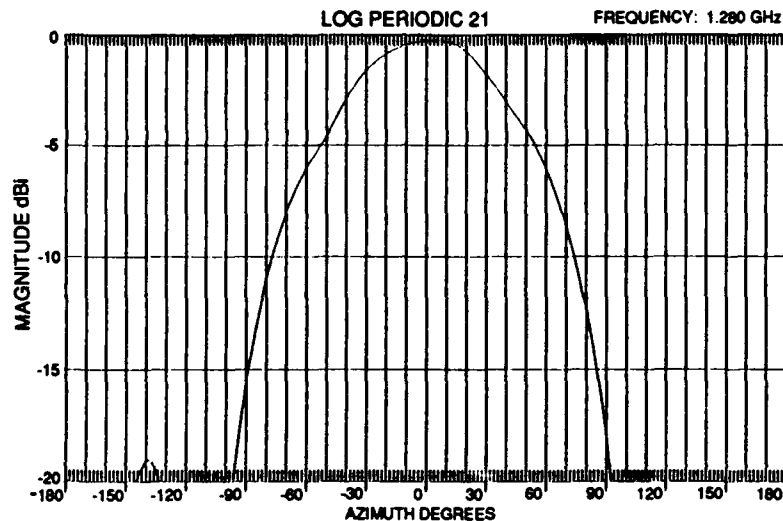


Figure 5. Measured H-plane Radiation Pattern of 10-monopole Array



Figure 6. 21-Element LP Folded Monopole Array Interferometer

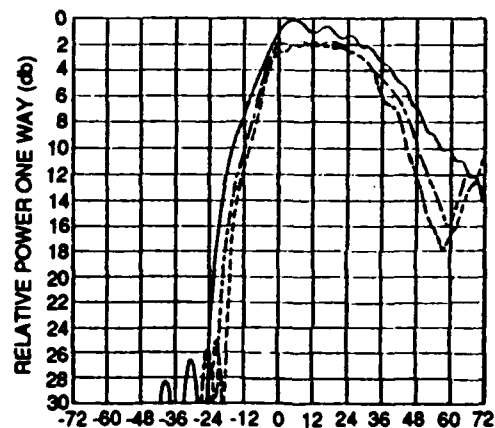


Figure 7. Measured E-plane Radiation Pattern of 21-element Monopole Array

A BROADBAND PANEL ANTENNA FOR HDTV APPLICATIONS

**Rex L. Niekamp
Manager of Antenna Operations
Harris Corporation
Harris Allied Broadcast Division
Quincy, Illinois**

1.0 ABSTRACT:

Current NTSC television transmission will begin to make way for a new system of broadcasting-- High Definition Television (HDTV)-- once the Federal Communications Commission (FCC) adopts a standard for the United States. The FCC decision is anticipated in 1993.

Conversion to the new HDTV system will be gradual, extending well into the 21st Century. The FCC has ruled that commercial television stations planning to convert to HDTV will be required to simulcast NTSC and HDTV signals during the transition. Current U.S. broadcasters will be able to apply for a second channel for HDTV transmission, then essentially will operate two channels instead of one per station for a period of time.

Given the associated cost, most broadcasters will need to consider trade-offs based upon available financing. Three potential options include: 1) quickly implementing a second "full power" HDTV channel (i.e., immediate large investment in HDTV equipment); 2) adding a low power HDTV signal which serves the primary coverage area and protects the channel for future expansion (i.e., immediate medium investment in HDTV equipment), or 3) simply maintaining an NTSC signal at risk of not being able to obtain license for an HDTV channel in the future.

For stations planning to offer some level of HDTV, antenna planning will be a major consideration. Many existing towers are already loaded to capacity, and most broadcasters will want to avoid the expense and the complexities of building a new tower. One solution-- using a common (multiplexed) wideband antenna which is sidemounted on a tower to free up valuable tower space-- will be a viable and attractive option to new tower construction¹.

Multiplexing channels for either NTSC or HDTV will require an antenna with wide impedance and pattern bandwidth, and high average and peak power handling capability. This paper will discuss one type of antenna-- the Deltawing panel antenna for VHF and UHF frequencies-- which is optimally suited to this requirement.

2.0 DESCRIPTION OF THE DELTAWING PANEL ANTENNA

The Deltawing antenna is an array of Deltawing panel radiators. When arranged properly on a tower, these radiators produce the desired azimuthal and elevation pattern. Patterns are produced by controlling the panel's physical orientation, its power and its phase. Figure 1 shows a typical panel arrangement of a Deltawing antenna.

A branch feed system is used to control the amount of power and the phase to each panel. Radiated power is horizontally polarized.

Operation of the Deltawing panel antenna may best be understood by analogy to the standard dipole antenna, and is very similar to the classic turnstile antenna element². Basically, the Deltawing radiator is a modification of the standard center fed dipole.

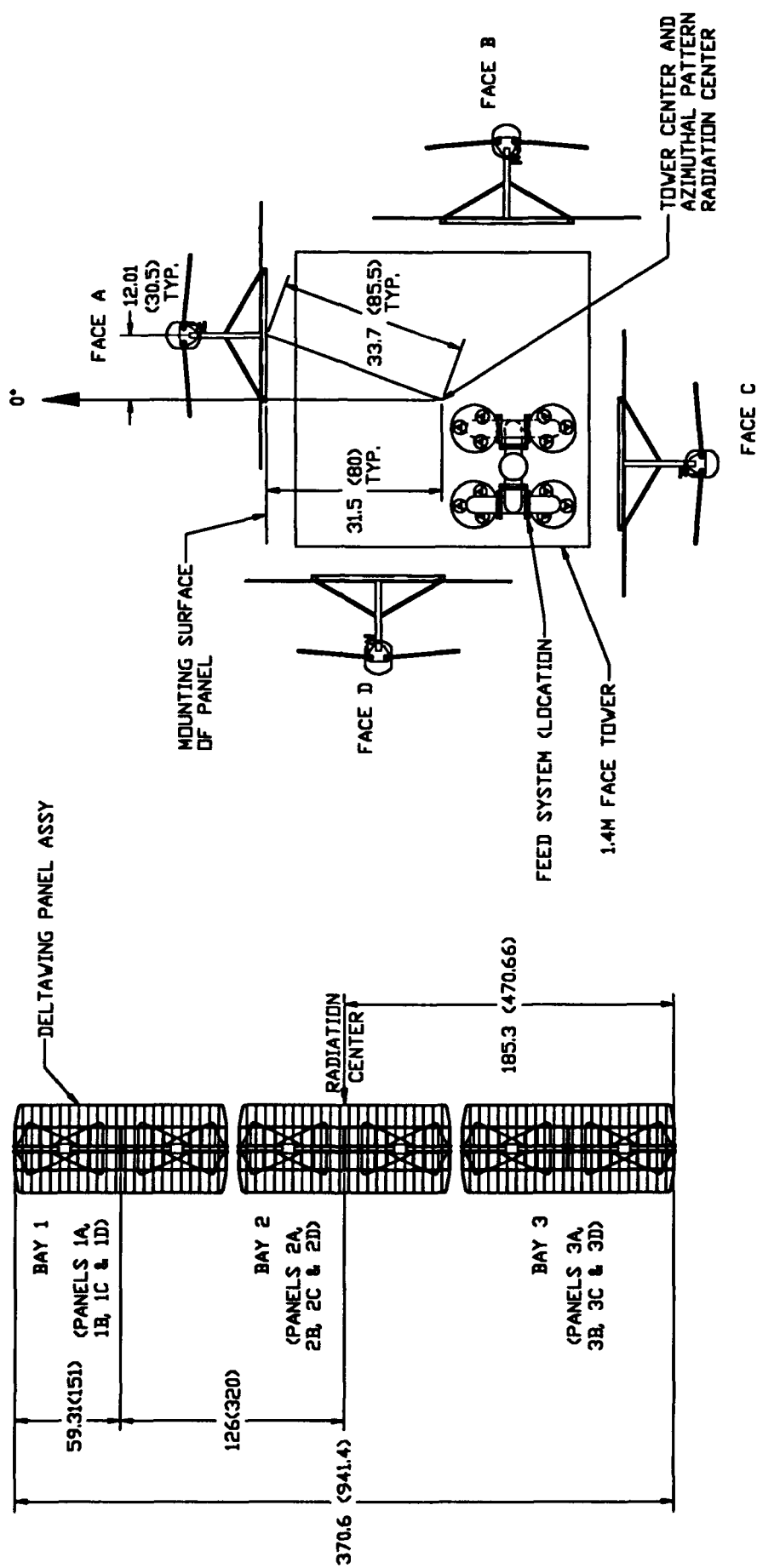
First, assume the arms of the dipole are made fan shaped, resembling a cross-section of two cones fed at the apex. This increase in element width correspondingly increases bandwidth over that of a standard dipole. If two quarter wavelength stubs are connected to the feedpoint, bandwidth is further increased.

Then, assume the dipole is filled solid to the stubs and notched in the middle to obtain uniform current distribution. In order to reduce the wind loading on VHF models, a sufficient number of horizontal elements are added across the radiators to simulate a solid element. The resulting element is called a Deltawing radiator.

Deltawing radiators are mounted approximately $1/4$ wavelength in front of a backplane. A grid design is used on the backplane of VHF antennas. This design minimizes weight; minimizes windloading, and minimizes back radiation into the tower structure and the related pattern distortion that may be caused by reflections from tower members and feed system components.

Lightning protection is provided by attaching the backplane to the top and bottom of the radiators, points of zero RF potential (DC ground). Impedance of each antenna is determined by panel geometry and mutual impedance presented by adjacent panels.

The element's azimuth pattern width can be controlled by varying the width of the backplane, the height of the radiator above the backplane, and the angle of the radiator in relation to the backplane. Two useful pattern widths have been developed: A 90° (Cos^2) beamwidth for "4 around" (four panels mounted to square tower) applications (Figure 2), and a 120° (Cos) beamwidth for "3 around" (three panels mounted to triangular tower) applications (Figure 3). Single Cos^2 and Cos panels with appropriate dimensions are shown in Figures 4 and 5 respectively.



PLAN VIEW

ELEVATION VIEW
TOWER NOT SHOWN FOR CLARITY

FIGURE 1: TYPICAL PANEL ARRANGEMENT

RELATIVE FIELD AZIMUTH PATTERN

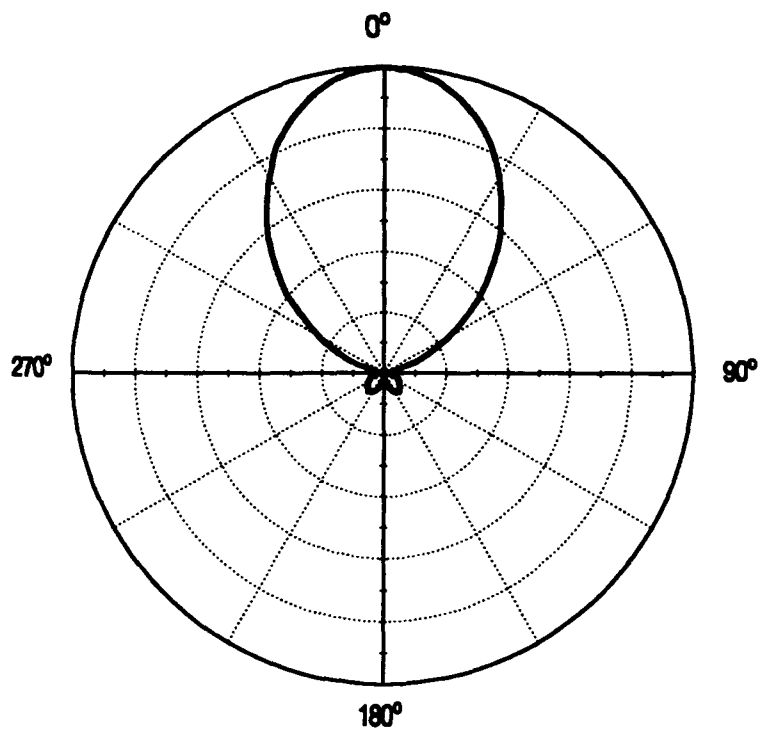


FIGURE 2: SINGLE PANEL, $\cos^2\theta$

RELATIVE FIELD AZIMUTH PATTERN

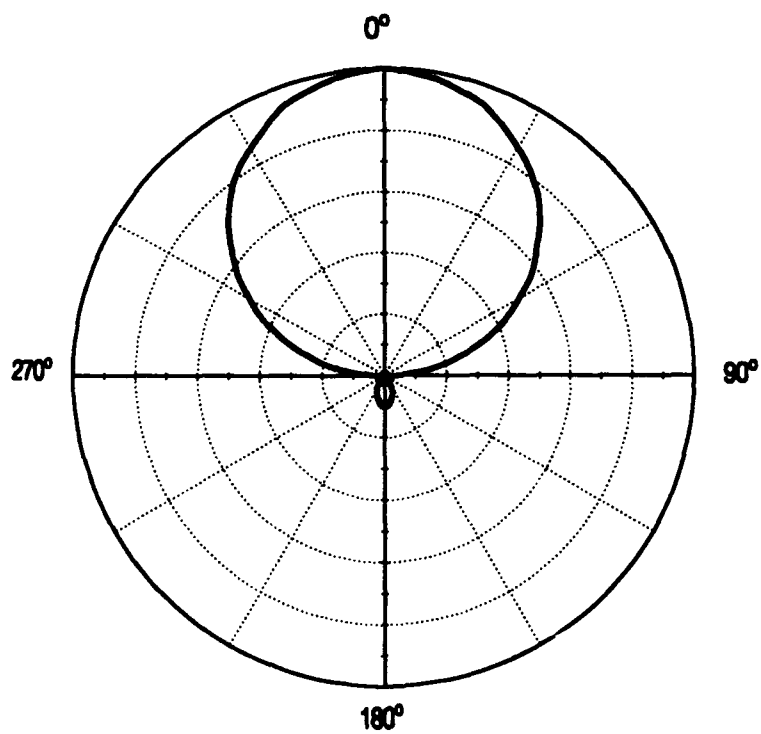


FIGURE 3: SINGLE PANEL, $\cos\theta$

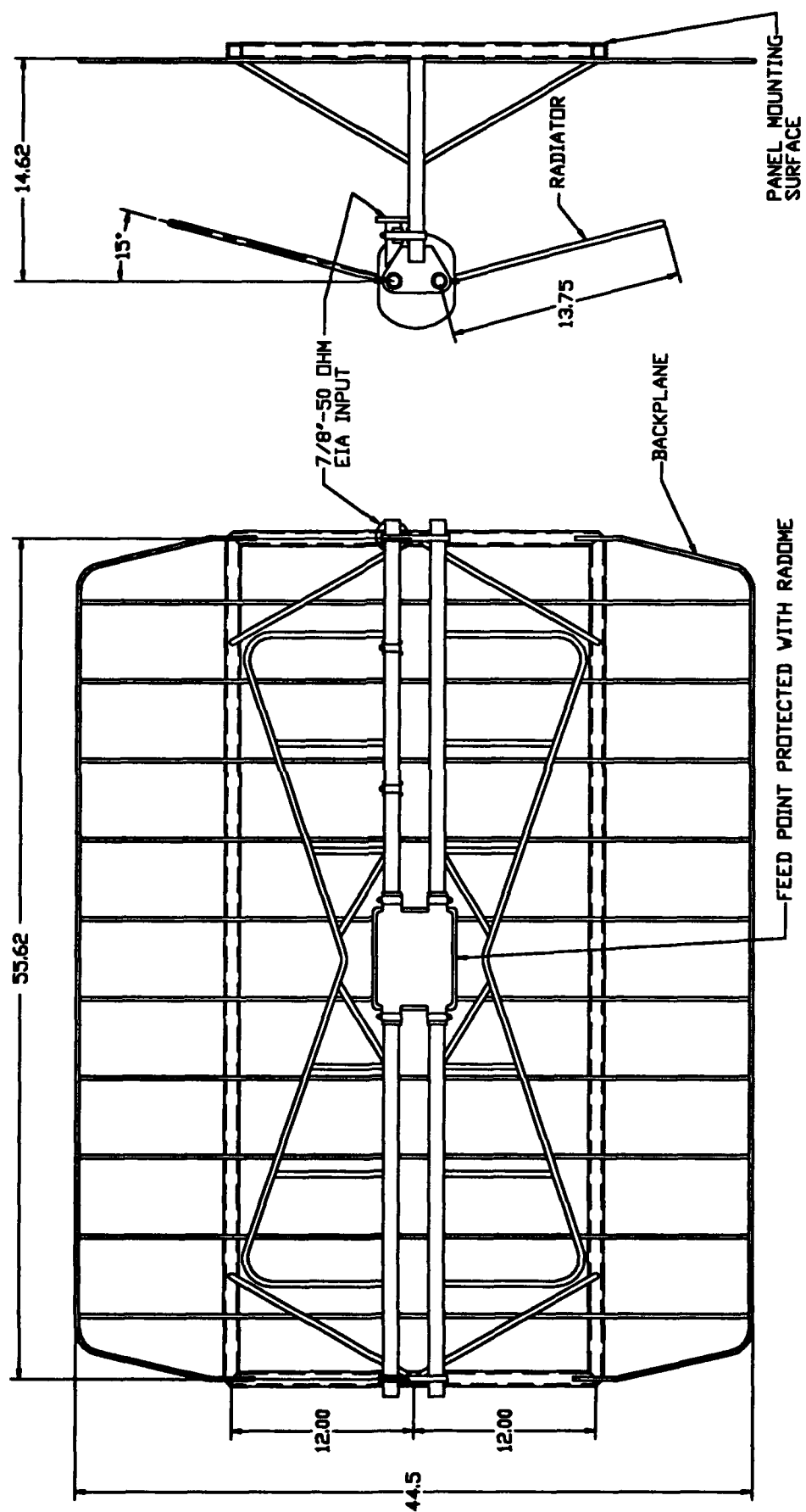


FIGURE 4. SINGLE PANEL, COS² 174-230 MHZ
DIMENSIONS IN INCHES

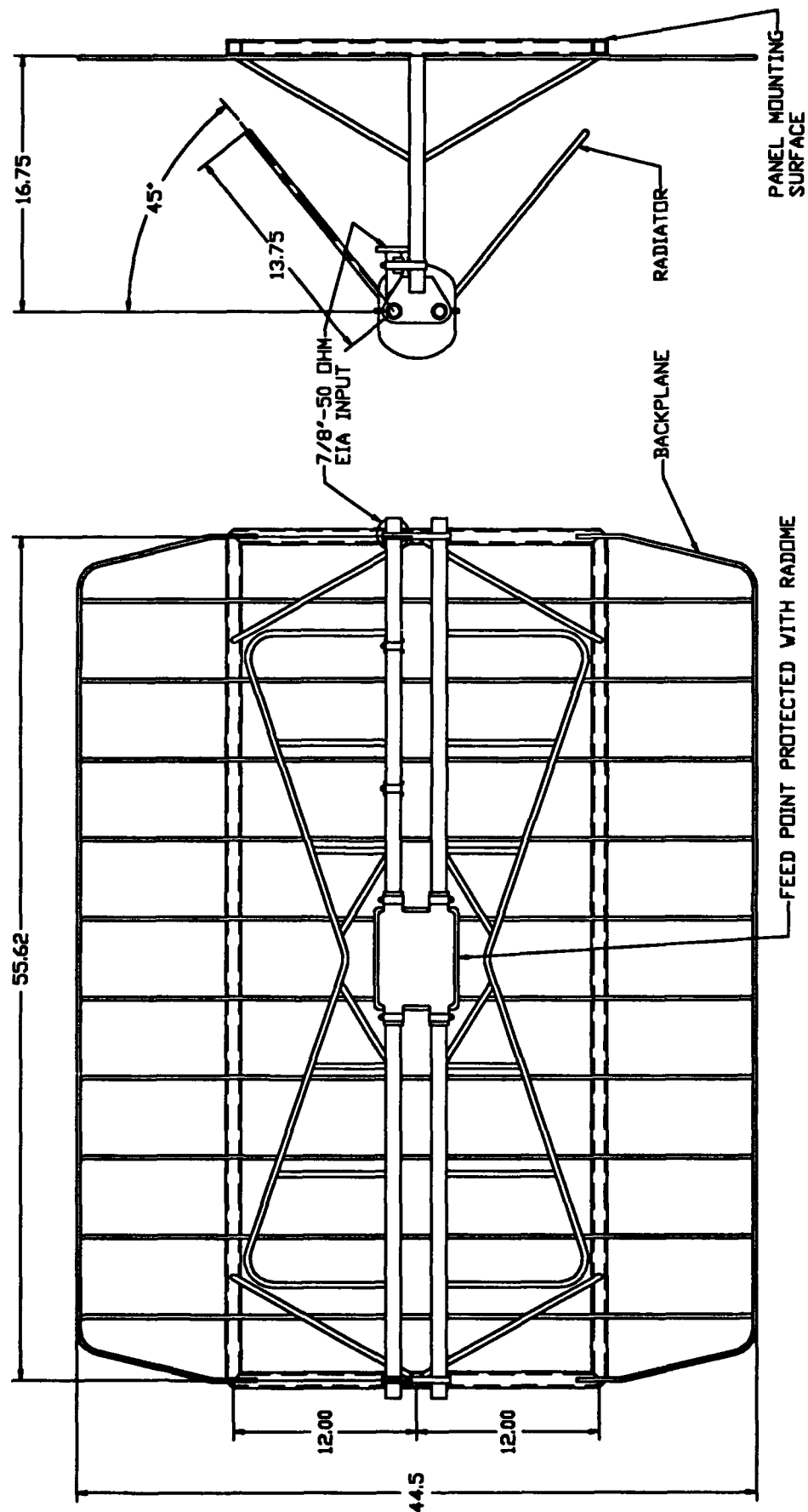


FIGURE 5: SINGLE PANEL, CDS 174-230 MHZ
DIMENSIONS IN INCHES

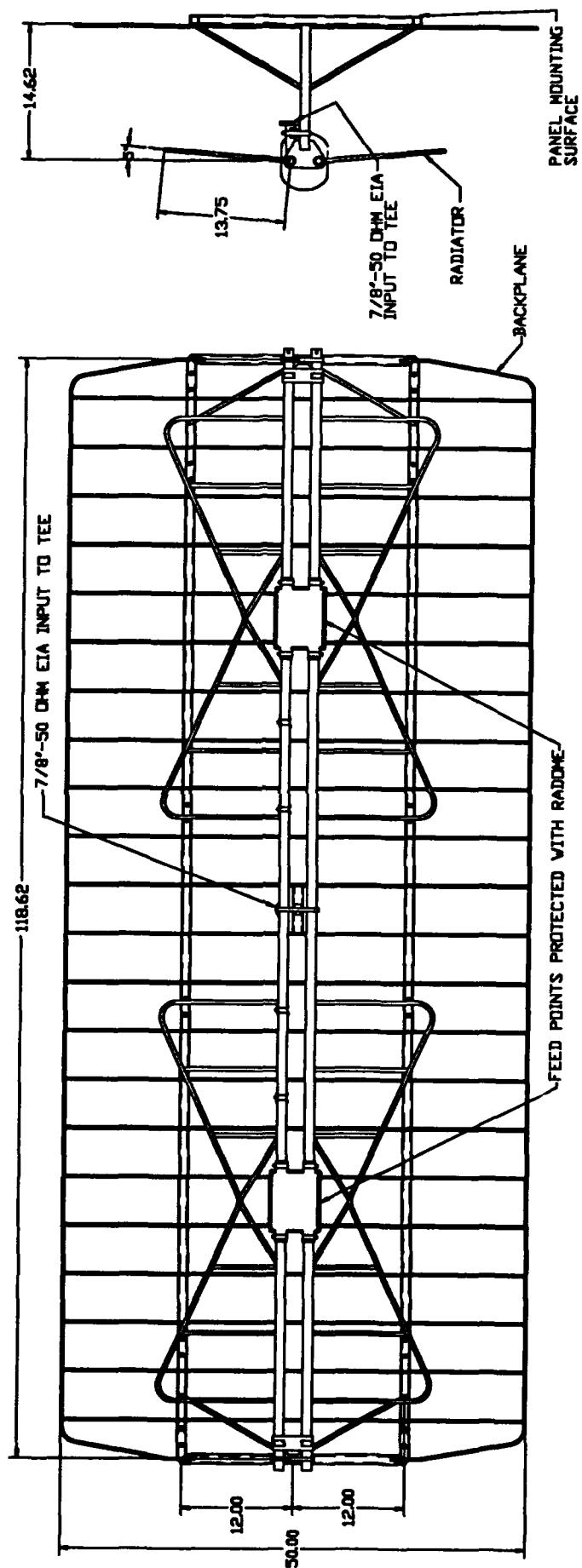


FIGURE 6. DUPLET PANEL, COS² 174-230 MHZ
DIMENSIONS IN INCHES

3.0 THE DUplet DELTAWING PANEL

Reliability of the antenna system can be improved by reducing the number of components in the system. In order to reduce the number of feedsystem components and to improve overall panel impedance, two Deltawing panels can be placed end to end and fed from a common "Tee" input (Figure 6). This configuration is called a "duplet" panel. Highband VHF and UHF versions of the duplet Deltawing panel have been developed using this concept. Figures 7 and 8 show the return loss for a typical Cos^2 and Cos Duplet element, and Figures 9 and 10 show the azimuth patterns at low and high frequency for the Cos^2 and Cos Highband Duplet element.

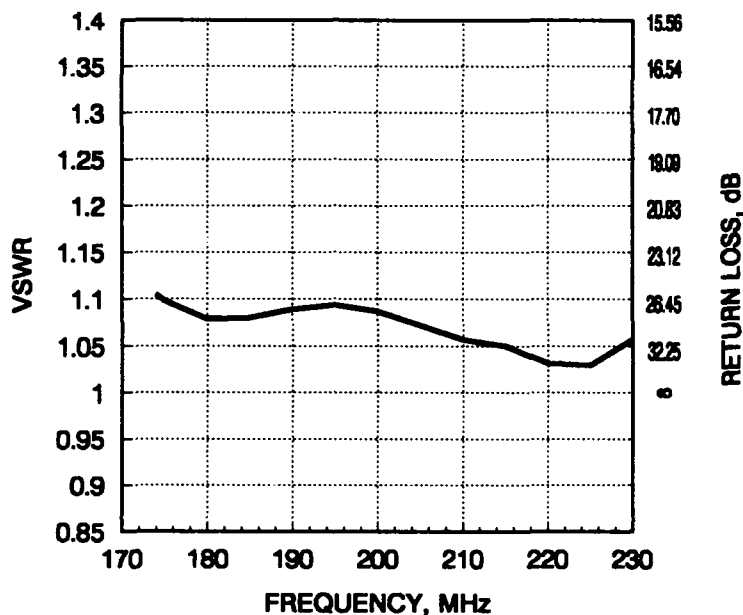


FIGURE 7: MEASURED PANEL VSWR, $\text{Cos}^2\theta$

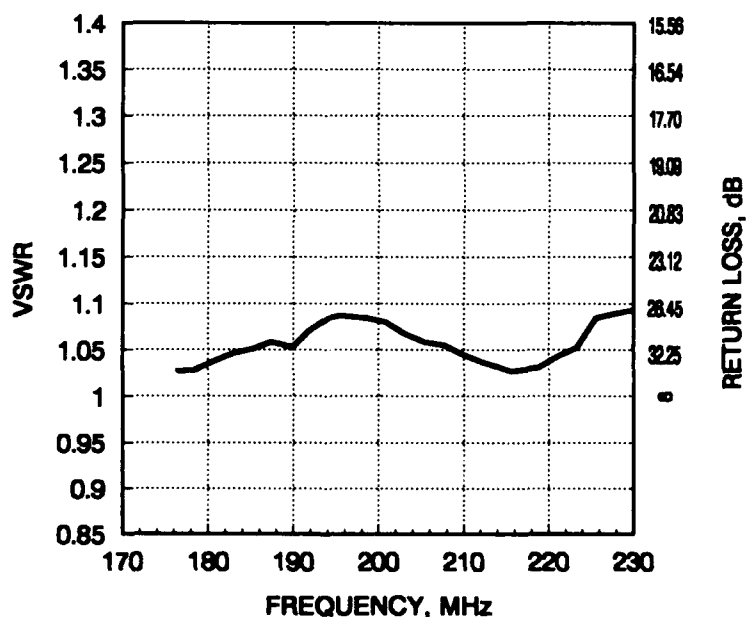


FIGURE 8: MEASURED PANEL VSWR, $\text{Cos}\theta$

RELATIVE FIELD AZIMUTH PATTERN

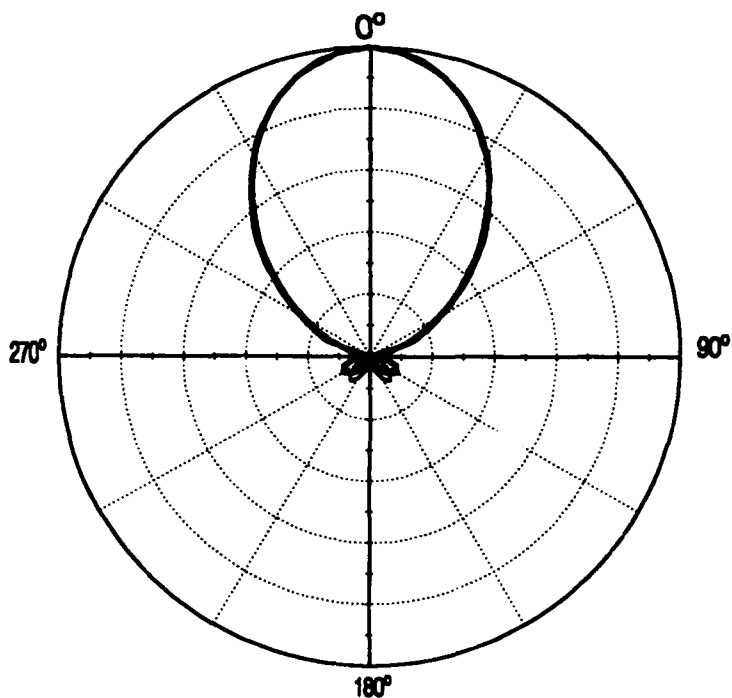


FIG. 9: $\cos^2\theta$ PANELS, 174 & 230 MHz

RELATIVE FIELD AZIMUTH PATTERN

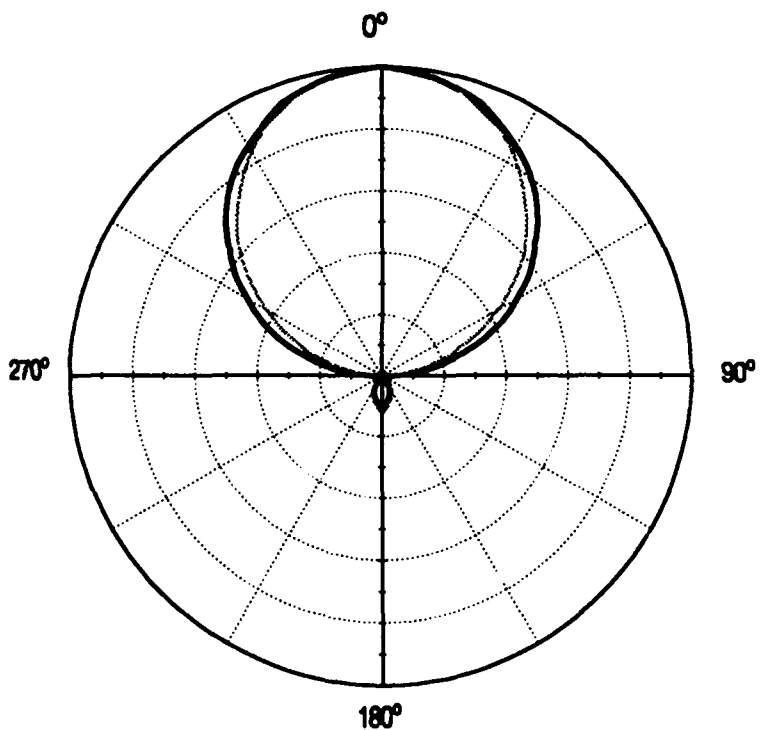


FIGURE 10: SINGLE $\cos\theta$ PANELS, AT 174 & 230 MHz

4.0 ARRAY MEASUREMENTS

The Deltawing panel can be used in an almost infinite number of array options. The number of panels, orientation, phase and amplitude can all be varied to provide many useful patterns. This feature will allow each station to tailor the resulting pattern to meet geographic and demographic needs. Figure 11 shows calculated array patterns using measured element patterns for several of these configurations. Because pattern changes as a function of frequency are minimal for the Deltawing antenna, more than one channel can be multiplexed into the antenna with nearly identical pattern characteristics.

In addition to pattern requirements, a low antenna input VSWR will be required to minimize problems with the transmitted signal³. Typical NTSC specifications call for 1.08 maximum at visual and chrominance carrier with the rest of the 6 MHz channel at 1.10 or less. HDTV systems should have the same VSWR specifications or better.

Through use of properly matched elements, feed cables and power dividers, and quadrature phasing techniques, an extremely low antenna input VSWR can be achieved over a wide frequency band. Figure 12 shows the measured return loss at the antenna input for an array of 24 Duplet Highband panels.

RELATIVE FIELD AZIMUTH PATTERN

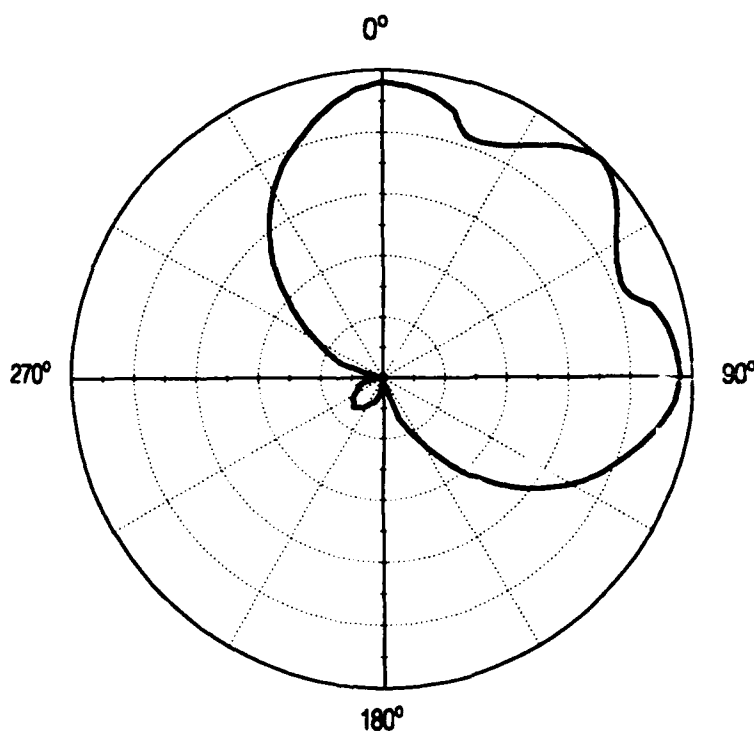


FIGURE 11A: $\text{COS}^2\theta$ PANELS AT 0° & 90°

RELATIVE FIELD AZIMUTH PATTERN

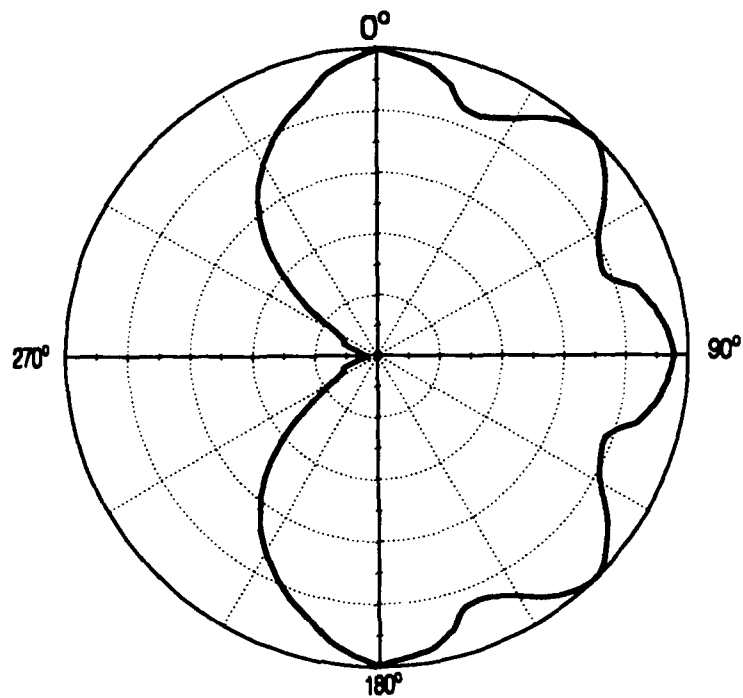


FIGURE 11B: $\cos^2\theta$ PANELS AT $0^\circ, 90^\circ$, & 180°

RELATIVE FIELD AZIMUTH PATTERN

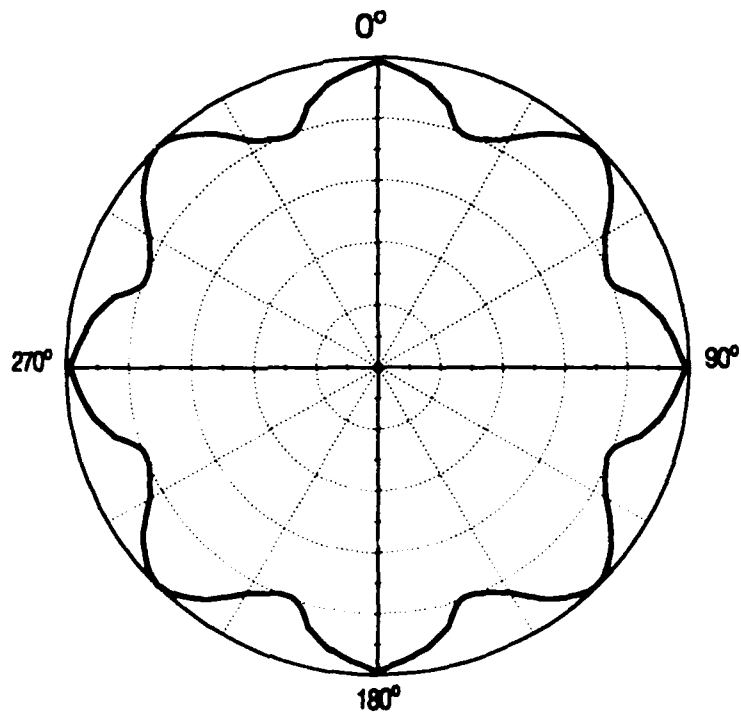


FIGURE 11C: $\cos^2\theta$ PANELS AT $0^\circ, 90^\circ, 180^\circ$ & 270°

RELATIVE FIELD AZIMUTH PATTERN

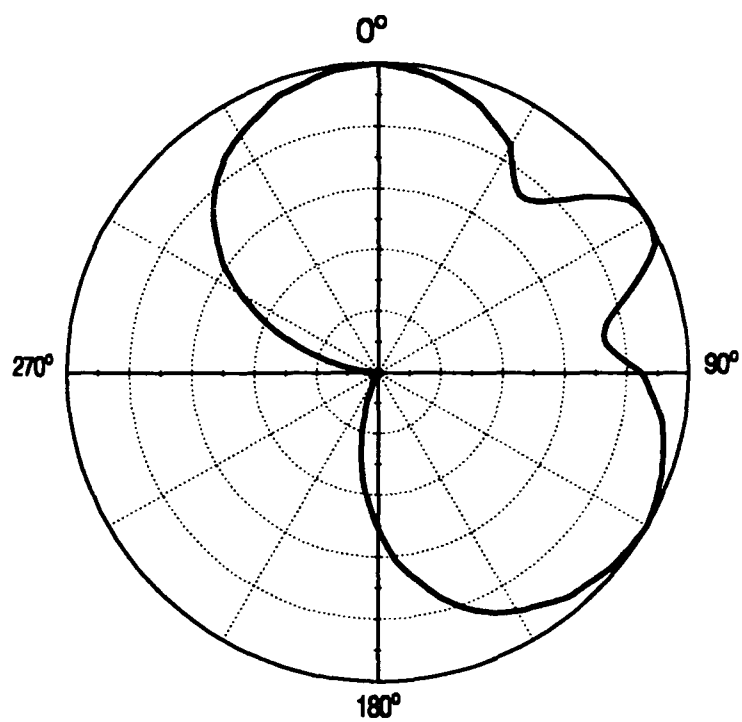


FIGURE 11D: COSθ PANELS AT 0°, 120°

RELATIVE FIELD AZIMUTH PATTERN

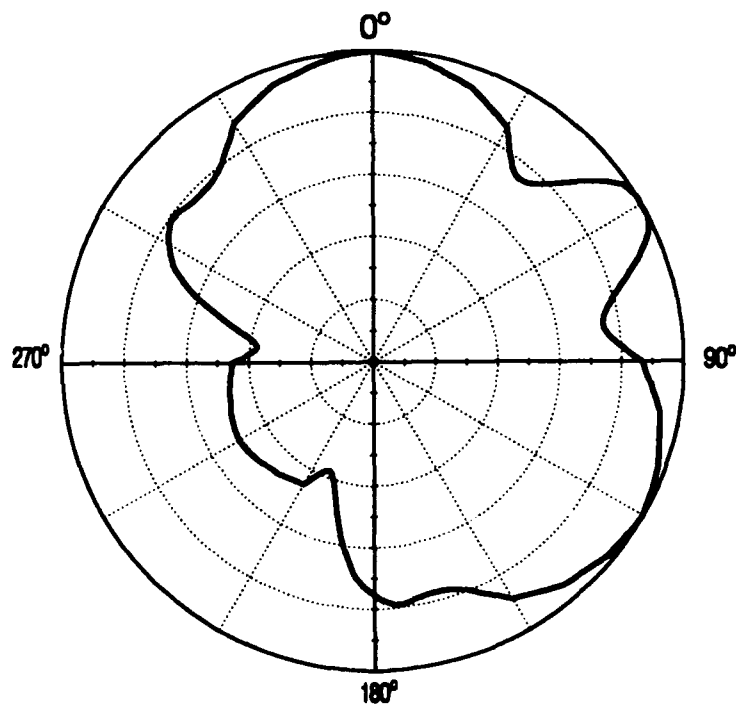


FIG. 11E: COSθ PANELS AT 0°, 120°, & 240° (25% POWER)

RELATIVE FIELD AZIMUTH PATTERN

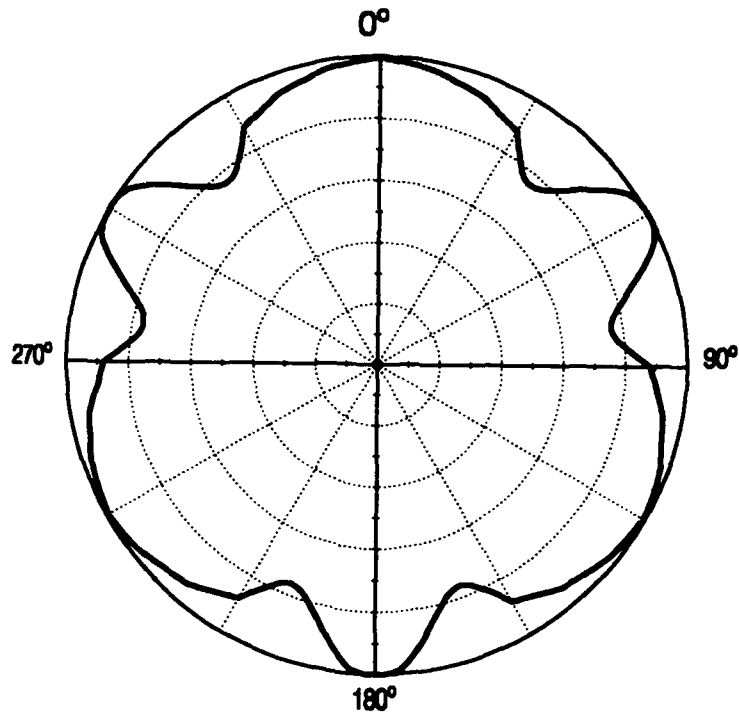


FIGURE 11F: COSθ PANELS AT 0°, 120° & 240°

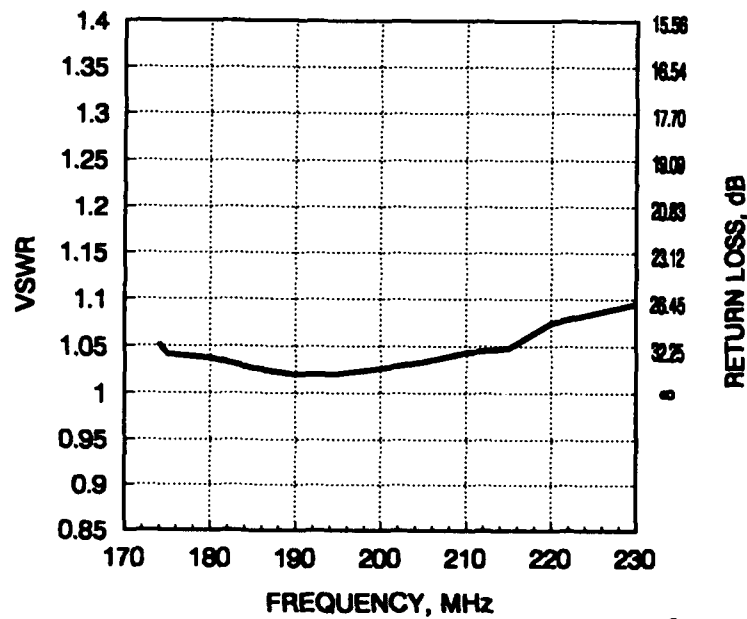


FIGURE 12: MEASURED SYSTEM VSWR, COS²θ

5.0 POWER HANDLING

When more than one channel is multiplexed into the same antenna, high average and peak power handling can become a significant issue. Transmitter powers of 30 to 60 kilowatts may typically be used for each channel. The Deltawing antenna's branch feeding network distributes 30 to 60 kilowatts of power to individual radiators. Power to individual panels is small enough to allow 7/8" 50 Ohm EIA connectors to be used. The entire feeding system is pressurized to ensure high average and peak power rating.

6.0 PANEL MOUNTING ARRANGEMENT

The Deltawing, with panel type construction, is ideal for side mounting on new or existing towers. The backplane construction and mounting configuration minimize the effect of the tower structure on array patterns and impedance. The low windloading of the Deltawing panel minimizes tower loading, allowing cost-effective implementation on existing tower structures.

7.0 CONCLUSIONS

- 7.1 The NTSC-HDTV simulcast period will require operation of two channels instead of the one now operated.
- 7.2 Multiplexing signals into a common antenna on existing but currently loaded towers will be a viable option to new tower construction.
- 7.3 With its wide impedance, wide bandwidth, and high average and peak power handling capability, the Deltawing panel antenna is matched to HDTV requirements.

8.0 ACKNOWLEDGEMENTS

I would like to thank the Engineering and Marketing department personnel at Harris Allied Broadcast Division for their assistance in preparing and reviewing this paper.

REFERENCES

- 1. Niekamp, R.L. (1992) Preparing antenna systems for NTSC and HDTV Simulcasting, 1992 NAB HDTV World Conference Proceedings, Las Vegas, NV.
- 2. Masters, R.W. (1946) The Super-turnstile Antenna, Broadcast News, 42, January, 1946.
- 3. Plonka, R.J. (1992) Comparing NTSC Transmitter specifications with requirements for HDTV, 1992 NAB HDTV World Conference Proceedings, Las Vegas, NV.

NECESSARY AND SUFFICIENT CONDITIONS FOR PORTABLE, RELIABLE, AND USEFUL CEM SOFTWARE

Edmund K. Miller
Los Alamos National Laboratory, Group MEE-3
PO Box 1663, Los Alamos, NM 87545

1.0 ABSTRACT

During the nearly three decades of the existence of what has come to be called computational electromagnetics (CEM), the volume of CEM-related software developed worldwide has grown at an exponential rate. This growth is demonstrated by the increasing numbers of papers involving CEM that appear in the literature and the proliferation of periodicals and meetings devoted to the topic. Because of the wider availability of computers and support of CEM in the United States, it is probably safe to conclude that US activity constitutes at least half of the total effort expended to date in developing and applying CEM models, although activities in other countries seem to be gaining rapidly. Although that situation might be viewed as positive, it only emphasizes the lamentable sparsity of the present CEM software base, compared with what might have been logically expected to evolve from the resources thus far invested in CEM. In this discussion, we speculate on the ingredients that might be missing from the approach that has been followed by CEM funders, developers, and users, and then consider in more detail what the goals of future software development should be if a longer-lasting payoff is to be expected.

2.0 INTRODUCTION

"The most important aspects of an EM computer code are the formulation, approximations, and numerical implementation on which it is based, and the accuracy that it provides." This observation might be made by most code developers, and it demonstrates a *how-* or process-oriented perspective of CEM. Code users, on the other hand, might observe that "the most important aspects of an EM computer code are the effort required to prepare input data and verify its correctness, the complexity and size of problems that a given code can solve, the computer resources it requires, and the accuracy that it provides." This observation demonstrates a *what-* or results-oriented perspective of CEM. Because these two viewpoints can be so disparate, it is not surprising that developers and users of EM modeling software might have different priorities and needs concerning how codes should be designed. Our purpose in this discussion is to examine some of the more important issues in CEM, as the author perceives them, in terms of developer and user priorities, focussing on the interplay of analysis and computation.

But before addressing developer and user issues, we must acknowledge that a third viewpoint might be expressed as "the most important aspect of any CEM software project is that it solve with acceptable accuracy a problem of current interest." This is an expectation that might be articulated by a sponsor or funder of electromagnetics R&D whose concern is not necessarily the software that might originate from a project but the project outcome itself. Our discussion begins with this third perspective, the funder's viewpoint of CEM software, which in the US, it should be noted, is primarily the federal government.

3.0 EM CODES FROM A FUNDER'S PERSPECTIVE

Based on my own experience, which is possibly not representative, some funders of CEM software development are interested in extending modeling capabilities beyond what's needed to solve their immediate problems. Unfortunately, however, that attitude does not seem to characterize the typical government approach where instead software developed under contract is most often regarded as a means to an end, rather than as comprising part of the end product itself. The result is that much of the investment that has been made in modeling software over the years has been lost. In the mid-1970's, we at Lawrence Livermore National Laboratory, in connection with a Computer Code Newsletter activity that was funded for a time by Defense Nuclear Agency, attempted to initiate an expanded activity to address this problem. Although we gained some support for the concept we were proposing from various government agencies, ultimately we were unsuccessful. More recently, attention has again been directed to the general problem of CEM software, by among other *ad hoc* groups the ElectroMagnetic Code Consortium (EMCC). Perhaps the time is more appropriate now for finally doing something significant to make CEM software development a better investment from the funder's perspective.

There seem to be essentially three issues of concern regarding the present status and future trends in CEM software development [as articulated at an informal workshop held at the 1992 AP-S Symposium attended by Robert Mallioux (Phillips Laboratory, Hanscom AFB, MA); Lou Libello (Harry Diamond Laboratory, Adelphi, MD); Larry Martin (Vulnerability Analysis Laboratory, White Sands Missile Range, NM); Helen Wang (Naval Air Weapon Center, China Lake, CA); Andreas Cangellaris (U. of Arizona, Tucson); and the author]. The following assessments were reached there:

- 1) There is a continuing need to develop the technology required to solve larger, more complex problems using both current and anticipated future computers, especially those computers employing parallel designs;
- 2) There is an overall need for a modeling environment or methodology that provides access to the collection of useful CEM models, with a set of supporting utilities, such as graphical

interfaces, validation and error-sensing routines etc., all from a common and standardized front end; and

3) There is a need to address the requirement that software produced on US government contracts (1) will satisfy well-defined design, documentation, and delivery specifications and standards, and (2) that a mechanism will be implemented for continuing maintenance of that software, using as two extremes of potential approaches, the NEC/GEMACS^{1,2} models on the one hand and commercial software marketers on the other.

Clearly, some government sponsors will continue to focus on item 1 because of their immediate needs, for example, in Low-Observables (LO) technology. Other sponsors may have different problems, ranging from accelerator technology to medical applications to MMIC (millimeter integrated circuits) technology; there are, of course, many others. As long as each separate funding agency continues to address only its own specific problems, the leverage and commonality from which all would benefit were a broader view to be taken will be lost.

Item 2 highlights the conclusion reached by some of us in the CEM community that a more far-reaching perspective is required if we are to progress toward the goal of achieving a capability in the modeling world that parallels what is now available in the experimental world. The status of most present modeling software is that the user/electromagneticist must become well-acquainted with the characteristics peculiar to a given model and computer environment. In other words, not only must they understand the EM functionality of the CEM model to be used, but they also must become sufficiently familiar with its software and hardware design to use it at all, let alone to use it effectively and reliably. Contrast this situation with that in electromagnetic experimentation, where the experimentalist does not need to know how to design a spectrum analyzer or network analyzer, for example, but only what these instruments can be used for and what their limitations are. Although it always helps to have a thorough understanding of all the tools being used, modern technology has become too complex for that to be a realistic expectation. Instead, computer power should enable the electromagneticist to concentrate on electromagnetic physics rather than computer programming. It must be obvious that the computer should be a tool for computational electromagneticists, i.e., a means to an end, not the end itself. To me, this means that we need to pay much more attention to designing an appropriate and standardized interface to CEM models than we have thus far.

The importance of item 3 cannot be overstated, because if we fail to address this concern, progress in the other two areas will have much-reduced benefit. Of the innumerable EM modeling codes developed over the years, it is useful to identify three main categories. The first, and probably most numerous category, is comprised of those

codes developed for a narrow, specific purpose, for example, for obtaining results for a Ph.D. dissertation. Most such codes are not retained for later use, having been written to achieve certain research goals.

A second, still sizable category consists of codes having a somewhat broader application, such as those associated with some ongoing research and development project (for example, radar scattering, wire-antenna design, or large-reflector analysis). While such codes might achieve an extended lifetime, they often become company or organization proprietary. Consequently, they are usually unavailable to outside organizations, oftentimes even to the government agencies that have supported their development.

A third code category, by far the smallest of the three, includes those codes that have become recognizable "products" because of the ongoing support provided by some government sponsor(s) or because they have become the products of commercial enterprises. The former type might be typified by NEC (Numerical Electromagnetics Code)¹ and GEMACS (General ElectroMagnetic Analysis of Complex Systems),² and the latter by Maxwell (Ansoft Corporation), EM-Sim (EESof), or Ansys (Swanson Analysis Systems, Inc.).³ These examples exist because they have received ongoing development, validation, and application by a wide circle of users. Furthermore, they have become widely used because:

- 1) the codes are readily available;
- 2) they are well documented;
- 3) they receive regular updating and fixing; and
- 4) user support is provided.

These attributes are true of a sharply decreasing number of codes, where the attributes are ranked here in decreasing order of occurrence from 1 to 4. It should be noted that a code library or distribution center can offer only item 1; the other services must be provided by the developers. Note also that a code distribution center operates in much the same way as a library. A library stores the books, but the expertise concerning a book's contents remains, for the most part, the province of a book's author(s). Without a concerted effort by funders, either to ensure ongoing support for some of the key codes they have instigated, or to actively promote their transfer to private industry where a profit motive will influence their survival, it is highly probable that the U.S. government will never reap the benefits it should from its investments in CEM. Items 1-4 listed above are necessary, and probably sufficient conditions as well, for beginning to develop portable, reliable, and useful CEM software.

Having looked at the issue of CEM software development from the more

general strategic viewpoint of government funding agencies, we now turn to the more specific issues of how CEM models might be characterized from the viewpoint of the developers and users, with the most attention being given to the latter.

4.0 EM CODES FROM A DEVELOPER'S PERSPECTIVE

4.1 Fidelity, Operation Count and Complexity

In CEM, as in other kinds of real-world phenomena, a continuous physical reality is represented by a discretized, sampled, numerical approximation intended to replicate behavior of the actual problem to some appropriate accuracy and resolution, originating from some appropriate analytical starting point. A small number of generic steps can be identified in developing essentially any physical model, as outlined in Table 1 (Ref. 4). In electromagnetics specifically, the physical principle on which CEM models are based begins with a field "propagator" as a means of developing source-field relationships as shown in Table 2. Additional factors affecting decisions in model development are the intended application and the problem type for which the model is intended, as also summarized in Table 2.

The *fidelity* (F) with which the numerical replication of the physical problem is accomplished involves a solution effort, or *complexity* (C), which is driven by:

- 1) the number of *spatial samples* (X) used;
- 2) the number of *numerical operations* (N), or FLOPs, required per sample in the initial solution process;
- 3) the number of *bits* (b) required per operation to maintain roundoff error at or below the intrinsic modeling error; and
- 4) the number of *additional operations per right-hand side* (S) needed to acquire the total information sought.

Together, these factors allow the characterization of a model application, in terms of the run time on a given computer as affected by the computation bandwidth, and they suggest that a figure-of-merit be defined for using a particular model on a given problem, as done below.

4.2 Running Time and Computation Bandwidth

The elapsed *time* (T) needed to perform the total number of *bit operations* (*BitFLOPs* or *BLOPs*) involved in the computation is determined by the effective throughput or FLOP rate R of the computer used and its word length W relative to b , whose product we call the *computation bandwidth* ($B = RW$). The specific dependence of T on these factors is a function of both the formulation and algorithm (or code) and of the computer used for the computation, and it increases with increasing X , N , b , and S , and decreases with increasing R and W , for which the simplest functional

form might be

$$T \sim (XNb + S)/RW \sim C/B, \quad (1)$$

where the numerator terms are all possibly functions of *frequency* (*f*). Because there is a practical acceptable upper limit on *T* that depends on the application and value of the numerical result, an ongoing challenge in CEM is to find ways to reduce *X*, *N*, *b*, and/or *S* without sacrificing acceptable fidelity and by exploiting computers with larger values of *R* and *W*.

4.3 Figure-of-Merit

We might also define a *figure-of-merit* (*M*) to characterize application of a particular code to a given problem, as

$$M \sim F/C, \quad (2)$$

where the code that maximizes *M*, everything else being equal, would be the best choice.

5.0 EM CODES FROM A USER'S PERSPECTIVE

The characteristics described in Section 4 are important to a user only insofar as they influence or limit the results that can be productively obtained from a given code. From a result's-oriented perspective, the properties that determine the value of a computer code might be stated in many ways, but they ultimately reduce to the following basic attributes:

ACCURACY/FIDELITY--To what extent can the code produce results on which design decisions can be based?

EFFICIENCY/PRODUCTIVITY--How expensive is the code to use in terms of both the computer resources needed to run it and the human resources needed to use it?

UTILITY/APPLICABILITY--In terms of complexity and size, what problems can be reliably modeled using a given code?

In some respects, the user of a computer code who needs to do EM numerical modeling and simulation is in a position similar to that of an experimenter who needs various kinds of hardware to perform measurements. Both understand the needs of their particular problems in terms of the basic EM phenomena involved, but both also need to depend on complex tools designed by others to accomplish their own particular goals. The experimentalist doesn't need to know how to design a spectrum analyzer, but rather, needs only to understand its functionality to use it productively, and, analogously, a code user shouldn't need to know *how* the code works but rather *only what* it can do. In another context, a person needs only a well-defined, yet limited, subset of knowledge about how an automobile works in order to be an

effective driver. Many of the improvements being made by the automotive industry today are designed to improve automobile performance, yet they require no further expertise on the part of the driver, i.e., the *how* part of the improvement is transparent to the user. In a fashion analogous to EM experimental hardware or automotive systems, user-oriented or user-friendly, EM modeling software should also be designed to improve modeling performance without requiring users to be experts in the underlying modeling principles. In the next sections, we consider some of the ways by which this might be accomplished.

5.1 Code Design for User Friendliness

For modeling software to be more applications oriented, it should exhibit the attributes of accuracy, efficiency, and utility in some appropriately optimized mix relative to the users and applications for which it is intended. Although utility will be very application sensitive, accuracy and efficiency are more generic qualities whose characteristics can be described essentially independent of the formulation and numerical treatment used for any given computer code. Summarized here are some of the more obvious ways by which these individual attributes might be improved.

5.2 Improving Accuracy

There are several ways to improve the overall accuracy of the modeling process, including the following.

5.2.1 Verifying Code Correctness--The first problem encountered by anyone using "imported" codes, or even codes that they have written that have been ported to a new machine or when some changes have been made to the system they normally use, is verifying that the code produces the same results as it previously did. This is now usually done by comparing the results of certain test problems with past runs whose results are stored, most often as computer printouts or in various reports, in graphical or tabular form. If differences are found, it may become necessary to backtrack through the code, checking various stages of the computation for errors. Often, this must be done without the original, correct results because, usually, the only results considered worth saving are the final answer. While clearly a useful indicator of an error somewhere in the computation, the final answer alone does not necessarily point to its source. It would be much more helpful to keep results from intermediate steps in the computation. Furthermore, these results should be provided not only in the hardcopy documentation, but also electronically as part of the code itself, so that anyone receiving the code will have all the test cases in an easy-to-use form. Finally, as part of the code's "self-validation" software, another option should be built-in that enables the user to automatically compare the various checkpoints of the present computation with the prestored values to obtain a set of diagnostic results that will indicate where the likely problems occur. Such a comparison would require only one run to identify the problem

unless there are multiple errors.

5.2.2 Verifying Input Data--After becoming convinced that the code is running as designed, the next problem faced by most users is whether or not the input data are correct for a new problem. There are actually two parts to this question: one is how to avoid violating the constraints and limitations inherent in any code, and the other is how to ensure that the input data is entered as actually intended. In both instances, accuracy could be substantially assisted by including various kinds of error-trapping procedures in the code. The former would include checks such that preclude user-selected modeling parameters from inadvertently exceeding any known bounds and to ensure their conformance to recommended modeling guidelines. Some of these guidelines are summarized in Table 3 (Ref. 4) below. Included here, for example, might be checks to ensure that physically adjacent patches or segments are electrically connected for models that employ multi-zone basis functions. As a further aid to the user, the code might have a number of the more commonly encountered geometries included in a model library so that repetitive labor can be avoided and input errors can be reduced. By helping the user conform to code requirements and to avoid the more predictable input errors, both the accuracy and efficiency of modeling will be enhanced.

5.2.3 Validating the Computation--Both code users and developers would agree that validation of the computed results is essential to developing confidence in any model. Developers tend to consider that this essential step is largely accomplished when some of the more generic analytically solvable cases have been checked. Users, on the other hand, will more often attempt to solve problems for which analytic solutions are unavailable and for which experimental data may be an unreliable validation source. Therefore, validation is a continuing requirement, especially as larger and more complex problems are attempted, and it is an activity in which analysis and computation have, perhaps, their most synergistic interplay. Although the errors that occur in CEM can be defined in various ways, they can be assigned to the two general categories of physical modeling error and numerical modeling error, as defined in Table 4 below.⁴

One source of validation results is the various analytically based internal checks that can be performed within literally any modeling code. For moment-method models, for example, these checks might include evaluating how well boundary conditions are satisfied across penetrable interfaces or at conducting surfaces. While these are the most fundamental checks that might be performed (since they are not only necessary but also sufficient properties of a valid solution), they are also generally the most expensive. Internal checks that are useful at other stages of the computation include computing the matrix condition number and checking the equation error after the system matrix has been factored or inverted. Other checks, necessary but not sufficient to ensure

model accuracy, are various reciprocity and power-balance tests. All such internal checks could be options, available in any code, to be exercised at the user's discretion, either when modeling a new problem or for troubleshooting routine applications when it appears necessary. If these kind of checks were conveniently available, the user could decide when increased confidence in the model's validity would be worth the incremental computer cost that these checks impose. Some of the available internal checks are summarized in Table 5.

An even larger selection of external checks might be used in validating a given computation. Whether these checks are experimental, analytical, or computational, an important issue in expediting their use is developing and presenting the results in terms of sampling and data formats that simplify their comparison with other data sources. This is a topic of such breadth and importance that it requires a more in-depth discussion than can be accommodated here, although some aspects of using external checks are summarized below in Table 6.

5.2.4 Interpreting the Output--As computers grow in power to make possible the solution of ever larger and more complex problems, the sheer volume of output drastically increases the need for better ways of accessing the model results. While it is a user-efficiency issue, flexible output presentation is also an issue of accuracy, as without it, the user might miss important aspects of the results that provide clues about nonphysical behavior or other indicators of erroneous results. On the other hand, correct results might be misunderstood if the user is unfamiliar with what is a correct, but unexpected, solution behavior. One way of improving output interpretation is to make more use of graphical presentations, as these are tremendously easier to evaluate than are the equivalent information presented in numerical form. Another is to provide the user community with a solved-problem data base whose expansion over time would increase the probability that already-solved problems might be available to a user that provide clues about how to model a new problem.

5.3 Improving Efficiency

Model efficiency is determined by two factors: 1) the computer resources, and 2) the user effort needed to model a given problem. The former can always exceed the latter if problem complexity becomes large enough, but for many routine modeling applications, the cost of user effort can exceed, or even dominate, the total money cost of modeling.

5.3.1 Computer resources--Users usually have little control over this cost component other than varying the number of samples or unknowns used for the model, and even then they have little quantitative expectation for how the model fidelity is affected by changing the sampling rate. At a minimum, a code might include the capability to show the estimated accuracy versus the time/storage cost to enable the user to make a more explicit tradeoff between them. Even better, a code

would permit the user to specify the accuracy desired so that the numerical model description can then be automatically developed from the previously given geometrical/electrical description. Perhaps the most important way for the user now to control cost is to select, from two or more codes that may be applicable to a problem, the one that offers the best figure of merit. For example, all things being nearly equal, a code having a reliable iterative-solution procedure should always be more efficient for antenna (single right-hand side) problems than one which does not. More efficient codes, still to be developed, will probably employ hybrid and adaptive techniques, having as their overall design goal that of reducing the operation count for any given application.

5.3.2 User effort--This category includes all activities required of the user to prepare the input data, monitor code operation, and utilize the output. An interactive, graphics-based interface is the most efficient way for these tasks to be performed. Input devices such as digitizer tablets and graphical displays not only increase the efficiency of entering the problem description, but improve the accuracy as well. Graphical output, using various devices such as color monitors and hardcopy printers, provides thought-enhancing access to the electromagnetic physics being modeled. If users can examine the input and output from a variety of perspectives and in various formats, and manipulate those presentations easily and rapidly, the effort required will be immeasurably less, especially when questions arise and clarification is needed.

These operations should be accomplished using a *standardized* interface through which the *physical* problem description is developed independent of a specific code. Code-specific translators would then develop the *numerical* model description needed to run a particular code. Among other benefits, this approach would ensure that all codes used for a given problem would start from a common data base. It would also mean that different models could be compared more easily by translating the numerical-model results back onto the common problem description. In addition, users could acquire equivalent facility with a wider variety of codes since all of the codes would use the same interface. Ultimately, this could become an EM Modeling Environment (EMME) in which all the codes are accessible as part of the same package.

5.4 Improving Utility

Utility refers to all factors other than accuracy and efficiency, specifically including the complexity and variety of problems for which a code can be used. The basic information needed by the user is help in identifying which of those codes resident in EMME might be most appropriate for a given application. Such information might be provided within EMME by creating a matrix of available modeling capabilities versus the various codes it contains as part of the documentation, or as a menu-based query system that can be exercised by the user.

6.0 THE BOTTOM LINE

It's clear that computational methods in EM will continue to develop and grow in importance regardless of whether any of the features discussed above are ever incorporated. It's equally clear however, that the overall value of CEM is ultimately limited by the accuracy, efficiency, and utility of the available codes, whose improvement relies on the intersection of analysis and computation. If the research and analysis aspects of code development are eventually to be realized in routine application and design, user-oriented EM modeling software is needed that provides a robust computational environment for EM engineers who are not, themselves, code developers. Only by providing

- 1) a single interface and problem-description language giving access to a variety of different codes;
- 2) built-in input-data checking and error trapping;
- 3) user-selectable optional validation checks;
- 4) graphical and interactive output displays; and
- 5) a convenient way to compare the results of different models,

will CEM become more accurate and efficient for all users. This can happen only if the funders, developers, and users of CEM software work in concert to bring about these advances. Of these three constituencies, the funders have the most important role because their support determines the directions taken by the rest of the CEM community.

7.0 REFERENCES

1. Burke, G. J. and A. J. Poggio (1980), *Numerical Electromagnetics Code (NEC -- Method of Moments*, NOSC TD 116, Naval Ocean Systems Center, San Diego, CA 92152, revised, (NEC-2).
2. Kadlec, D. L. and E. L. Coffey (1983), *General Electromagnetic Model for the Analysis of Complex Systems (GEMACS) User Manual (Version 3)*, RADC-TR-83-217, Rome Air Development Center, Griffiss Air Force Base, NY 13441. ADA137461
3. Cendes, Z. J. (1990), "EM simulators = CAE tools," *IEEE Spectrum*, November.
4. Miller, E. K. (1988), "A Selective Survey of Computational Electromagnetics", Invited Tutorial-Review Paper, *IEEE Transactions Antennas and Propagation*, September, 1281-1305.

Table 1
CLASSIFICATION OF MODEL TYPES IN CEM

<u>FIELD PROPAGATOR</u>	<u>DESCRIPTION BASED ON:</u>
Integral operator	Green's function for infinite medium or special boundaries.
Differential operator	Maxwell Curl Equations or their integral counterparts.
Modal expansions	Solutions of Maxwell's Equations in particular coordinate systems and expansions.
Optical description	Rays and diffraction coefficients.
<u>APPLICATION</u>	<u>REQUIRES:</u>
Radiation	Determining the originating sources of a field.
Propagation	Obtaining the fields distant from a known source.
Scattering	Determining the perturbing effects of medium inhomogeneities.
<u>PROBLEM TYPE</u>	<u>CHARACTERIZED BY:</u>
Solution domain	Time or frequency.
Solution space	Configuration or wavenumber.
Dimensionality	1D, 2D, 3D.
Electrical properties of medium and/or boundary	Dielectric; lossy; perfectly conducting; anisotropic; inhomogeneous; nonlinear.
Boundary geometry	Linear; curved; segmented; compound; arbitrary

Table 2
STEPS IN DEVELOPING A COMPUTER MODEL

<u>STEP</u>	<u>ACTIVITY</u>
Conceptualization	Encapsulating observation and analysis in terms of elementary physical principles and their mathematical description.
Formulation	"Fleshing out" of the elementary description into a more complete, formally solved, mathematical representation.
Numerical implementation	Transforming into a computer algorithm using various numerical techniques.
Computation	Obtaining quantitative results.
Validation	Determining the numerical and physical credibility of the computed results.

Table 3
GENERIC GUIDELINES FOR IE MODELING

MODELING PARAMETER OR ISSUE	NOMINAL RANGE OR VALUE	TYPICAL REASON
Wire Length, L	$L > 10d$	Neglect of end caps in thin-wire treatment.
Wire diameter, d	$\lambda > \pi d$	Neglect of circumferential effects.
Wire segment length, Δ As related to diameter	$\Delta > d$	Use of thin-wire kernel in integral equation. Can be relaxed by use of extended kernel ¹
As related to wavelength	$\Delta < \lambda/2\pi$	Necessity of sampling current densely enough per unit wavelength.
Step change in wire radius, δa	$\Delta > 10 \delta a$	Neglect of sources on stepped surface (similar to end-cap problem).
Source location	Do not place on open-ended segment	Avoids non-physical situation of driving wire at open end.
Angle of wire bend, α	$\alpha > 2\pi a/\Delta$	Keeps adjacent wires from occupying too large a common volume.
Axial separation of parallel wires, r With match points aligned	$r > 10a$	Avoids placing one match point (when using delta-function weights) in error field of another
Otherwise	$r > 3a$	Neglect of circumferential current variation.
Wire mesh model of solid surface Mesh size $\Delta x \Delta$	$\Delta < \lambda/10$	To reduce field "leakage" to acceptable level.
Wire radius	$a = \Delta/2\pi$	To have wire area equal to surface area of solid.
Surface-patch area, Δ_s Frequency domain Time domain	$\lambda > 2\pi\sqrt{\Delta_s}$ $\lambda_{\min} > 2\pi\sqrt{\Delta_s}$	Need to sample currents densely enough in wavelengths.
Piecewise model of curved wire or surface of radius of curvature R	$R > \Delta$ $R > \sqrt{\Delta_s}$	Necessity of sampling a circular arc at least 6 times per 2π radians.
Starting time in time-domain solution, t_{st}	$E^{inc}(t) _{\max} \geq 10^X E^{inc}(t_{st})$	To achieve numerically smooth buildup of exciting field and 10^X accuracy.

Table 3, Continued
GENERIC GUIDELINES FOR IE MODELING

MODELING PARAMETER OR ISSUE	NOMINAL RANGE OR VALUE	TYPICAL REASON
Stopping time in time-domain solution, t_{sp}	$E^{inc}(t) _{max} \geq 10^x E^{inc}(t \rightarrow \infty) - E^{inc}(t_{sp}) $ or $I(t)$ reaches steady state	Stabilize final response.
Time step in TD solution, δ	$c\delta \leq \Delta$	Satisfies Courant stability condition. Required for explicit solution.
Maximum frequency of transient source in TD solution, using Gaussian excitation, i.e. $E^{inc}(t) = \exp(-g^2 t^2)$	$g \sim 2f_{max}$	Ensures source spectrum does not exceed upper frequency for model validity.

Table 4
ERROR TYPES THAT OCCUR IN COMPUTATIONAL ELECTROMAGNETICS

CATEGORY	DEFINITION
Physical Modeling Error, ϵ_p	Arises because the numerical model used is normally an idealized mathematical representation of the actual physical reality.
Numerical Modeling Error, ϵ_N	Arises because the numerical results obtained are only approximate solutions to that idealized representation, and is comprised of two components: 1) Solution error--The difference that can exist between the computed results and an exact solution even if the linear system of equations is solved exactly, because a finite number of unknowns has been used; and 2) Equation error--The equation mismatch that can occur in the numerical solution (1) because of roundoff that occurs in finite-precision computations, or (2) when using an iterative technique because of limited solution convergence.

Table 5
INTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION
VALIDITY

MEASURE	EXAMPLE	TESTS	PROPERTIES
CONVERGENCE MEASURES			
Local	$\lim[I(s)], \lim[E(r)]$ as $N \rightarrow N_{\max}$	Convergence of input impedance, current, fields, etc.	Reasonable measure of solution behavior, but can yield non-monotonic result.
Global	$\int I(s)I^*(s)ds$ or $\int E(r)E^*(r)dx^n$	Convergence over entire object of current or convergence of field in $n=1,2$, or 3 dimensions.	A more complete measure of convergence.
Random (local or global)	$\Sigma F(r_n)$, with $F(r_n)$ a field quantity which is function of a random variable r_n .	Convergence of any field quantity measured by a random observation variable.	Permits estimation of convergence and uncertainty of convergence estimates.
OTHER MEASURES			
Power Balance	$P_{in} + P_{loss}$ $= P_{radiated}$	Whether supplied power equals sum of radiated plus dissipated power.	Provides good check on antenna source model for radiation resistance. A necessary, but not sufficient condition.
Boundary- Condition Matching	$E_{tan}(r) = 0$, r on object modeled	Degree to which specified conditions on the boundary are satisfied.	Most fundamental check on solution. Consistency requires use of same weight function as for model itself. Can be computationally expensive. Necessary and sufficient condition.
Reciprocity	$E(\phi_1^{inc}, \phi_2^{scat})$ $= E(\phi_2^{inc}, \phi_1^{scat})$	Whether interchanging observation and source locations yields identical results.	Useful check for antenna and bistatic scattering patterns. Necessary but not sufficient condition.
"Nonphysical Behavior" of Solution	—	Whether computed results exhibit physically reasonable behavior.	Can be a subjective check. One example is provided by spatial oscillation in current when thin-wire approximation is violated.

Table 6
EXTERNAL CHECKS USEFUL AS MEASURES OF SOLUTION
VALIDITY

MEASURE	EXAMPLE	TESTS	PROPERTIES
ANALYTICAL			
Any observable provided by a formally exact solution.	Sources, near and far fields.	Any observable provided by the computer model.	Provides a necessary and sufficient condition for solution validity. Available for only special geometries, but gives canonical benchmarks.
COMPUTATIONAL			
Far fields.	Radiation pattern, bistatic and monostatic scattering pattern.	Consistency of the quantity least sensitive to solution errors.	A useful test, but one which is often subject to angle shifts between results from two models.
Near fields and sources.	Near-field cuts, current and charge distributions.	Quantities most often directly computed by model.	A more demanding test for comparison, but one which often exhibits spatial shifts between models.
Input impedance/ — susceptibility.		Source models and single-port input characteristics.	Especially sensitive measure in terms of input susceptibility. Highly advisable to examine over a range of frequencies because shifts in frequency also occur.
EXPERIMENTAL			
Same observables as used for computational checks.	—	Physical modeling error and relative correlation of actual problem with numerical model.	Perhaps the most reassuring check to make, but also often the most difficult.

Advances in Modeling and Simulation of Complex Radiating Structures and Platforms: The Role of Computational Electromagnetics*

by

Michael P. Hurst, Louis N. Medgyesi-Mitschang, John M. Putnam
McDonnell Douglas Research Laboratories
St. Louis, Missouri 63166-0516

Abstract

Advanced modern antenna systems are characterized by ever increasing complexity. These include multi-functionality, wide bandwidth, and high radiating efficiency. In addition, for arrays there is also the desired beamwidth, flexible electronic steering, and agile beam forming and shaping. In many cases, radome effects must also be considered. For airborne and vehicular applications, the finiteness and complexity of the platform on which the antenna or array is mounted must be taken into account. In general, such applications require a lightweight topologically compact radiating system.

Many of the foregoing requirements translate into conflicting constraints necessitating engineering compromises to yield cost-effective solutions. In the past few years, advances in computational electromagnetics have allowed these design compromises to be carried out through numerical modeling and simulation. This methodology complements, and in some cases, supplants traditional experimental methods. In this paper the principal, most widely used computational techniques, such as the method of moments and various hybrid approaches are summarized. The focus here is their application for large, complex radiating structures and platforms.

I. Introduction

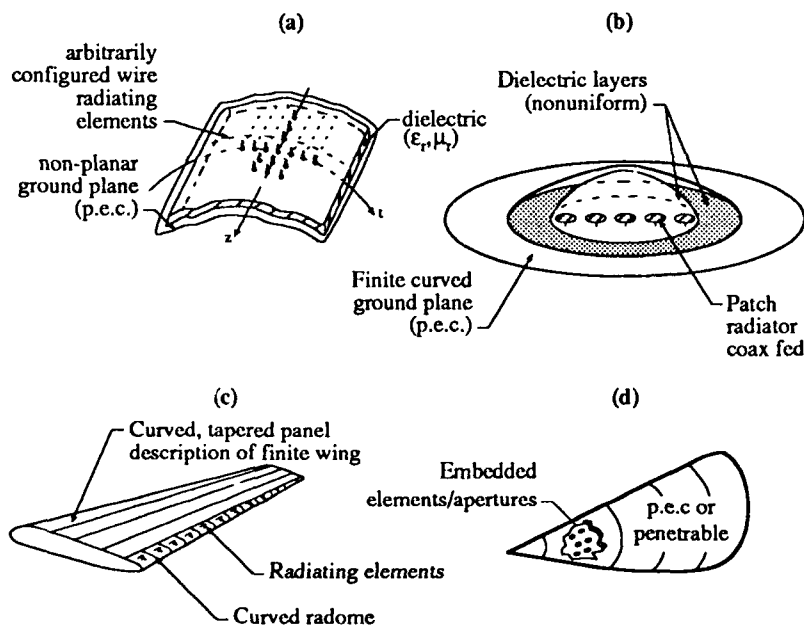
The recent rapid advances in antenna technology and associated avionics have been overshadowed by even faster-paced developments in computing power available to the antenna engineer and system designer. The ubiquity and low cost of this computing power have given rise to a myriad of developments. A recent special issue of the IEEE PROCEEDINGS is devoted to these advances^[1]. The enormity of the antenna literature is staggering. More than a thousand archival papers appear yearly in the English language. Many more appear in other publications and conferences.

The most striking aspect of these advances is that "numerical" experiments have taken precedence over traditional laboratory and range experimentation. In some cases the latter have been largely supplanted. The dominant trend is toward "full complexity" simulation of antennas, arrays, and in some cases their platforms. For certain

* This investigation was carried under the McDonnell Douglas Independent Research and Development Program.

applications, especially airborne arrays steered to the horizon, the effect of the entire aircraft must be incorporated into the analysis. It is well known that sophisticated beam steering and beam forming require inclusion of array/platform interactions. The advances in computer hardware have paced the developments in new electromagnetics analysis and software. Increased user-friendliness and availability of parallel computing architectures (SIMD or MIMD) have accelerated new developments and opened up new major uses of older methods such as the method of moments and finite element methods.

In this paper, we first outline in general terms the methods available to solve complex antenna and platform interaction problems. We highlight some of the applications of these computational methods to problems studied at our laboratory. Some of the array configurations handled are depicted in Figure 1. Specifically, the problems considered include large arrays, the impact of radomes on end-fire arrays, and antenna pattern prediction where platform (vehicle) coupling is incorporated.



92-223-246

Fig. 1 Classes of complex antenna/array problems amenable to CEM: (a) wire radiators in dielectric layers; (b) patch antennas on a finite radome enclosed ground plane; (c) airfoil embedded array with radome; (d) radiators/apertures on partially penetrable 3-D body.

II. Methods in Computational Electromagnetics (CEM)

The principal "numerical" or low frequency methods currently used in CEM are the method of moments (MM)[2], the finite element method (FEM)[3], and the finite difference time domain (FDTD)[4] technique. In addition, there are the well known high frequency techniques based on extensions of either geometrical or physical optics (GO and PO). The former was extended by Keller and other investigators[5,6,7] resulting in the geometric theory of diffraction (GTD). Subsequently, GTD was modified to its full potential as a uniform theory of diffraction (UTD)[8]. By historical coincidence, Ufimtsev[9] developed the physical theory of diffraction (PTD) contemporaneously. The domains of application of these methods were greatly expanded by the formulation of hybrid methods[10-14], providing a consistent generalizing mathematical framework. Indeed after a hiatus of several years, hybrid methods have again become an important area of research in electromagnetics. Their extension and application to antenna problems are self-evident.

Method of Moments Technique

This technique, probably more than any other, helped to usher in advanced computational methods to the antenna analysis field. A thorough discussion of this method as applied to antennas is explicated in a recent review paper by E. K. Miller[15]. The attractiveness of this method is its mathematical and numerical robustness for a great variety of very complex radiation problems. An additional virtue is that generally the MM formulation can be rapidly implemented on a computer and numerous specialized MM codes now exist for many classes of antennas.

Some recent MM advances are notable. Today, one can achieve full 3-D analysis capability with the MM approach. Building on earlier work by Rao et al. [16], the 3-D surfaces can be facetized by flat triangular or curvilinear patches[17]. While the emphasis has been largely to scattering problems, these methods are adaptable to antenna problems as well.*

Combined formulations have been developed to eliminate the spurious resonance problems experienced with earlier electric or magnetic field integral equation formulations (EFIE or MFIE). Ultra-thin substrates and superstrates, encountered in microstrip antennas and with certain radome applications, can now be handled. This was made possible by singularity extractions in the evaluation of the principal value integrals for the self-terms of the MM impedance matrix.

A complex antenna geometry or platform requires flexibility in the choice of basis functions used in the MM technique. For example, to achieve high accuracy and yet maintain computational economy, different representations must be used for wire or

* At our laboratory MM-based codes using both surface descriptions have been implemented. The CARLOS™ (Code for Analysis of Radiators on Lossy Surfaces) handles arbitrary 3-D surfaces that are p.e.c., partially or totally penetrable, and may be discontinuous, inhomogeneous, and layered.

aperture radiators and the surfaces on which they reside. Specialized representations are needed for accurate modeling of active and parasitic elements in microstrip antennas as well as at junctions. A number of formulations incorporating mixed domain expansions to address these problems have now been developed.[18-20].

Finite Element Method (FEM)

This method is well suited for analyzing antennas when complex, inhomogeneously loaded cavities are involved. In the analysis of transient phenomena, one begins with the Maxwell curl equations, multiplies each equation by a test function (not necessarily the same one), and integrates over the entire computational domain. One then expands the electric and magnetic fields in a series of basis functions. Taking these basis functions to be the same as the test functions results in an implicit system of ordinary differential equations (ODE). If one uses polynomial basis functions of degree one or less, then it is possible to write this implicit system as an explicit system, which is much easier to deal with in practice. As is often done in finite difference methods, it is preferable to stagger the placement of the electric and magnetic fields within the mesh. In external problems the computational domain must be truncated. Recently, approximate absorbing boundary conditions (ABC) have been developed for this purpose[21,22].

To analyze wave phenomena in the frequency domain, we begin with a modified form of the vector Helmholtz equation, which is satisfied by either the electric or magnetic fields. As is done in time domain, one multiplies the partial differential equation by a testing function and integrates over the entire domain. Upon applying Green's theorem, a weak form results. In the Galerkin scheme the basis functions are the same as the testing functions and a sparse linear system of equations results.

The advantages of using the FEM for electromagnetic modeling are that when using triangular elements (tetrahedrals in 3-D), modeling complex body-shapes and inhomogeneous materials is easily accomplished and utilizing sparseness in both time and frequency domains is advantageous to reduce CPU time and storage. Over the last decade, investigators have tried to use the FEM to model electromagnetic waves either as a stand-alone method or more recently merged with MM codes in a hybrid formulation[23,24] as discussed below.

Finite Difference Time Domain (FDTD) Methods

This method provides a direct solution to Maxwell's time-dependent curl equations. Conceptually, it is simple and elegant. Second-order central differencing for space and time derivatives in the curl equations is used. The FDTD is a time-marching procedure, where at each time step, the system of equations for the electric and magnetic fields is fully explicit. This avoids storing and solving linear systems of equations. The computer memory required in this technique is proportional to the number of field unknowns in a given volume analyzed and is therefore quite modest. Many variations of the FDTD scheme exist but all incorporate the Yee interleaved field discretization method.[25]

Advances in computer storage capacity and computational speed have made FDTD a viable computational technique in the last few years. There have been many analysis advances as detailed in the review paper by Taflové et al.[4]. The so-called "staircase" approximations of earlier FDTD schemes have been obviated through use of locally distorted grid models, globally body fitted grids, and unstructured grids. Also significant advancements have been made to minimize the computational domain around an object by ABC truncations[26]. While scattering has been the primary focus of FDTD investigations, recently results appeared for wire, microstrip, and aperture antennas.[27-29]

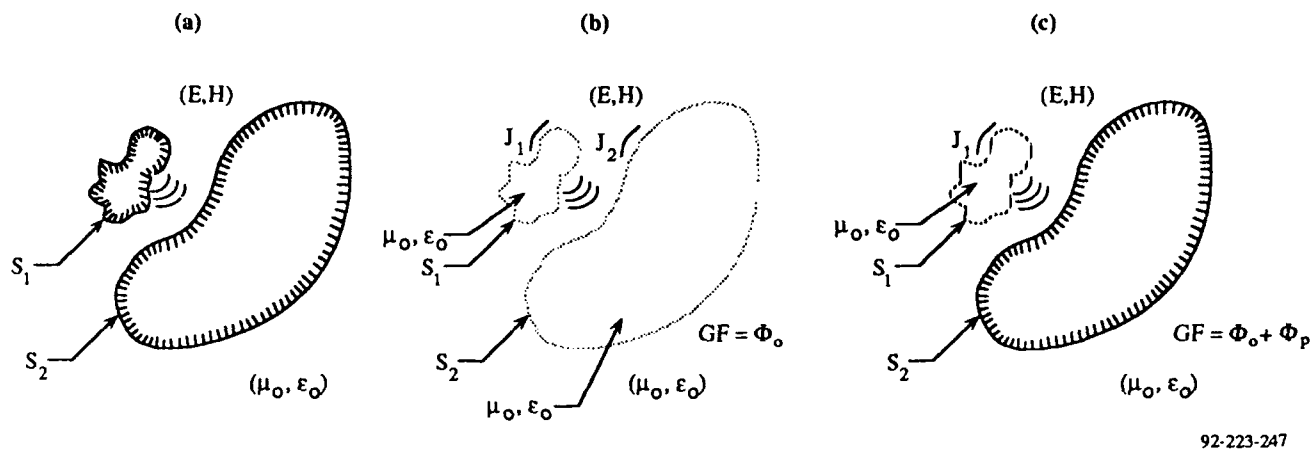
Hybrid Methods

Combining Surface Integral Formulations and Asymptotics

The earliest applications of hybrid methods were antenna problems. A recent review of this field is given by Thiele in [10] with additional details given in [11]. The earlier work in this area focused on combining the low and high frequency methods, i.e., the MM approach with the optics derived ones such as GTD, UTD, PTD, and the Fock theory. Conceptually, the radiation (antenna) problem is formulated in terms of a surface integral equation (SIE). As shown in Figure 2, the original problem can be reduced by the Schelkunoff equivalence to two sets of equivalent problems depending whether the SIE kernel is the free space Green's function Φ_0 or a perturbed one $\Phi_0 + \Phi_p$. The latter incorporates the effect of the second body. A similar equivalence can be explicated for an aperture problem as shown in Figure 3.

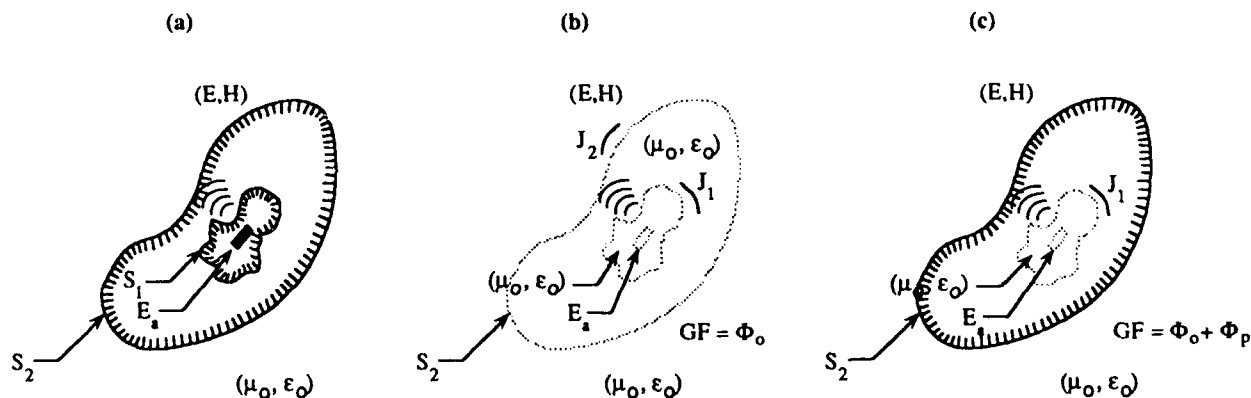
In general, the concept is to represent the complex part of the problem (i.e., the radiator, constituting S_1 in Figure 2, and the aperture and its proximal region as S_1 in Figure 3) with a MM (Galerkin) representation. The remainder of the geometry (i.e., S_2) is represented with an Ansatz. If the Ansatz is embedded in the perturbed Green's function, $\Phi_0 + \Phi_p$, the hybrid formulation is often called "field-based", since the perturbation is obtained from the GTD or UTD derived diffraction effects of the fields. On the other hand, if the free space Green's function is used in the SIE, then the Ansatz manifests itself in the PO, PTD or Fock theory derived currents used in SIE kernel spanning S_2 . This variant of the hybrid formulation is often denoted as "current-based". In both cases, the problem is so partitioned that S_2 is a surface for which canonic solutions exist either as diffracted fields from GTD/UTD or asymptotic currents from PO, PTD, or the Fock theory. There are specific considerations in choosing these Ansatzes and hence the respective partitioning of a problem. These are discussed at length in [11].

The foregoing approaches have been extensively used. A few representative examples are given later. Two important points can be made about these hybrid approaches. First, by using Ansatz currents or a perturbed Green function, the computational domain over which the MM (Galerkin) expansion is used is substantially reduced. This translates to very modest matrix storage requirements. Typical storage requirements are 10 to 100 fold less than those required if one used the MM expansion everywhere. The second



92-223-247

Fig. 2 (a) Original two-body problem - detached radiating surface S_1 ; (b) equivalent two-body problem with free-space Green's function, Φ_0 ; (c) equivalent one-body problem with perturbed Green's function, $\Phi_0 + \Phi_p$.



92-223-248

Fig. 3 (a) Original problem - embedded aperture in S_1 ; (b) equivalent problem with free space Green's function, Φ_0 ; (c) equivalent problem with perturbed Green's function, $\Phi_0 + \Phi_p$.

point is that hybrid methods are inherently well suited to massively parallel computing. Unfortunately space limitations preclude a discussion of these aspects here.

Combining Surface and Partial Differential Equation Formulations

As noted before, recent advances in partial differential equation (PDE) solvers open up new avenues for hybrid formulations. In cases where the material properties of a penetrable body change from point-to-point or there is great geometrical complexity such as in microstrip feeds and antenna cavities, SIE formulations may be inappropriate. The FEM or FDTD formulations are often better suited to these problems. In FEM, the method requires many unknowns per wavelength (typically 20 to $50/\lambda$ /linear dimension) to span the computational (volume) domain. However, the resulting matrices are sparse, unlike the fully populated matrices arising from the MM approach. One drawback is the difficulty in satisfying the Sommerfeld (radiation) condition at infinity for external problems. Radiation problems are inherently external problems.

To overcome this limitation, the foregoing problems can be recast in a hybrid formulation. For illustration, let the problem involve a cavity-backed antenna embedded in an electrically large surface. Consider again Figure 3. The exterior surface S_2 is represented by a SIE formulation, spanned by the MM (Galerkin) expansion. The interior of the cavity is represented by a PDE formulation solved by the FEM technique. The boundary between the exterior and the interior region is S_1 , spanned by both the MM and FEM basis sets. The SIE formulation incorporates the radiation boundary condition rigorously. The FEM formulation captures the complexity of the antenna/cavity region. Thus the two methods complement each other. Since S_2 may be electrically large, the MM approach also provides the more efficient representation for the electrically large region. Details on use of this approach are given in [24].

III. Examples of Application of CEM

In this section, we illustrate the uses of some of the analysis tools described above for a number of complex radiation problems. In general, the complexity of these problems requires numerical and hybrid methods.

Simulation of Large Arrays

Simulation of very large arrays poses special challenges. While present computers and modeling software often allow rigorous analysis of very large arrays of complex antenna elements, the simultaneous simulation of an array and its platform is rare. However, predictions which are sufficiently accurate for engineering purposes can still be obtained through the judicious use of simplifying approximations. For example, planar arrays on a conducting surface can often be treated using image theory and ray-tracing methods to account for antenna-platform interactions, thus obviating the need to explicitly determine current distributions on the platform. Also, a complex radiating element can usually be described in terms of a few parameters which can be used in a simplified system model.

In this way the number of unknowns required to model a large antenna array and its platform can be reduced by orders of magnitude, thus allowing numerical analysis to proceed.

Circular Planar Array - An Example

As an example of the above approach, the analysis of a 1306 element array on a circular ground plane 36.6λ in diameter will be described. (See Figure 4.) This array was built and tested by Ball Communications Systems Division as reported in [30]. Since the array was designed to be steered to endfire (i.e., form a beam near the plane of the array), extreme mutual coupling effects were anticipated, thus requiring detailed analysis to predict array performance. The antenna elements in this case were top-loaded monopoles (i.e., short wires coupled to circular disks) with a coaxial ground plane feed-through designed for matching the monopole at the desired operating frequency (4.8 GHz). The matching network does an impedance transformation and also cancels the reactive part of the monopole input impedance, which is large since the monopole is electrically small.

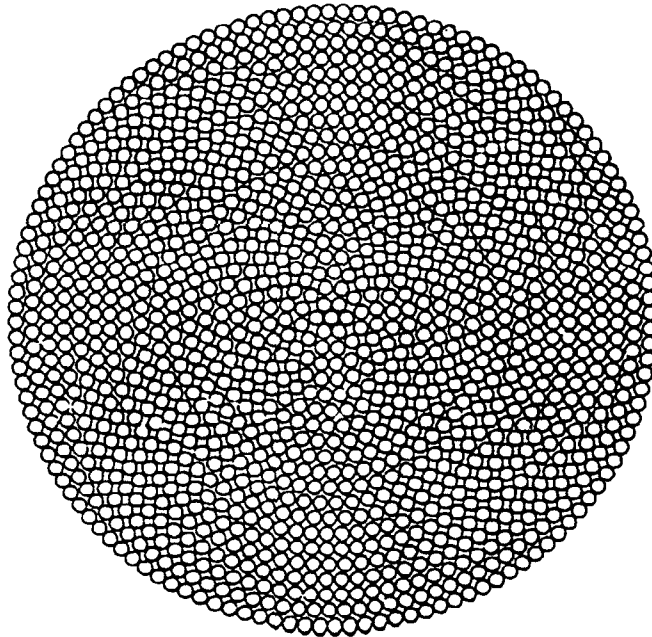


Fig. 4 Concentric ring array geometry.

The two most significant quantities characterizing an antenna element are its input impedance and radiation pattern. Since the radiation pattern of this antenna does not depend strongly on the details of how it is fed, a simplified model consisting of a disk and a grounding stub (post) was used. The element was excited by impressing a voltage across an infinitesimally thin gap between the stub and the ground plane (the so-called

delta-gap feed). To account for the effects of the matching network and source impedance, a lumped element model was used. In the MM a series impedance in a wire can be modeled by enforcing an Ohm's law relation between the total tangential electric field and the current on a small section of the wire. When subsectional basis functions are used, this is implemented by modifying a single element in the MM impedance matrix.^[19] A circuit model of the antenna/feed network and its Thevenin equivalent are shown in Figure 5. V_s and V_m are the source and mutual coupling voltages, respectively; R is the output impedance of the source; Z_{ant} is the monopole input impedance, and Z_{eq} is an equivalent impedance whose value can easily be determined if the parameters of the matching network are known.

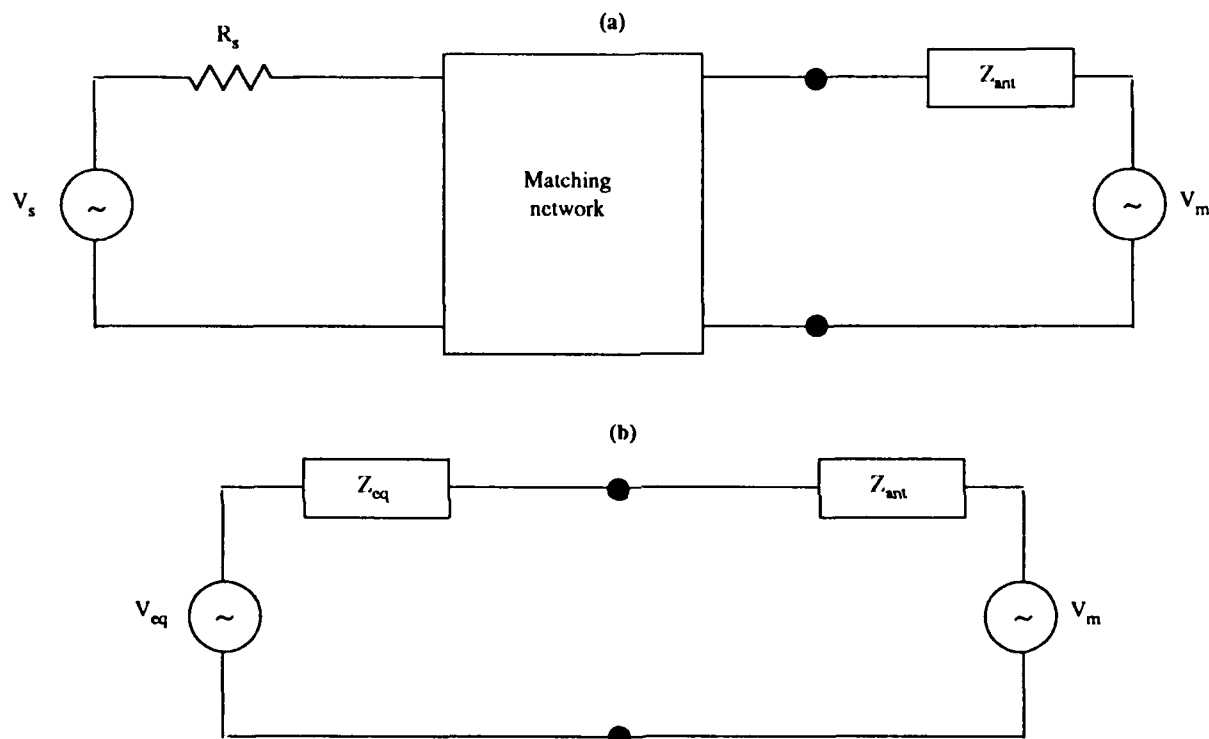


Fig. 5. (a) Equivalent circuit for an element in an array and (b) Thevenin equivalent circuit.

The value Z_{ant} is ideally found by theoretical or numerical analysis. In the present case the complexity of the geometrical configuration made accurate calculation of Z_{ant} difficult and measured data from a test fixture was used instead. Because it is only possible to make measurements at the input ports of the matching networks, it was necessary to measure both the input impedance of an isolated element and the mutual coupling between elements in order to determine the circuit parameters. The model parameters were adjusted to match the calculated and measured input impedance of an element in a small (7 element) array.

With the simplified element model in place, the entire array was analyzed via the MM technique under the assumption of an infinite ground plane. The current on each monopole was described by a single unknown. The current was assumed to be uniform on the post and flow radially from the center of the disk, decreasing linearly to zero at the edge of the disk, with no azimuthal variation. Rigorous numerical analysis showed that the actual distribution of the current is not significantly different from the assumed distribution for top-loaded monopole antennas which are actively fed or illuminated by a vertically polarized wave at grazing incidence, i.e., the two cases of interest in the antenna array problem. Since image theory was used to account for currents on the ground plane, the order of the MM system matrix was equal to the number of elements in the array.

In order to calculate the absolute gain and efficiency of the array, the total radiated power and the power available from the sources must first be determined. The total radiated power from an array of N active antenna elements is given by

$$P^{rad} = \sum_{i=1}^N \left[\frac{1}{2} \text{Re}(V_i I_i^*) - \frac{1}{2} \text{Re}(I_i Z_i^s I_i^*) \right],$$

where V_i is the total voltage across the terminals of element i and I_i is the current. The second term accounts for the power dissipation internal to the sources. The total power available from the sources is given by

$$P^{avail} = \sum_{i=1}^N \frac{1}{2} \text{Re} \left[V_i^s \left(\frac{V_i^s}{Z_i^s} \right)^* \right],$$

where V_i^s is the source voltage at element i . The directive gain of the array is then given by

$$D(\theta, \phi) = \frac{\frac{1}{2} \text{Re}(\bar{\mathbf{E}} \times \bar{\mathbf{H}}^*) \cdot \hat{\mathbf{r}}}{P^{rad} / 4\pi r^2}$$

and the absolute gain is given by

$$G(\theta, \phi) = \frac{\frac{1}{2} \text{Re}(\bar{\mathbf{E}} \times \bar{\mathbf{H}}^*) \cdot \hat{\mathbf{r}}}{P^{avail} / 4\pi r^2}$$

The efficiency of the array is P_{rad} / P_{avail} . This definition of gain assumes that the generator is matched to a transmission line connected to the antenna through an isolator and any power not accepted at the antenna terminals is dissipated.

Since the actual platform is finite, the effects of diffraction from the edges of the ground plane must be taken into account. These effects are particularly important for the case of endfire steering. The Uniform Theory of Diffraction (UTD)[7] was used to calculate the diffracted fields, which were then added to the direct radiation fields to obtain the array pattern. To simplify the calculations, the edge of the circular ground plane, which was many wavelengths from the closest antenna element, was approximated as a straight edge. Since the diffracted field depends on the distance from the source to the edge, a separate term for each antenna element was calculated for each observation angle.

Figure 6 shows the calculated and measured absolute gain of the array with 64 elements in a band driven and phased to form an endfire beam. The remaining 1242 elements were terminated in reactive loads to serve as parasitic beam forming elements. The agreement of experiment and theory is good except for elevation angles below the plane of the array. This discrepancy is to be expected since no attempt was made to model the support and corporate feed structure on the backside of the array.

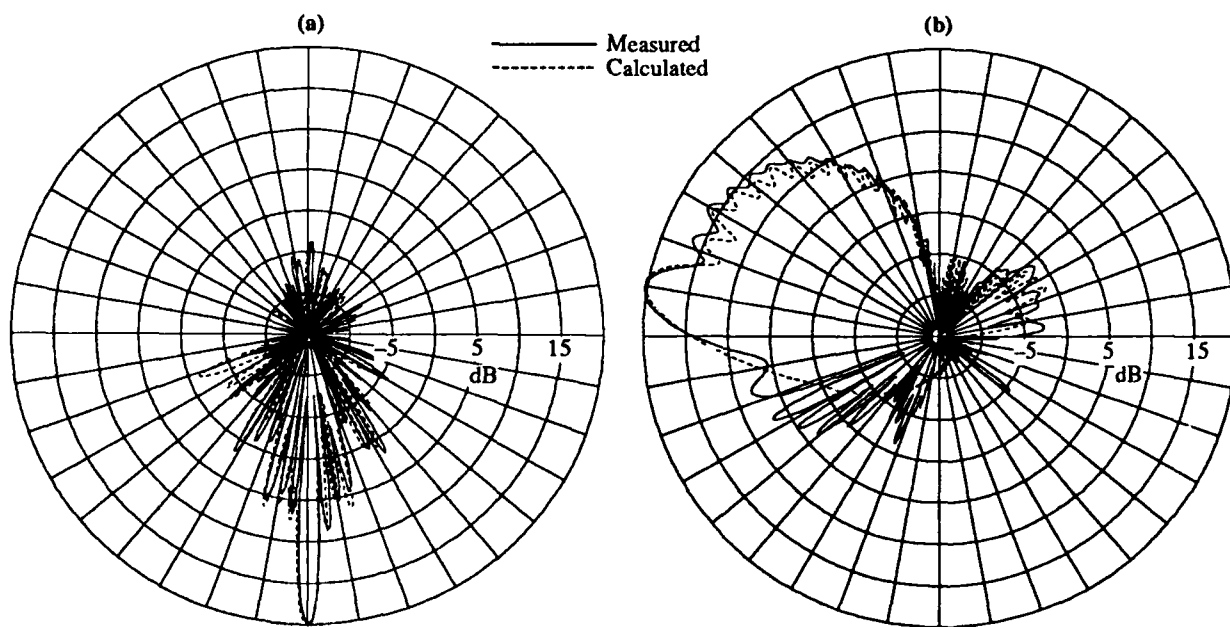


Fig. 6 Gain pattern of a 1306-element phased array on a circular ground plane 36.6 wavelengths in diameter; (a) azimuth cut (10° elevation) and (b) elevation cut.

Radome Effects

Radomes can modify the performance of arrays in several ways. Among them are lensing effects introducing beam distortion and termination effects yielding unwanted diffraction sidelobes. If the radome is in close proximity to the antenna elements, the antenna radiation resistance may change. Termination effects can be severe for

microstrip antennas where a radome in close proximity may function as a superstrate, thus increasing surface wave losses and reducing the antenna efficiency. Radome effects vary in degree, depending on the specifics of its geometry, its electrical properties and those of the array. A comprehensive discussion of these issues is outside the scope of this paper. However, a few examples follow illustrating the use of CEM techniques to incorporate radome effects.

For clarity of illustration, we consider a simple, seven element linear array. The elements are quarter-wave monopoles situated on a 10λ ground plane. The following four cases are considered: (a) no radome, (b) a circular shell radome of 0.1λ thickness, (c) a non-uniformly thick radome, and (d) the elements embedded in the radome. Two sets of radomes are considered having a permittivity of $\epsilon = 1.5$ and 3.0 . In all cases, the array elements are spaced 0.352λ apart and are Hansen-Woodward phased to endfire in the absence of the radome. The respective radiation patterns are elevation cuts for θ -polarization.

The results for case (a) are shown in Figure 7, comparing with those for an infinite ground plane. As can be seen, the finite ground planes introduce considerable re-radiation filling in the customary null at $\theta = 90^\circ$ for a linear monopole array. For case (b) in Figure 8, the thin shell radome distorts the main beam only slightly, however, the back lobes are enhanced significantly, particularly for the shell with $\epsilon = 3.0$. (For cases (b)-(d), the patterns are shown only for above horizon.) As seen in Figure 9, the non-uniform radome in case (c) has clearly deleterious effects on the main beam while the retrodirective (back) lobes fall generally within the envelopes of cases (a) and (b). Obviously, the non-uniform radome profile chosen here is not a desirable one for endfire. It is used for illustration only to emphasize that the radome termination can be an important issue. Finally for case (d) in Figure 10, the elements are part of the radome. Such may be the case for microstrip antennas and also for low profile, low volume airborne installations where the antenna elements are embedded in the skin of the aircraft.

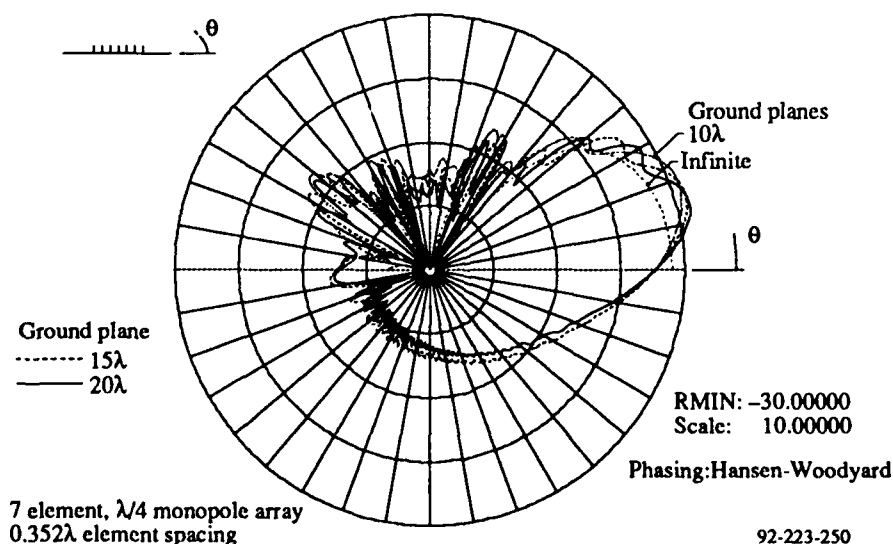


Fig. 7 Endfire array gain of a linear array: effect of a finite p.e.c. ground plane. No radome. θ -polarization.

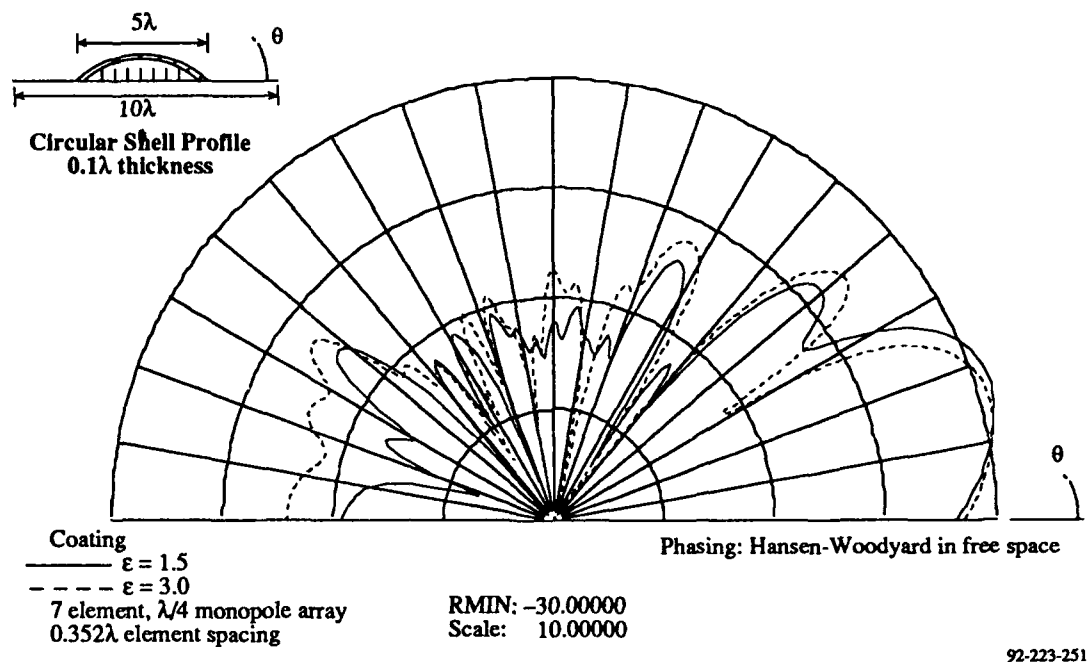


Fig. 8 Endfire gain of a linear array, enclosed with a thin shell radome. θ -polarization.

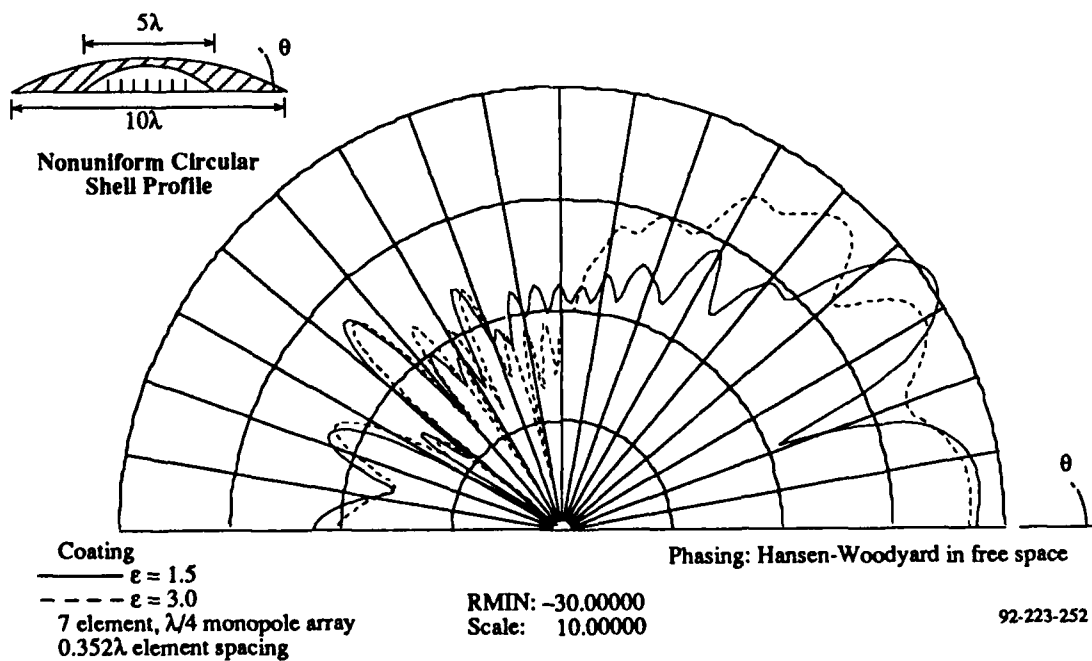


Fig. 9 Endfire gain of a linear array, enclosed with a nonuniform radome. θ -polarization.

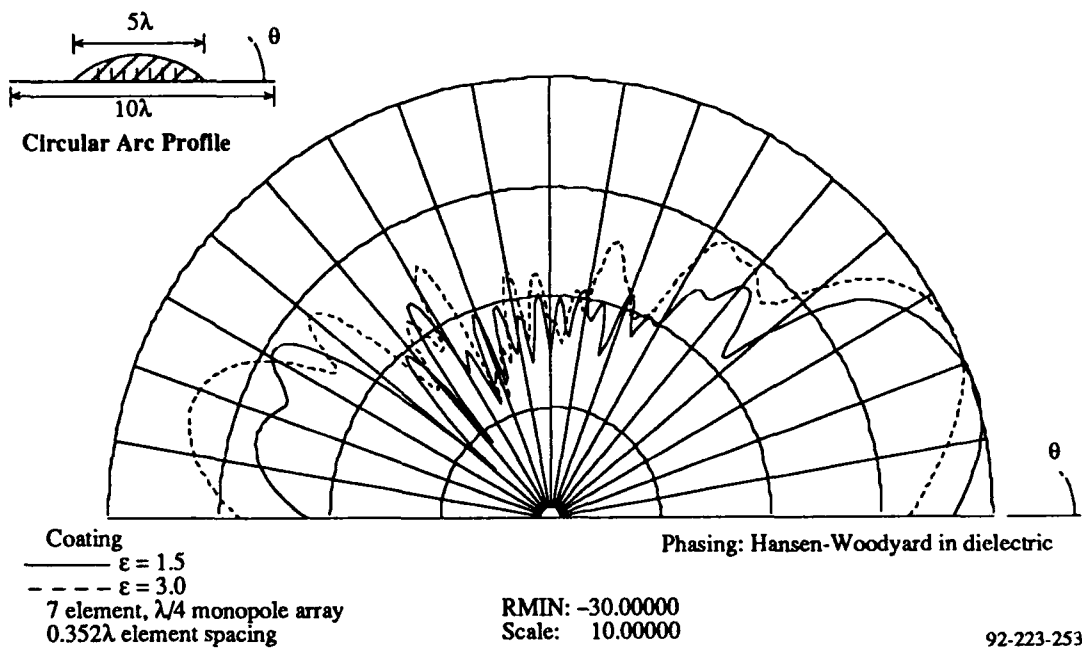


Fig. 10 Endfire gain of a linear array, with radome embedded elements. θ -polarization.

The calculations for the foregoing cases were carried out using the MM approach, rigorously incorporating all the boundary and junction conditions. As noted in the earlier circular array example, the element and platform (including radome) interactions must be included to achieve high fidelity simulations. This requirement gets stricter as the array becomes larger, its elements more complex (multi-functional), and the scanning angles wider.

Pattern Synthesis

Much work has been done in this area since the pioneering work of Schelkunoff^[31] and Dolph^[32]. Recent advances in this field were reviewed by Hansen^[33]. Today well-developed methods exist for synthesizing pencil beams for linear arrays and likewise for shaped beams. The Orchard-Elliott synthesis is widely used for the latter.^[34]

Here we highlight some results for shaped beam synthesis based on the Tihonov regularization^[35] adapted for discretized system by Mautz and Harrington^[36]. Earlier, Deschamps and Cabayan^[37] treated antenna pattern synthesis in general functional spaces. From the mathematical standpoint it is noteworthy that this problem is an example of a classic ill-posed one in the Hadamard sense. It can be shown that the CEM

methods discussed in Section II can be adapted to synthesis problems. For brevity, we confine the discussion to application of the MM approach.

Let the desired pattern at (θ_m, ϕ_m) be specified at $m = 1, \dots, M$ points on the radiation sphere as

$$g(\theta_m, \phi_m) = a_m e^{j\beta_m}$$

then a least mean square (LMS) solution implies minimization of the function[36]

$$\min_{f_n} \epsilon = \sum_{m=1}^M W_m \left| \sum_{n=1}^N T_{mn} f_n - a_m e^{j\beta_m} \right|^2$$

where W_m are the synthesis weights chosen and T_{mn} is a matrix operator discussed below. The currents on the radiator are given by the MM (Galerkin) expansion as:

$$I(S) = \sum_{n=1}^N f_n v_n(S)$$

where f_n are the unknown coefficients associated with the basis functions $v_n(S)$ spanning the radiator surface S . Note in general, this surface encompasses both the antenna (array) elements as well as its platform. In the notation of Figure 2, $S = S_1 \cup S_2$. In synthesis there are two important cases. If the magnitude (a_m) and phase (β_m) are constrained, the LMS solution yields the following condition on the Galerkin coefficients:

$$f_n = \sum_{m=1}^M \left\{ ([\tilde{T}_{nm}^*][W_{mm}]^{-1}) [\tilde{T}_{nm}^*][W_{mm}] \right\} g_m$$

with $g_m \equiv g(\theta_m, \phi_m)$. If alternately the magnitude is constrained, the condition on the phase becomes:

$$e^{j\beta_m} = \frac{\sum_{n=1}^N [T_{mn}] f_n}{\left| \sum_{n=1}^N T_{mn} f_n \right|}$$

In the above equations, T_{mn} denotes an operator that maps the fields from the n -th Galerkin expansion function to the m -th observation point (i.e., at θ_m, ϕ_m). This operator contains all the interactions between the various parts of the antenna (array) and its platform. In this case the coupling between elements is explicitly included in the synthesis procedure.

The computation of the T operator can be carried out in a number of ways. If an SIE formulation is used to describe the physics of the problem and the SIE is solved by the MM technique, then T is related to the MM system (impedance) matrix. Similar relationships hold if the SIE is solved by hybrid methods involving asymptotic current or field Ansatzes or a hybrid SIE/PDE solution (as in the MM/FEM approach). The T operator has been obtained for classes of problems depicted in Figure 1. Space constraints preclude a discussion of these computations here.

To illustrate the above synthesis procedure, we consider the earlier generic monopole array, i.e., case (d) in Figure 10. Let the desired pattern be uniform for $0 \leq \theta \leq 30^\circ$ and zero otherwise. Using the LMS phase synthesis, incorporating inter-element and radome coupling, yields the pattern in Figure 11. As can be seen in Figure 12, the synthesized pattern is superior to the Hansen-Woodward one, particularly for the back lobes. The corresponding comparisons where the radome permittivity is $\epsilon = 3.0$ is depicted in Figure 13. As noted before, the foregoing results include all the coupling effects. If coupling is omitted, the corresponding results are markedly degraded as seen in Figure 14.

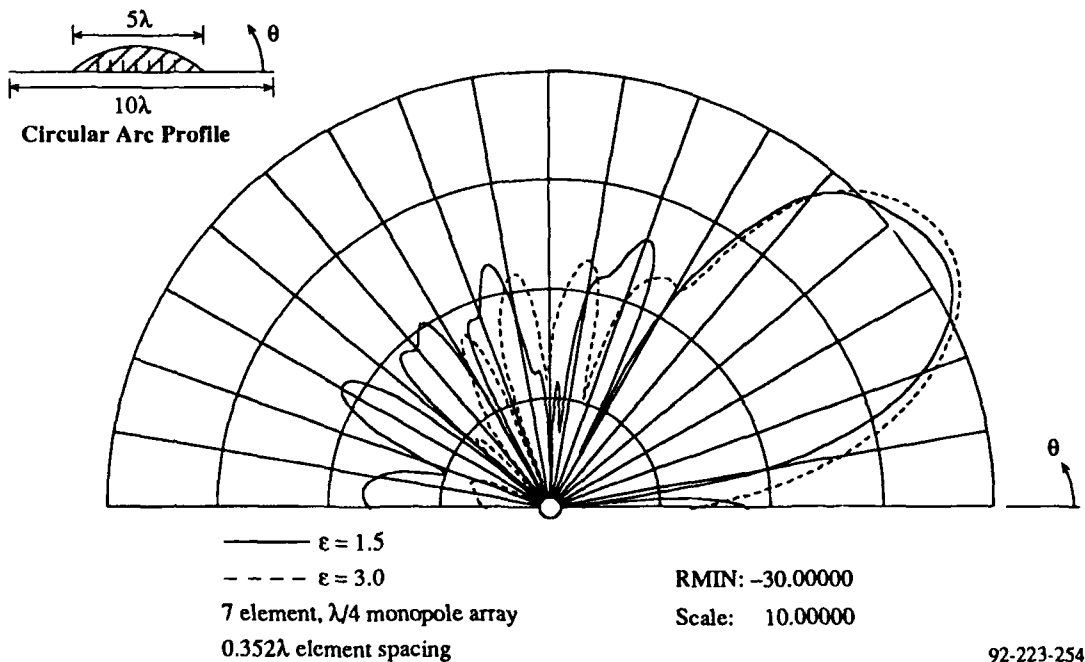


Fig. 11 Synthesized pattern for embedded array elements. Main beam steered to 35° . θ -polarization.

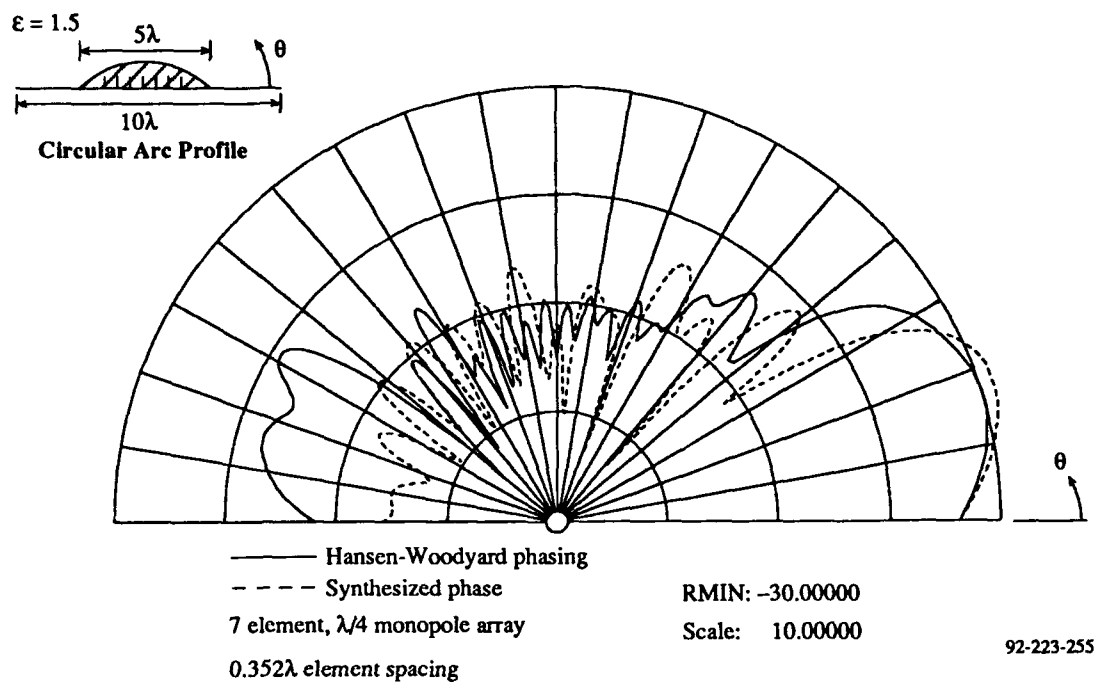


Fig. 12 Comparison of synthesized and Hansen-Woodward phased array patterns. Array steered to endfire. Dielectric $\epsilon = 1.5$. θ -polarization.

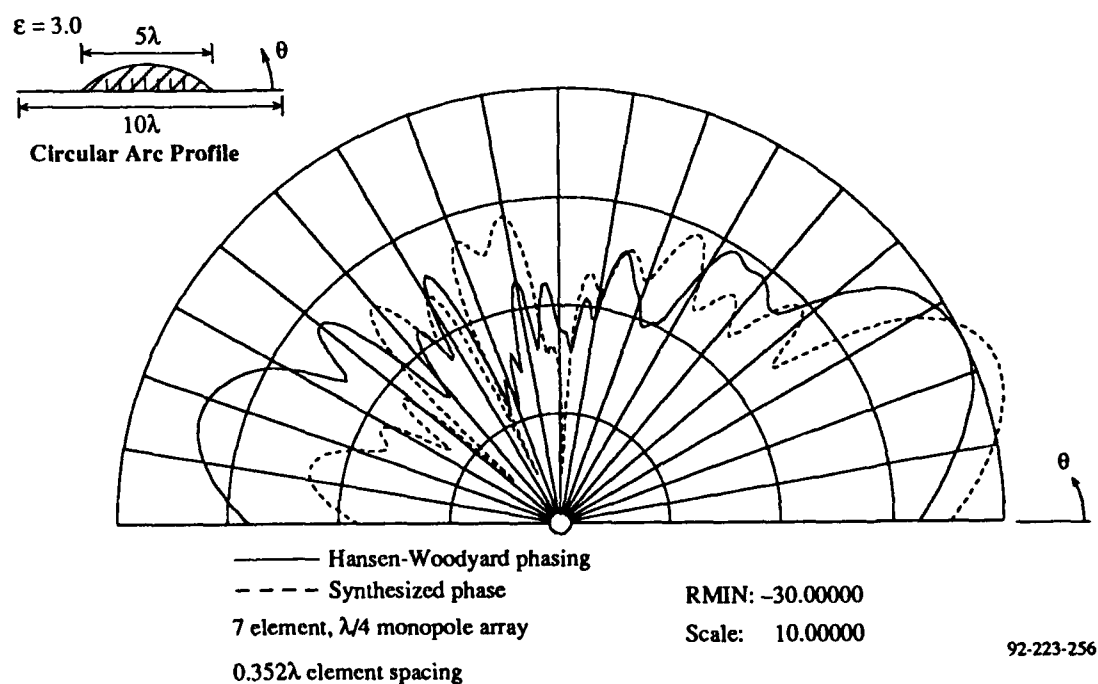


Fig. 13 Comparison of synthesized and Hansen-Woodward phased array patterns. Array steered to endfire. Dielectric $\epsilon = 3.0$. θ -polarization.

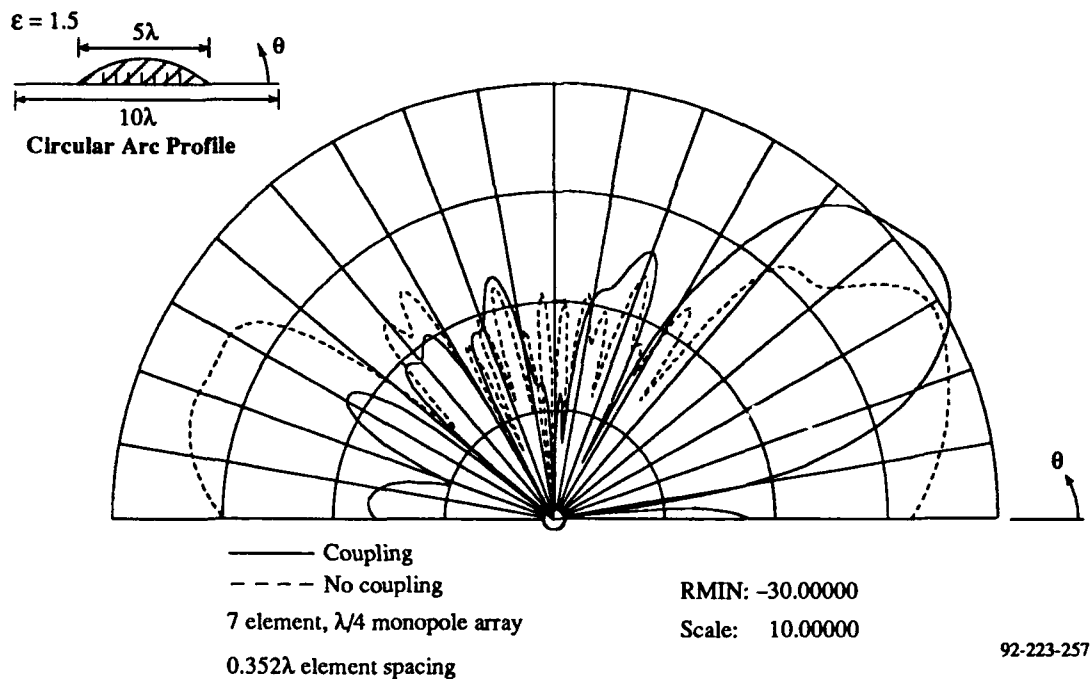


Fig. 14 Effect of coupling between radiating elements on steered beam. Main beam steered to 35° . θ -polarization.

Analysis of Complex Configurations

Several examples are considered next where asymptotic and hybrid methods provide effective solution procedures.

For electrically large problems, asymptotic methods are often useful. Standard asymptotic methods (e.g., GTD, PTD, and UTD) are limited, however, to conducting surfaces and a small number of other types of canonic surfaces. Hence, these methods cannot be applied in many cases of interest. Recently, a method known as the Numerical Uniform Theory of Diffraction (NUTD) has been developed for treating discontinuities in arbitrary surfaces^[38]. In this method diffraction coefficients are extracted from data calculated via the MM technique and applied within the UTD framework. Figure 15 shows an example calculation of the radiation pattern of a microstrip antenna on a doubly curved substrate. The effects of surface wave diffraction from the edges of the substrate were calculated using NUTD. Figure 16 shows a second example of a NUTD calculation. In this case the radiation pattern of a line source behind a conducting screen with a resonant metallic radome window is shown. The window is an array of tripole slots ("chicken tracks") arranged in a triangular lattice. In both cases there is excellent agreement with the MM results.

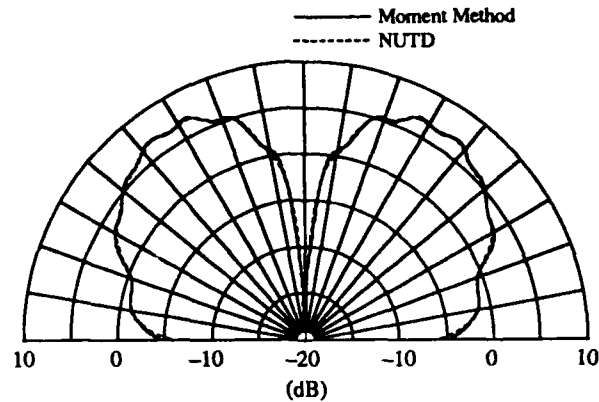
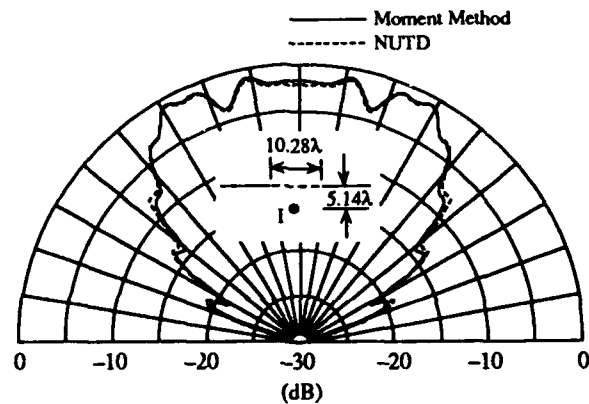


Fig. 15. Gain pattern of a microstrip disk antenna conformal to a spherical surface. Vertical polarization.



992-223-258

Fig. 16 Gain pattern of a line source behind a resonant metallic radome window in a conducting screen.

The final two cases are examples of a hybrid approach where the free space Green's function is perturbed by the presence of the antenna platform. In one case, the platform is a finite wing section with active and parasitic monopoles (Figure 17). This case fits conceptually the problem depicted in Figure 2 where the surface S_1 corresponds to the monopoles and S_2 is the wing section. The Green's function perturbation can be computed numerically by various means including the MM technique. If the latter approach is used (as was done here), it is often denoted as the hybrid MM/matrix Green's function approach.^[12] Depending on the electrical size of S_1 and S_2 , this approach may be identical with an MM formulation. If the radiating structure and/or its platform is electrically large, the MM/matrix Green's function will be an approximation. In Figure 17, the results coalesce with the standard MM formulation.

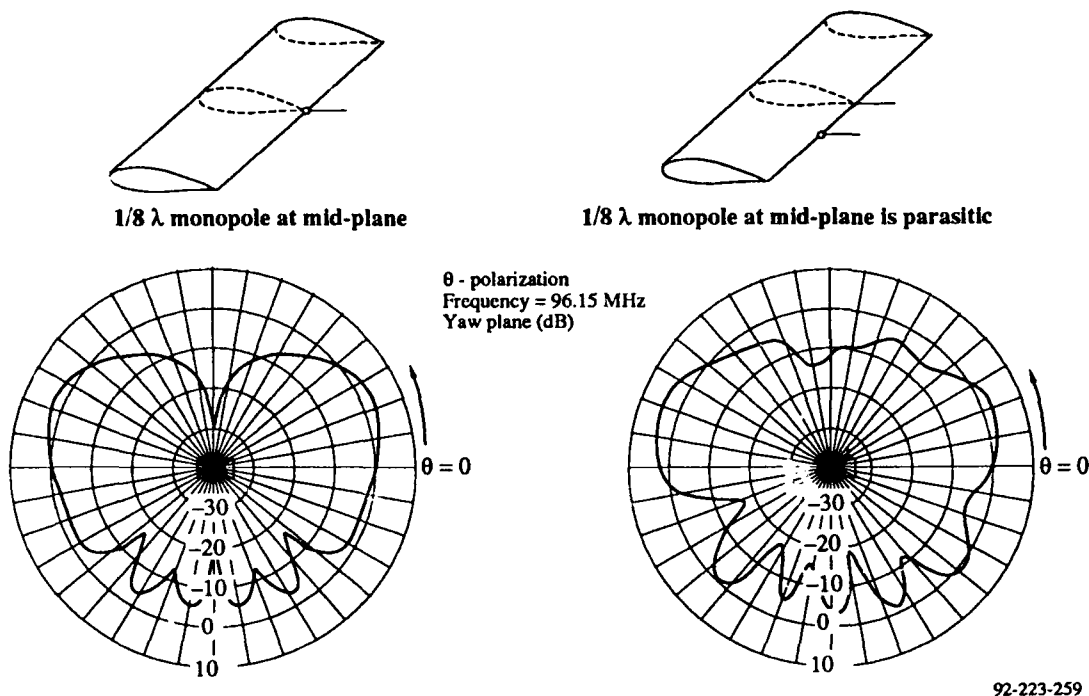


Fig. 17 Calculated radiation patterns for wing-mounted monopole antennas. Results are for active and parasitic elements.

The second case depicted in Figure 18 can be thought of as an antenna problem near an interface. If the antenna were a simple one (i.e., monopole) above a lossy earth, the perturbation in the Green's function can be obtained from the classical Sommerfeld formulation.^[39] Clearly in this case at HF frequencies, the loop antenna as well as the helicopter airframe are closely coupled and together act as a radiating structure. It can be shown that in this case, the perturbation to the free space Green's function due to the proximity of the earth (i.e., S_2 in the nomenclature of Figure 2) can be approximated via the reflection coefficient method (RCM).^[40-42] In addition, a matrix Green's function is used in the MM representation of the antenna and airframe. As can be seen in Figure 18, the horizontal plane results are in good agreement with the measured data. Experimental measurement of the pattern in the vertical plane was not feasible.

The foregoing cases are but a few examples of the broad diversity of antenna problems that can be handled by asymptotic and hybrid methods. These methods can be powerful analysis tools for problems where the complexity and/or electrical size of the configuration make application of low frequency techniques such as MM, FEM, and FDTD by themselves impractical.

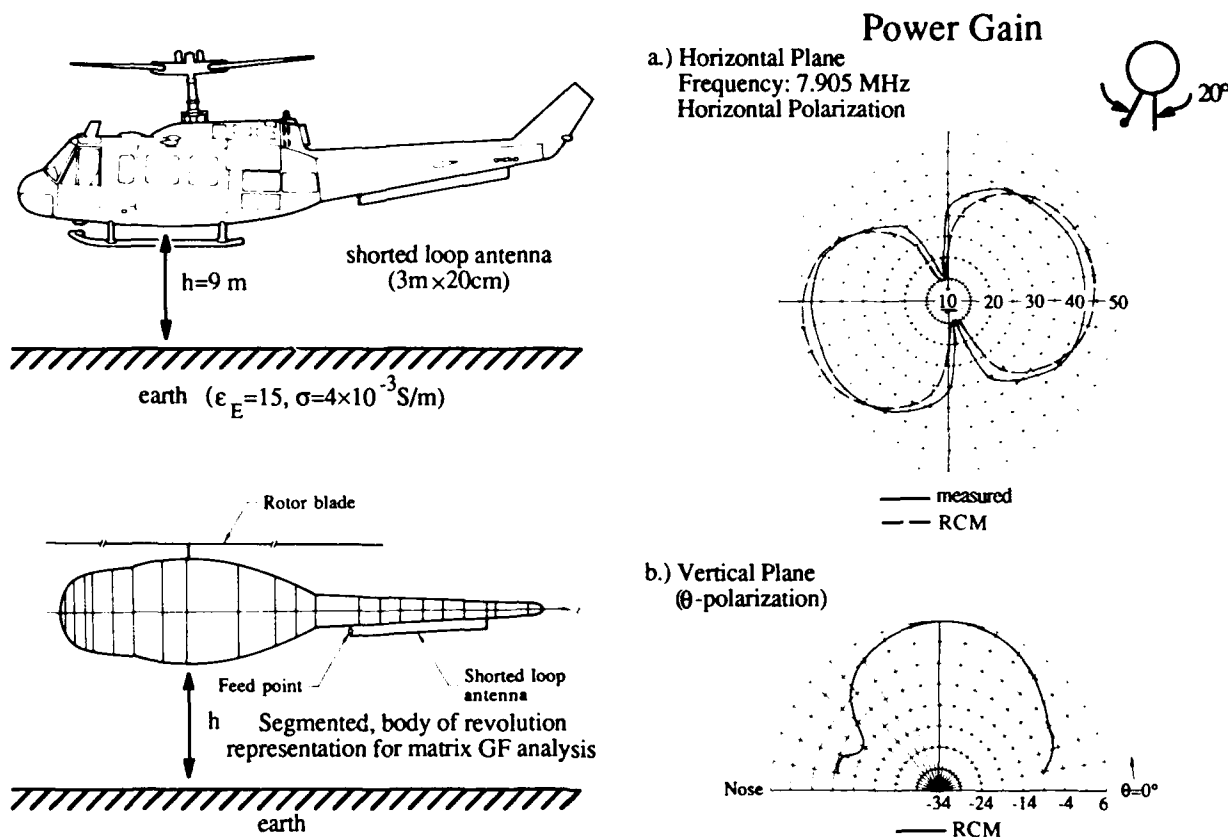


Fig. 18 Calculated and measured radiation patterns for a shorted loop antenna on a hovering helicopter. (a) horizontal plane pattern, and (b) vertical plane pattern.

IV. Summary

In this paper, we summarized key approaches and recent advances in CEM as applied to complex antenna and array problems. In particular, the discussion focused on low frequency "numerical" methods such as MM, FEM, and FDTD. Hybrid approaches linking these numerical techniques with asymptotic methods and incorporation of surface integral equation and partial differential equations (volume) formulations were outlined. Application of these CEM analysis tools to complex antenna/array problems was illustrated by representative examples.

References

1. Special Issue on Antennas, PROC. IEEE, 80, January 1992.
2. E. K. Miller, L. N. Medgyesi-Mitschang and E. H. Newman (eds.), Computational Electromagnetics, IEEE Press, 1991.
3. P. Silvester and R. Ferrari, Finite Elements for Electrical Engineers, Cambridge Press, Cambridge, 1990.
4. A. Taflove and K. R. Umashankar, "The FDTD method for numerical modeling of electromagnetic wave interactions with arbitrary structures," in Finite Element and Finite Difference Methods in Electromagnetics Scattering, Progress in Electromagnetic Research, vol. 2, New York: Elsevier, 1990; Chap. 8.
5. J. B. Keller, "Geometrical theory of diffraction," JOSA, 52, pp. 116-130, February 1952.
6. R. C. Hansen (ed.), Geometric Theory of Diffraction, IEEE Press (1981).
7. P. H. Pathak, "High frequency techniques for antenna analysis," PROC. IEEE, 80, pp. 44-65, (1992).
8. R. G. Kouyoumjian and P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface," IEEE PROC., 62, pp. 1448-1461, 1974.
9. P. Ya Ufimtsev, "Method of edge waves in physical theory of diffraction," Air Force System Command, Foreign Techn. Div. DOC. ID FTD-HC-23-259-71, (1971).
10. E. P. Ekelman and G. A. Thiele, "A hybrid technique for combining the method of moments treatment of wires with the GTD for curved surfaces," IEEE Trans. Antennas Propagat., AP-28, pp. 831-839, (1980).
11. L. N. Medgyesi-Mitschang and D. S. Wang, "Review of hybrid methods on antenna theory," ANN. Telecommun 44, no. 9-10, pp. 445-455, (1989).
12. E. H. Newman, "An overview of the hybrid MM/Green's function method in electromagnetics, PROC. IEEE, 76, pp. 270-282, (1988).
13. L. N. Medgyesi-Mitschang and D. S. Wang, "Hybrid solutions for scattering from perfectly conducting bodies of revolution," IEEE Trans. Antennas Propagat., AP-31, pp. 570-583, (1983).
14. L. N. Medgyesi-Mitschang and D. S. Wang, "Hybrid solutions for scattering from large bodies of revolution with material discontinuities and coatings," IEEE Trans.. Antennas Propagat., AP-32, pp. 717-723 (1984).
15. E. K. Miller, "Low frequency computational electromagnetics for antenna analysis," PROC. IEEE, 80, pp. 24-43, (1992).
16. S. M. Rao, D. R. Wilton and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," IEEE Trans. Antenna Propagat. AP-30, pp. 409-418, (1982).
17. D. S. Wang, "Current-based hybrid analysis for surface-wave effects on large scatterers," IEEE Trans. Antennas Propagat., AP-39, pp. 839-850 (1991).

18. J. F. Shaeffer and L. N. Medgyesi-Mitschang, "Radiation from wire antennas attached to bodies of revolution: The Junction Problem," *IEEE Trans. Antennas Propag.*, AP-29, pp. 479-487, (1981).
19. L. N. Medgyesi-Mitschang and J. M. Putnam, "Formulation for wire radiators on bodies of translation with and without end caps," *IEEE Trans. Antennas Propag.*, AP-31, pp. 853-862, (1983).
20. J. M. Bornholdt and L. N. Medgyesi-Mitschang, "Mixed-domain Galerkin expansions in scattering problems," *IEEE Trans. Antennas Propag.*, AP-35, pp. 216-227, (1988).
21. J. D'Angelo and I. Mayergoyz, "On the use of local absorbing boundary conditions for RF scattering problems," *IEEE Trans. Mag.*, MAG-25, pp. 3040-3043, (1989).
22. R. Mittra, O. Ramahi, A. Khebir, R. Gordon, and A. Kouki, "A review of absorbing boundary conditions for two- and three-dimensional electromagnetic scattering problems," *IEEE Trans. Mag.*, MAG-25, pp. 3034-3040, (1984).
23. W. Boyse and A. Seidl, "A hybrid finite element method for conductors with thin dense coatings," *App. Comp. Electromagnetics Soc.*, 4, pp. 59-72, (1989).
24. T. Cwik, "Coupling finite element and integral equation solutions using decoupled boundary meshes," *IEEE Trans. Antennas Propag.*, to appear.
25. K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, AP-14, pp. 302-307, (1966).
26. T. G. Jurgens, A. Taflove, K. Umashankar, and T. G. Moore, "FDTD modeling of curved surfaces," *IEEE Trans. Antennas Propag.*, AP-40, pp. 357-366, (1992).
27. J. G. Maloney, G. S. Smith, and W. R. Scott, "Accurate computation of the radiation from simple antennas using the finite-difference time-domain method," *IEEE Trans. Antennas Propag.*, AP-38, pp. 1059-1068, (1990).
28. A. Reineix and B. Jecko, "Analysis of microstrip patch antennas using finite difference time domain method," *IEEE Trans. Antennas Propag.*, AP-37, pp. 1361-1369, (1989).
29. A. Taflove, K. Umashankar, B. Beker, F. Harfoush, and K. Yee, "Detailed FD-TD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens," *IEEE Trans. Antennas Propag.*, AP-36, pp. 247-257, (1988).
30. G. S. Sanford and P. Westfeldt, Jr., "Circular phased array development for electronic scanning of an endfire beam," *Proc. 1991 Antenna Applications Symposium*, Sept. 1991.
31. S. A. Schelkunoff, "A mathematical theory of linear arrays," *Bell System Tech. J.*, 22, pp. 80-107, (1943).
32. C. L. Dolph, "A current distribution for broadside arrays which optimize the relationship between beamwidth and side lobe level," *Proc. I.R.E.*, 34, pp. 335-348, (1946).
33. R. C. Hansen, "Array pattern control and synthesis," *PROC. IEEE* 80, pp. 141-151 (1992).
34. H. J. Orchard, R. S. Elliott, and G. J. Stern, "Optimizing the synthesis of shaped beam antenna patterns," *IEE Proc. (London)*, Pt. 132H, pp. 63-68, (1985).

35. A. N. Tihonov, "Solution of incorrectly formulated problems and the regularization method," *Sov. Math.* 4, pp. 1035-1038, (1963).
36. J. R. Mautz and R. F. Harrington, "Computational Methods for Antenna Pattern Synthesis," *IEEE Trans. Antennas Propagat.*, AP-23, pp. 507-512, 1975.
37. G. A. Deschamps and H. S. Cabayan, "Antenna synthesis and solution of inverse problems by regularization methods," *IEEE Trans. Antennas Propagat.*, AP-20, pp. 268-274, (1972).
38. M. P. Hurst, "Improved numerical diffraction coefficients with application to frequency-selective surfaces," *IEEE Trans. Antennas Propagat.*, AP-40, pp. 606-612, (1992).
39. R. Mittra, P. Parhami, Y. Rahmat-Samii, "Solving the current element problem over lossy-half-space without Sommerfeld integration," *IEEE Trans. Antennas Propagat.*, AP-28, pp. 778-782, (1979).
40. E. K. Miller, A. J. Poggio, G. J. Burke, E. S. Seldon, "Analysis of wire antennas in the presence of a conducting half-space, part I. The vertical antenna in free space," *Can. J. Phys.* 50, pp. 879-888, (1972).
41. E. K. Miller, A. J. Poggio, G. J. Burke, E. S. Seldon, "Analysis of wire antennas in the presence of a conducting half-space, part II. The horizontal antenna in free space," *Can. J. Phys.* 50, pp. 2614-2627, (1972).
42. G. J. Burke, E. K. Miller, L. N. Brittingham, D. L. Lager, R. J. Lytle, J. T. Okada, "Computer-modeling of antennas near the ground", *Electromagnetics* 1, 29-49, (1981).

Benchmark Problems for Electromagnetics Software Validation: An Overview of the Joint ACES/IEEE-AP Effort

*Andrew F. Peterson
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250*

Abstract: Improved validation procedures for electromagnetics software are needed to enhance the degree of confidence in numerical solutions. The Applied Computational Electromagnetics Society (ACES) Software Performance Standards Committee and the IEEE Antennas and Propagation Society (IEEE-AP) Committee on Electromagnetic Modeling Software have been created as a joint committee with the initial objective of establishing a database of benchmark problems, and the long-term goal of enhancing the practice of electromagnetics software validation. This presentation will summarize the organization of these committees, recent and ongoing activities, and ancillary issues related to software validation.

1. INTRODUCTION

During the past three decades, computational electromagnetics has matured to the degree that it is presently considered as important for electromagnetic engineering as the disciplines of theoretical analysis and experimental measurement. However, there is currently a proliferation of computational techniques and specific computer codes that are offered to the research and user communities without systematic procedures for verifying their results. To a large extent, the extreme complexity of the mathematical foundations of these procedures (for instance, integral equation formulations) makes it impossible to establish precise error

bounds or other self-contained accuracy tests. Consequently, attempts to establish the accuracy of numerical results are usually based on comparisons with solutions obtained by other, independent means. Few practical geometries yield exact analytical solutions, and data from other numerical solutions may be subject to the same uncertainties and approximations as the specific code under consideration. Ultimately, the most desirable comparison may be with measured results obtained by careful experimental procedures. Yet measured data may not be available in a form suitable for direct comparison with the output of a particular code, and the cost of carrying out a measurement program for validating a specific code may be prohibitive to the average code developer. Even when an effort is made to conduct a detailed program of validation, the absence of systematic procedures for validation often leaves a question mark in the minds of code users. As reported by Rautio, "most published validation efforts are viewed with skepticism" and must be duplicated "at considerable expense" by each subsequent user [1].

The increasing difficulty of validating the results of available computer codes has motivated several professional societies with interest in electromagnetics (EM) applications to investigate and promote better procedures for EM software validation. The Applied Computational Electromagnetics Society (ACES) Software Performance Standards Committee and the IEEE Antennas and Propagation Society (IEEE-AP) Committee on Electromagnetic Modeling Software were established during 1988 with E. K. Miller as Chairman. (Although these societies are distinct, the two committees actually function as a single entity and will be denoted "the committee" in the following.) A. F. Peterson succeeded Miller as Chairman in 1991. The initial objective of the committee has been to establish a database of "solved" problems that could be used as standard benchmark solutions for validating EM codes. Although the ACES charter is broad in scope, the committee has concentrated primarily on wire antennas, scattering, and other high-frequency applications. Presently, there is a growing participation of the low-frequency community with interests such as eddy current problems.

Other organizations are also collecting benchmark problems in electromagnetics. As an example, each issue of the *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering* contains an "MIC Simulation Column" coordinated by I. J. Bahl, to which readers are encouraged to submit canonical problems and solutions. A second example is the ongoing series of Testing Electromagnetics Analysis Methods (TEAM) workshops, coordinated by an international group of researchers with interests in nuclear reactor and magnetic applications modeling. ACES and TEAM have collaborated on several recent workshops and the current ACES/IEEE-AP workshops are largely patterned after the TEAM model. A third example is the Electromagnetic Code Consortium (EMCC), which is primarily concerned with electromagnetic scattering applications [2]. As part of their validation activities, the EMCC requests that their code developers analyze a number of specific canonical geometries.

To date, the ACES/IEEE-AP activity has resulted in three workshops at the summer AP symposia, shorter workshops at annual ACES symposia, and several joint workshops with TEAM. Two special issues of the ACES Journal have been devoted to validation issues [3-4], the second consisting entirely of proposed canonical problems.

2. COMMITTEE ORGANIZATION

Because of the wide range of possible application areas, the committee has been organized around "working groups" that focus on specific areas of interest and narrow the scope of the task facing any particular group. Initially, focused working groups were established in the application areas *Wires, Surfaces & Penetrable Objects*, and *Time Domain*. (These areas are concentrated in antennas and scattering, but the committee has a broad charter and participation from other application areas is welcomed and expected.) Additional focused working groups will be established as needed, and may be targeted toward much more specific application areas than those listed above — as specific as *Dihedrals and Corner*

Reflectors, for instance. In addition, procedural working groups were initially established in the areas of *Database Issues* and *Policy and Procedures*. The procedural working groups are not active at the present time and may be reactivated when the focused working groups reach a mature enough state to benefit from central coordination.

Based to a large extent on the TEAM model, the following format has been proposed: Each working group shall identify a short list of 3-5 precisely defined problems. These are publicized to the EM modeling community, which responds by attempting to solve them and routing their solutions back to the working group. (The solutions may be based on measurement, numerical modeling, analytical approximation, etc.) After a suitable time interval (perhaps one or two years) a particular working group may decide that a problem is "solved" to acceptable accuracy, archive the solution in the database, and remove it from the list of active problems. (The issue of what constitutes a "solved" problem remains open; see discussion below.) In this manner, the committee will gradually build a database of solved problems. Regular workshops will provide a forum for working groups to meet and discuss their business, including a review of new proposed solutions, decisions as to whether a problem should be considered solved or not, and the identification of new canonical problems to replace those removed from the active set. Regular publications in society newsletters will document the current "unsolved" problems and make the results available to the research and user communities.

Under the proposed organizational framework, the key unit is the working group. Each working group is a relatively autonomous entity, free to decide what canonical problems to propose and at what point a solution should be added to the database. Since working groups will conduct business concurrently at workshops, it will probably be necessary for each participant to be affiliated primarily with one group. Because workshops will be held in conjunction with various conferences, there is likely to be a rather large "turn over" of participants from one workshop to the next. Although expected, the turn-over is cause for concern, as some means

(detailed reports?) will be necessary to ensure the continuity of each group's activities over time.

3. SELF-CONTAINED VALIDATION

The fundamental issue behind validation efforts is clearly: *How accurate are the results of a particular computation?* An ideal "self-validating" code would provide error bars incorporating all known sources of error as a function of parameters including the geometry, the excitation, the discretization, any approximations incorporated into the procedure, and numerical round-off or truncation effects. While we will never achieve the utopian "self-validating" code, it is interesting that in many recent publications even the mildest sort of error analysis appears to be the exception rather than the rule. The author is aware of only one recent computational result, due to Ludwig, that includes approximate error bars [4].

The mathematical complexity of many EM formulations rules out precise error bounds for almost all structures of practical interest. Perhaps as a consequence of the difficulty of establishing an ironclad error bound, the EM community appears to tolerate the absence of any error analysis. However, despite the absence of a rigorous mathematical bound, systematic error analysis can be attempted.

For instance, a large part of the error associated with many numerical results is simple interpolation error. The error in a piecewise-polynomial approximation is $O(h^{p+1})$, where h is the largest cell dimension and p is the polynomial order [5]. For these problems, an estimate of the actual interpolation error can therefore be gleaned from two numerical results obtained with different cell dimensions. The estimated interpolation error must be considered a *lower* bound on the overall error, and not the only error in the result. For numerical formulations that produce results known to converge to the correct solution as $h \rightarrow 0$, this type of analysis may

engender confidence in particular results. Unfortunately, many formulations incorporate approximations that negate the convergence of the solution. Even in robust formulations, at some point as $h \rightarrow 0$ numerical truncation and round-off effects may alter the $O(h^{p+1})$ trend.

Under the present committee structure, each working group is expected to decide "What constitutes a solved problem?" in the context of their application area. In actuality, these groups will only be able to establish that, with reasonable confidence, a particular result is correct to some specified accuracy. The hope of the present effort is that *some form of error analysis will be provided with each contributed result*, and that the types of error analysis established in the course of this activity will prove acceptable to the EM community. In other words, the philosophical question "What constitutes a solved problem?" is replaced by "How much confidence do we place in a particular error analysis?" At the present time, there is no standard error analysis to apply; it is hoped that the activities of the committee may contribute to improved error analyses applicable to EM software. Assuming that error levels are estimated for all contributed results, the task of selecting a solution to be archived in the database becomes systematic and somewhat manageable.

4. VALIDATION BY EXPERIMENTAL MEASUREMENT

Data obtained by experimental measurement are expected to comprise a large portion of the archived canonical solutions. In common with other proposed solutions, experimental results must be accompanied by some sort of error analysis and will be subject to scrutiny by the appropriate working group. Measured data can potentially serve as a "primary standard" with which to compare numerical results, yet obtaining useful measured data may not be a trivial task. In a recent article [1], Rautio discusses the present state of experimental validation of software and raises several points that must be addressed in a meaningful validation activity. One is the issue of sensitivity, i.e., how sensitive are the measured results to the

particular feature of the geometry under consideration? Rautio also discusses the importance of properly combining distinct sources of error into an overall error index, the advantageous use of scale models, and the benefit of changing the typical experimental objective from the subjective goal of confirming that the computed result is "correct" to a quantitative identification of the error level associated with the computation. The discussion in [1] accentuates the fact that an experimental validation program generally requires substantial planning and effort.

5. SUMMARY

Although the ACES Software Performance Standards Committee and the IEEE-AP Committee on Electromagnetic Modeling Software have been in existence for several years, the task of collecting a database of "solved" problems for use as benchmark solutions is only beginning. Contributed solutions to several proposed canonical problems have been received, and participants are presently considering issues including relative accuracy, possible error analyses, and the desired format of numerical results.

Many of the issues identified in the course of the committee's activities extend beyond the scope of constructing a canonical problems database. Better understanding of these issues may ultimately enhance the scientific rigor of numerical modeling in EM, and lead to increased confidence in numerical models. Consequently, it is expected that the committee's ongoing activities will be of interest to the EM community, and that EM practitioners will find their self-interests served by participation. The committee is a volunteer organization, and interested participants are always welcome!

REFERENCES

- [1] Rautio, J. C. (1991) Electromagnetic validation of electromagnetic software, International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering 1:379-385.
- [2] Faison, J. C. (1990) The Electromagnetic Code Consortium, IEEE Antennas and Propagation Society Magazine 32:19-23.
- [3] Kubina, S. J. and D. E. Stein, Eds. (1989) Electromagnetics Computer Code Validation, Special Issue of the ACES Journal.
- [4] Sabbagh, H. A., Ed. (1990) The ACES Collection of Canonical Problems Set 1, Special Issue of the ACES Journal.
- [5] Strang, G. and G. J. Fix (1973) An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, 136-171.

COMPUTATIONAL ELECTROMAGNETICS— CONSIDERATIONS FOR THE NEXT GENERATION OF SUPERCODES

Kenneth R. Siarkiewicz
Rome Laboratory/ERPT
525 Brooks Road
Griffiss AFB NY 13441-4505

ABSTRACT

During the last two decades several large, multi-purpose, generally available codes have been developed for use in system-level electromagnetic analyses. These have employed various *numerical techniques*, either singly or in some hybridized combination. The development, maintenance, use, and enhancement of these codes have provided many lessons to the electromagnetics community regarding how the next generation of "supercodes" should look.

Several other factors also impact the design and capabilities of future CEM codes. Among these is the fact that the focus of electromagnetic studies is changing. Antennas have become apertures, and they now consist of linear, planar, and conformal configurations of simpler antenna types. Furthermore, these basic building blocks are increasing in geometric and material complexity.

- 1 -

This requires a new and radical approach to performing system-level electromagnetic analyses.

Greater detail and resolution are necessary to accommodate the evolving aperture technology and the uses to which that technology will be put. Therefore, geometric modeling capabilities and modeling strategies must also undergo further study and evolution.

Finally, computer hardware and software are advancing in great leaps. Conventional sequential processing is being replaced by computing on massively parallel processing (MPP) platforms. Algorithm development and numerical processing procedures are evolving to fully utilize the computational speed and throughput potential of MPPs. However, the employment of these new technologies in CEM codes must achieve and even extend the lower computational noise floors required by future numerical electromagnetics requirements.

This paper examines these and many other factors which must be considered during the design and development of a parallelized electromagnetic analysis code. These are presented in light of the current CEM technology and the projected analysis requirements demanded by emerging system concepts. A broad, top-level specification for the next generation system-level CEM tools is

proposed. Some factors and implications which affect the development of the new tools are also discussed.

1.0 INTRODUCTION

This paper presents an assessment of the capabilities of the current state-of-the-art of system-level computer-aided electromagnetic fields analysis codes and draws implications for the future of such codes in light of the concepts and systems that are being and will be developed during the next decade and more. The assessment is based on experience gained during the development, use, maintenance and enhancements of the General Electromagnetic Model for the Analysis of Complex Systems (GEMACS) computer program. Focusing on the experiences associated with this one computer program is by no means restrictive, nor does it invalidate the conclusions drawn and recommendations made. It is representative of the class of computational electromagnetics (CEM) codes called "supercodes." Other such codes include the Numerical Electromagnetics Code (NEC) and the Basic Scattering Code (BSC). These codes were developed to analyze system-level electromagnetic phenomena (e.g., antenna pattern distortion generated by platform scattering) using sophisticated numerical EM techniques

[Method of Moments (MOM) and Uniform Theory of Diffraction (UTD)].

Furthermore, the analysis is to be performed using geometric and electric models of the system components (antennas, engines, etc.) which possess some level of resolution and accuracy regarding structural detail and source excitation.

These codes have also been used for the conduct of electromagnetics research. This application, however, is not their main strength. Electromagnetics research is more efficiently served by the many special purpose codes developed by the researchers themselves.

Supercodes can be summarized as those codes which are:

1. designed for the conduct of system-level analysis;
2. meant to be generally used by electromagnetic, antenna, and systems engineers throughout the industrial, academic, and government community;
3. intended to incorporate a number of techniques to efficiently treat the analysis at hand;
4. well documented to assist debugging and enhancement;
5. supported to provide a long-turn stable base to the user community;
6. highly transportable across classes of computer (PCs, workstations, mainframes) and manufacturers (IBM, DEC, SUN, etc.).

2.0 SUPERCODE DESCRIPTION

GEMACS has been continuously evolving since 1973. It is now fielded to the U. S. and foreign countries as Version 5. The code implements in the frequency domain three widely used and accepted CEM techniques; MOM, UTD, and finite differences (FD).

Moreover, these three techniques are fully hybridized within the code. This means that MOM and GTD can be used to treat appropriate regions and elements of the exterior of a structure, and FD can be used to solve for electromagnetic observables of interest within the structure. Coupling through the surface is achieved via the matching of boundary conditions in either dielectric- or air-filled apertures. All electromagnetic interactions among all exterior and interior components are fully taken into account in the calculation of the observables (e.g., fields, currents, impedances, etc.)

In addition to fully hybridizing three major CEM techniques GEMACS also contains a number of solution techniques in an attempt to expeditiously solve for the unknowns of interest (e.g., fields, coupling, RCS, currents). These include a full-matrix lower-upper decomposition (LUD) routine, a banded-matrix iteration (BMI) routine, and routines to take advantage of any rotational or planar symmetry present in the geometry. There is also a model order reduction (MOR)

routine to minimize matrix sizes when the FD technique is used to solve an interior region within the problem. Problems with 2000 complex unknowns (i.e., real and imaginary) have been solved using GEMACS.

GEMACS is general in the sense that it can be used to characterize ground, air, missiles, sea, and space platforms. To do this it uses wires, plates, cylinders, and endcaps to describe the geometry. Incident electromagnetic fields, antenna excitations, and antenna patterns are used to specify the EM sources which illuminate the geometry. Special provisions have been incorporated to input the mechanical and electrical data associated with phased arrays.

In addition, a separate PC-based graphics package is available to assist the analyst in developing the geometry. GAUGE provides the capability to dynamically develop and edit the computer model, as well as allow the user to view all or any specific parts of the model from any perspective in three dimensional space. The output of GAUGE is a geometry input data file directly readable by the GEMACS input processor.

The output of GEMACS can be read by GAUGE. The analyst is then able to display currents on MOM wires and antenna patterns in a variety of user-specified formats. Hard copy printouts from a laser printer or a plotter are also an available option from within GAUGE.

Several types of documentation are available for GEMACS, and for GAUGE to a limited extent. These include:

1. A user manual which describes the syntax of all commands used by GEMACS to set up the geometry and solution process;
2. An engineering manual which provides a derivation and discussion of the various CEM formulations and solution processes incorporated in GEMACS;
3. A code documentation manual which contains a detailed description of the modules, subroutines, variables, common blocks, and data transfer throughout the GEMACS program; and
4. A Source Book which is a compendium of information regarding the computational analysis of a system, treating the subject from both a theoretical and practical point of view. Numerous generic problems are described in detail (e.g., antenna pattern and coupling analysis, RCS calculation, and missile end-game).

GEMACS possesses all the other characteristics used to describe a supercode. It is written in FORTRAN-77 and has proven to be highly transportable (PC jr, AT class of PCs, workstations, and mainframes). It has been supported by Rome Laboratory and a number of companies (as a private

venture, not under contract to the government for this purpose). It is widely distributed within the provisions of government export control restrictions. Finally, GEMACS is supported by a number of companies which provide on-site and remote consultation services and formal training classes on a reimbursable basis.

This then is a summary of the characteristics of the GEMACS computer program, as well as a description of what is referred to as a CEM supercode. It now remains to examine the performance of a supercode to determine how well it does its intended job (i.e., system-level electromagnetic analysis) with the intent of attempting to specify to no more than a first-level what supercode future development and direction should be in order to meet the needs of system-level design and development in the year 2000 and beyond.

3.0 CURRENT SUPERCODE ASSESSMENT

This section describes an on-going application of the GEMACS program. This study stresses the external environment analysis capability of the code and can therefore be used as a vehicle to assess the strengths and weaknesses of today's supercodes. Analogous comments and conclusions can be made regarding the internal analysis capability using the FD formulation. Several areas will be

evaluated, including geometry modeling, execution time, and system observable data quantity and quality.

The geometry of the problem is shown in Figure 1. The platform is a KC-135 aircraft, very similar in shape to a commercial Boeing 707 aircraft. It is desired to determine the pattern of a phased array mounted on this platform as a function of location of the array with respect to the aircraft.

Figure 2 shows a CEM model of the aircraft. Since the aircraft is electrically large at the frequency of interest (1250 MHz), the UTD formulation is used to conduct the analysis.

The fuselage is modeled as a cylinder whose radius of curvature closely approximates that of the fuselage near the top of the aircraft. The nose of the cylinder is a flat endcap, and the rear of the fuselage is modeled by a slated endcap. This is a compromise in GEMACS. The BSC code does have a cone frustum which can more closely approximate the nose section of an aircraft. The best that can be done in GEMACS is to use a series of flat plates, tapered to match the aircraft nose geometry. Diffractions from the plate edges, corners, and intersections are not computed neither used in the calculations of antenna patterns.

The control surfaces of the aircraft (wings, horizontal and vertical stabilizers) and the engine pylons are modeled as infinitely thin flat plates. The

shape of these plates very closely approximates the two-dimensional projection of the actual surface. The contour and thickness in the third dimension are lost in this model.

The engines are each modeled as two cylinders intersecting end-to-end. Thus, the taper of the nacelle is approximated by a two-step function. Since the

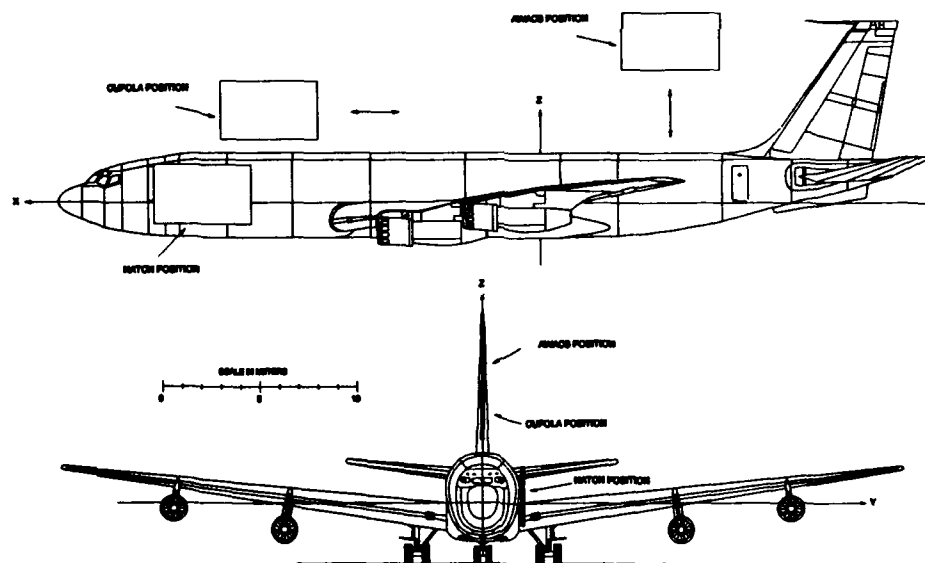


Figure 1. Pictorial Representation of Aircraft and Array Positions

engines are fairly well hidden by the wing from a top-mounted antenna, the use of a second cylinder may very well be an unnecessary embellishment for this study. Yet since the same aircraft model is used for antenna-below-the-fuselage studies (where the extra nacelle detail is needed), it was decided to use the more complicated model for consistency.

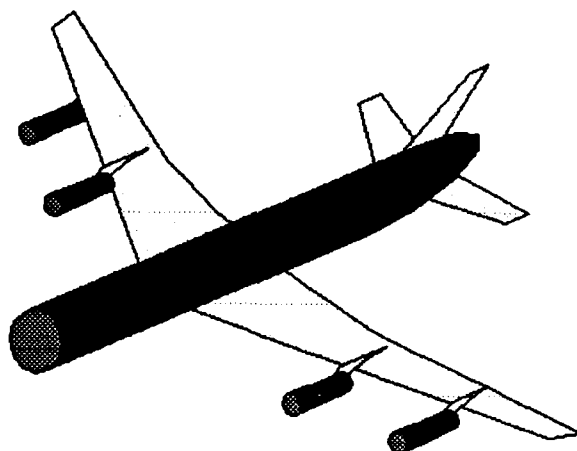


Figure 2. Computer Model of the KC-135

The array chosen as the test antenna is relatively simple. It consists of a grid of 900 half-wave dipoles mounted in a 30 x 30 configuration. A cosine-squared taper is applied to the array. In addition to boresight (normal to the plane of the array), the beam was scanned 45 degrees forward and aft of boresight. The array was variously mounted on the aircraft as shown in figure 1.

The array can be modeled as wires for the MOM formulation or as a series of equivalent half-wave dipole sources in the UTD formulation. The latter was chosen and a further approximation was introduced in order to reduce run times. The array is divided into 30 subarrays as shown in figure 3. This reduces the number of excitations from 900 to 30 and yields an identical free-space pattern. The effects of this approximation are only seen when blockage by the aircraft is present.

Figure 4 is an example of the type of data output by GEMACS for this problem. The figure shows the azimuth and elevation patterns through boresight for the array mounted on the aircraft (solid line) and the array in free space (dashed line). Without going into detail here, one can note pattern distortions. These are generally caused by wing and tail structure edges, as well as by blockage and reflection due to the surface area of the control surfaces.

Subarray Model

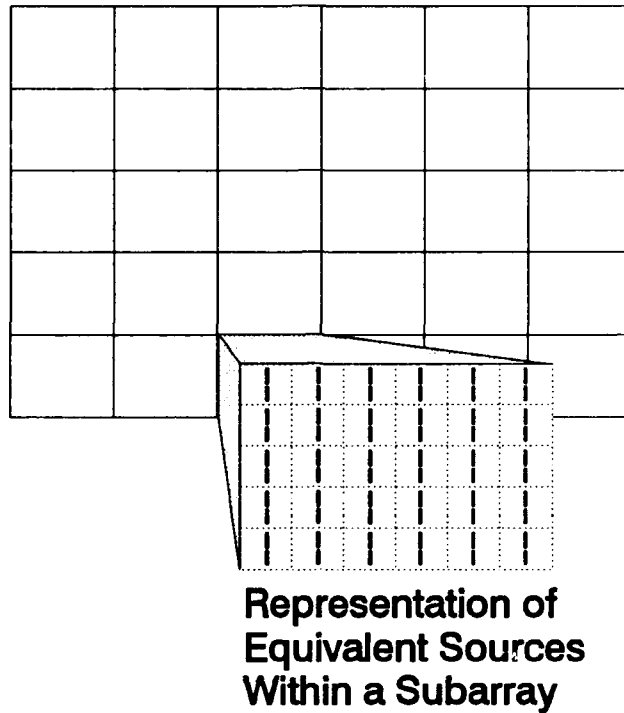


Figure 3. Subarray Representation of Array Antenna

This figure also demonstrates the effects of the subarray approximation. The approximation reduces the 30 distributed sources within a subarray to a single equivalent source at the physical center of the subarray. This serves to exaggerate blockage and other electromagnetic effects because of the electrically

large spacing among the equivalent sources for each of the 30 subarrays. This results in discontinuities in the pattern (e.g., at 160 degrees in the elevation pattern in figure 4), especially if an entire row or column of subarrays is

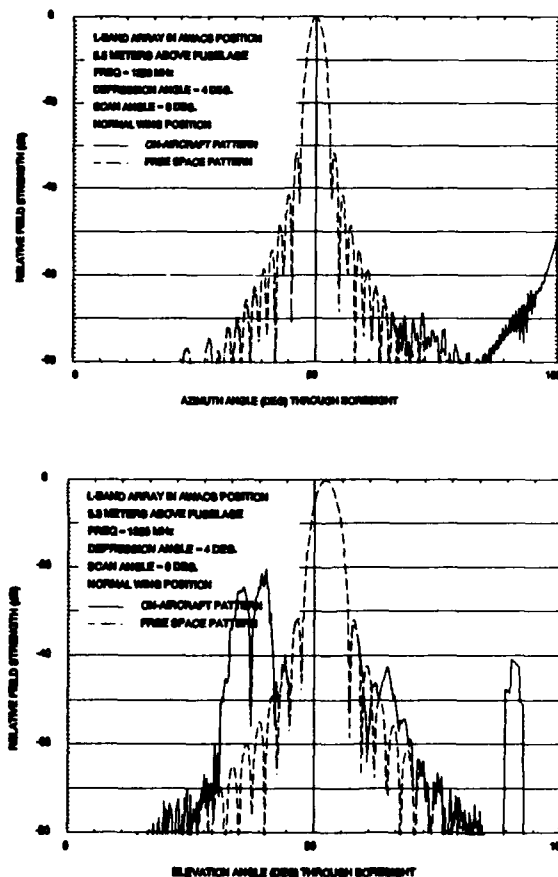


Figure 4. Test Array Sample Data

simultaneously affected.

An assessment regarding the capabilities and drawbacks of today's supercodes can be drawn from this study and from several others performed at Rome Laboratory. First, the studies demonstrate that a valid estimate of the electromagnetic radiating and scattering phenomena associated with a system can be obtained. The accuracy and resolution of the data are a function of many practical considerations, including costs for computer resources, time available for the study, modeling resolution and fidelity, CEM formulation utilized, equation solution technique, and computational numerics involved.

The point to emphasize, in spite of all these caveats, is that system-level analysis are possible and the data obtained from them provides an excellent engineering estimate of the electromagnetic performance of the system. Moreover, increasing system and antenna complexity as well as even more stringent operational requirements make them a necessity.

Other studies have validated through a comparison of measured and predicted data the hybridization scheme resident in GEMACS. It is possible to bring together and integrate the operations and results of a number of CEM techniques. This means that very complex problems can be broken down into component parts. Each of these is solved using the most appropriate technique,

taking into account the electromagnetic contribution of all other parts. The results of each process are then combined and the observables of interest are derived. This is shown in figure 5.

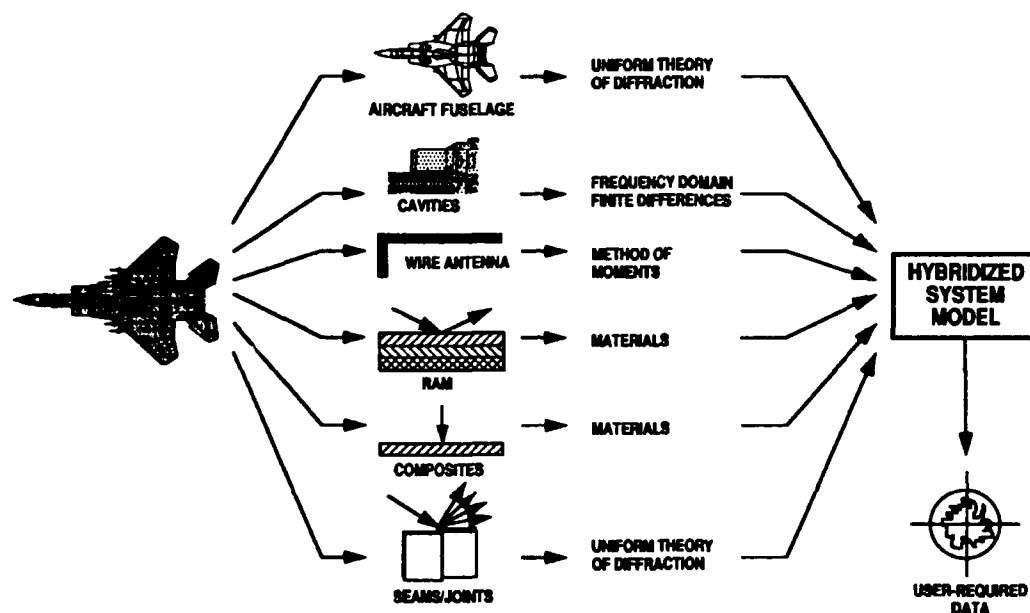


Figure 5. System Decomposition

This power and capability is not without a price. As useful as the current supercodes are, it is readily apparent that in their present form they are not able to meet the needs of the year 2000 and beyond. As a matter of fact, the vision outlined in the next section needs to be already in place to work on the design

problems now facing antenna engineers and system integrators.

That these codes are computationally intensive and long-running has already been pointed out. Analysts cannot afford the time required to fully, or even adequately, characterize a system to the depth and accuracy needed.

Second, the analysts have an extremely difficult and tedious task to accomplish when they begin to develop a computer model of the physical system, including its electric parameters. The graphical tools now available serve to make the job less miserable and prone to error. A tremendous amount of detail may be needed. One person dedicated for a month is required to generate even a very first-order, high-level approximation to the physical geometry.

Third, in many cases there is not available an adequate repertoire of modeling elements to generate a high fidelity computer model of the physical structure. Analyses of wire structures (antennas, transmission lines, etc.) using the MOM formulation are an exception. However, all geometries of interest (aircraft, ships, etc.) are complex and not reducible to cylinders, cones, cubes, and other available modeling elements. The study described in this section used a model whose radius of curvature matched that of the fuselage only in certain regions. The nose of the model is totally inadequate. All structural detail in the third dimension of the control surfaces is ignored. Such detail may be of extreme

importance when dealing with creeping waves and materials with arbitrary values of conductivity, permittivity, and permeability. The strong field shown at an azimuthal angle of 180 degrees in figure 4 is very likely a result of the thin edge used to model the leading edge of the vertical stabilizer.

Although not a major electromagnetic contributor to this particular study, the model of the engine nacelles is woefully inadequate. They have a complex shape at the inlet port. They are filled open-ended cavities. They have a complex taper from front to back. They cannot continue to be modeled as closed UTD cylinders for either antenna pattern calculations or radar cross section and target identification studies.

The list of modeling issues and inadequacies can be extended, especially when one considers radomes and the amount of dielectric and radar absorbing materials now being added to the basic structure for one purpose or another.

One final aspect is that regarding the computational numerics involved in today's supercodes. The process begins with exact integral or differential equations. Series approximations are used to translate these equations into a form that can be solved by a digital computer. Each type of series approximation carries with it some residual error. It has been a judgment call to say when a "small" residual error is small enough.

The original loss of accuracy generated by neglecting the residual error is compounded by the millions of multiplications and divisions which are carried out during the execution of the code. This occurs in matrix multiplications to generate a MOM interaction matrix, or it occurs in multiplying a diffraction coefficient by a propagation distance in the UTD formulation.

Finite length computer words also contribute to the problem. These are typically 32-bit words. Studies have shown that the use of double precision variables does eliminate or reduce the problem. This shows that there is a problem. It also implies that eventually even the double precision computation will break down.

A major portion of these supercodes was developed in the 1970 decade. Numerical processing algorithms were less precise than they are now. Thus, even the computational process adds unnecessary "noise" to the calculations. For example, in UTD calculations it quite often happens that two large, nearly equal numbers will be each squared and then the squares will be subtracted from each other. The result is a number which may be 10^{-4} or smaller. The square root of this number is then taken, and the result may then be cubed. Eventually this ever-decreasing number is lost, which may represent a phase error in near-field scattering.

The point is that other, more accurate ways of performing computations are now available. The coding practices used to generate the existing supercodes need to be carefully examined to determine if simple improvements can be made which will yield observable data with a greater degree of accuracy.

It needs to be restated at this point that in spite of the shortcomings that have been described, today's supercodes are nonetheless useful determinants of system operation. With intelligent use they do provide valid data, as comparisons with measured have shown.

However, the requirements of upcoming analyses beg for a new generation of CEM tools. These include the characterization of ultralow sidelobe antennas (ULSA) mounted on aircraft such as the KC-135, the characterization of bodies manufactured with low observable (LO) materials, ultrawideband (UWB) operation, and the migration to ever-higher frequencies of operation (e.g., 94 GHz) at which seams and cracks become apertures of significant electrical size. What do requirements such as these imply regarding the capability, and indeed the development, of the next operation of supercodes?

4.0 A TOP-LEVEL SPECIFICATION

This section contains one view of what is needed to arrive at a supercode

useful for the year 2000 and beyond. The primary function of the code is to analyze the performance of antennas mounted on complex, large platforms. Antenna performance parameters of interest include far-field patterns, gain, coupling to other antennas, coupling through apertures in the skin of the platform and ending up as unwanted power at the terminals of sensitive devices, and antenna terminal parameters.

It is furthermore desired to account for any parameters intrinsic to the materials out of which the structure is constructed. The representation of layered-structures by a surface impedance may be adequate for nearly all studies for which the code is used. For example, a 5-layer radome may be analyzed by a code developed specifically for that purpose, and an equivalent surface impedance derived from the data regarding the individual layers may be used by the CEM code.

This implies that the supercode will be concerned with structure elements on a macroscopic level, or at a level which averages structural characteristics and responses over an electrically large area (perhaps a radius of several wavelengths, depending on frequency). As is the case with current CEM analyses, the resolution of the geometric and electrical parameters associated with the system under analysis will be driven by the desired quantity and quality of the observable

data.

Beyond this consideration, several other major categories of design criteria must be factored into the generation of the next generation of supercodes. These are discussed under the headings of execution run-time, CEM formulation, modeling fidelity and strategy, and the computational numerics involved in the problem solution.

4.1 Execution Run-Time

This is an extremely critical factor today and will always continue to be the key to the completeness with which a system is characterized. Future systems will be very wideband, operate at higher frequencies and have antenna patterns with ultralow sidelobe levels. In addition, they will be mounted on platforms which have extremely convoluted shapes and are made of materials which are anisotropic on a somewhat macroscopic level. This latter already exists in that boron and carbon impregnated materials have the doping material running in only one direction, depending on the direction in which the extra strength is needed.

What will result then is electrically large structures requiring more detail, modeled with greater fidelity, and observable data requirements orders of magnitude greater than those which now exist. This implies more computations at higher levels of accuracy and resolution. To accomplish this the supercode

must be designed for and installed on massively parallel processors. Figure 6 shows a comparison of an identical analysis performed on a number of different machines. All but one, the 1024-node hypercube, are sequential processing platforms.

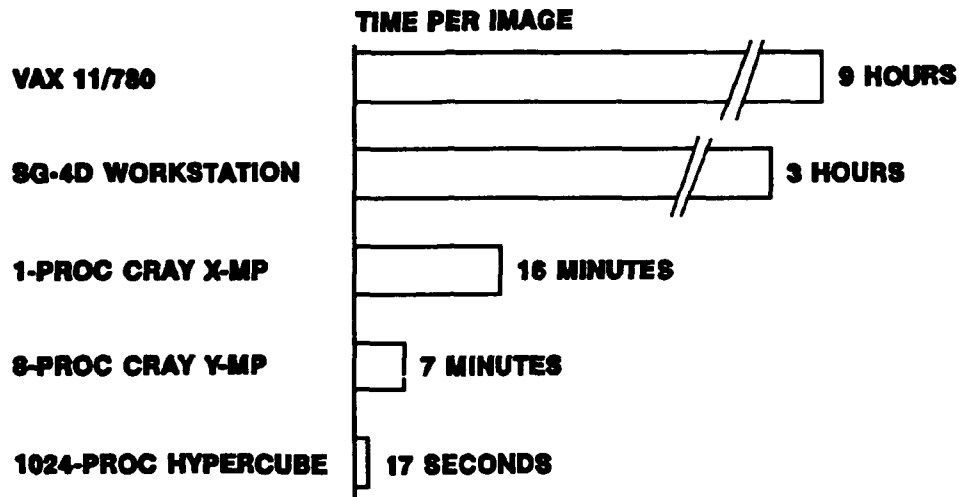


Figure 6. Execution Time Comparison

Irrespective of what other advances are made in the area of CEM, the move from sequential to massively parallel processors is critical.

Several other issues then arise. It is beyond the scope of this paper to deal with these issues, other than to name them. These issues include the development

of efficient, "low noise" algorithms and computation processes, node communications to achieve optimum use of the available processors, design for transportability across machine types and manufacturers, and design for growth and enhancement in order to incorporate the results of electromagnetics and computational numerics research. Complete documentation, in-depth training, and centralized distribution and maintenance must also be available.

4.2 CEM Formulation

The next generation of supercodes must include and hybridize a broader base of analysis techniques. These include physical optics, characteristic modes, and others. Each of these may be applicable to only a narrow range of problem sets, but as an aggregate they will allow an analyst greater freedom in employing the appropriate technique for each element of geometry into which the system has been divided (e.g. the cavity regions of an engine nacelle).

Time domain techniques need to be developed and incorporated into the structure of the supercodes. With very few exceptions, notably EMP studies, the current supercodes operate in the frequency domain. This will not be satisfactory for future systems which will operate across broad ranges of frequency and ultrawide bandwidths. There will be a need to characterize the system quickly in an attempt to isolate system resonances which can then be investigated in

greater detail using the frequency domain tools as appropriate.

4.3 Modeling Fidelity and Strategy

It will be necessary to have available as complete and detailed as possible a computerized description of the platform. This is already available for many aircraft platforms which use computer-aided design and computer-aided manufacturing (CAD/CAM) techniques. These will need to be augmented to include operational antennas and other subsystem-specific structures such as beamforming materials and fences.

It is not reasonable to expect that all this data will ever be used in any individual analysis. Its existence enables an analyst to simply "point and shoot" to select those elements of the structure which will eventually be transformed into the modeling elements of the CEM formulations which will be used. Thus, geometry input will be simplified and also made more accurate.

The limiting factor in geometry modeling will be the modeling elements available in the CEM formulations. Today an analyst can very accurately generate a computer model of the nose radome of an aircraft. However, there is not any element in the UTD formulation which can be used to describe this shape to the code in order to calculate its scattering, reflection, and creeping wave phenomena.

Therefore, electromagnetics research will need to develop appropriate scattering formulations for curved patches (curved in two dimensions). These will need to be suitable for both MOM and UTD formulations. They must also be capable of handling dielectric and magnetic materials, as well as containing filled or unfilled apertures within the boundaries of the doubly curved patch.

In general models for seams, cracks and electrically large apertures must be available. As frequencies of interest increase, the wavelength decreases and seams between similar and dissimilar materials become efficient radiators of energy both interior to and exterior to the skin of the platform.

This also holds true for any perturbations of the surface, such as aerodynamic stabilizing fins and lightning rods. Once again, as the wavelength decreases the dimensions of these objects become significant fractions of a wavelength. Their contribution to the total field may be small, but they may be significant if the system involves an antenna with ultralow sidelobe levels.

This is yet another justification for the inclusion of many additional CEM formulations, no matter how limited they may be in their applicability to general geometries. If they can efficiently and completely treat only the seam or the small fin, then they will relieve the computational burden incurred in the use of the MOM, UTD, or FD formulations to treat these specialized geometries. A

solved seam or fin can be used as an equivalent source in the UTD formulation or as the generator of an incident field in the MOM or FD formations.

However, what elements of the geometry to include in the analysis or what formulation to use on specific elements of the geometry, is a matter of judgment. Such electromagnetic decision-making is a matter of intuition and experience. In addition, with greater computing power available there may be the temptation to "over-model" the problem at hand. For example, in the experiment described in the previous section, the engine nacelles were modeled as two end-to-end intersecting cylinders. This was done in spite of the fact that the rear part of the nacelle was effectively blocked from the view of the antenna by the wing structure. This unnecessarily increased execution run-time without adding any electromagnetic energy or interaction to the analysis at hand.

A new science of modeling strategy in the form of rules and guidelines must be established. The next step would be the development of an artificial intelligence (AI) tool that guides and aids the analyst in the development of an adequate model suitable to the problem at hand. Such a tool would eventually be tied into the CAD/CAM database for the platform and could generate the complete data input stream for the supercode, via an interactive dialogue with the analyst. This would effectively reduce the level of expertise required of the

analyst and make the supercode more accessible to a wider community of antenna engineers and system designers/integrators.

4.4 Computational Numerics

The final area to be considered here has already been alluded to several times. As the number of computations increases and the analysis requirements for extended dynamic ranges (on the order of 10 million or more) become more established, the manner in which intermediate computations are carried out becomes critical. Increased computational power, fidelity and resolution of model, and the existence of suitable CEM formulations will be negated if computer "noise" masks the least significant decimal places of the intermediate data. This is especially true of phase-sensitive systems and multiple-bounce near-field diffractions and reflections. Other critical areas are radome effects on phased array beam patterns (especially when they are scanned off boresight) and dielectric-filled slots or open-ended waveguide (when used as elements in phased arrays).

As always has been the case, the problem becomes more critical at the higher frequencies of operation.

5.0 SUMMARY

This paper attempts to highlight some of the major issues that impact the area of computational electromagnetics. These include, but of course are not limited to, the following:

- Execution time to complete an analysis
- Modeling resolution and fidelity
- Electromagnetic models to characterize the phenomena
- Computational numerics to reduce the "noise floor"

There are other issues which surround a supercode, which are more practical rather than technical. These include:

- documentation, completeness and currency
- training, for both computer system managers and engineers
- transportability of the code across platforms
- maintenance and distribution
- growth potential
- user interface

One aspect of this development of the next generation of supercodes has

yet to be addressed. That aspect is the programmatic one, which is just as critical as any one of those already mentioned.

Four questions can be formulated to describe this aspect:

- Who will do it?
- How will they go about doing it?
- How much will it cost to do it?
- How long will it take to do it?

A discussion of the programmatic involved in the development of the next generation of electromagnetic supercodes is beyond the scope of this paper—and the crystal ball of its author.

Fast Algorithms for Wave Scattering Developed at the Electromagnetics Laboratory, University of Illinois *

WENG CHO CHEW
ELECTROMAGNETICS LABORATORY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
UNIVERSITY OF ILLINOIS
URBANA, IL 61801

Abstract

Several fast algorithms for wave scattering solutions developed at the Electromagnetics Laboratory at the University of Illinois are reviewed. The computational complexity of each of the algorithms is given together with some simulation results, validation of simulation results, and comparison of CPU time with methods that require Gaussian elimination. A brief mathematical description of an algorithm is given together with discussion on the significance of such research work.

1. Introduction

As computers grow faster and larger, more engineering and scientific problems are solved with computational techniques. Hence, computers have become an indispensable tool in both engineering and science because they expand our horizons in the understanding of problems around us. Although the speed of computers have grown by leaps and bounds in the past decades, there will be a stage that the algorithms used to solve many engineering and scientific problems are more important than the hardware itself.

Solving complex wave-scattering problems have been the preoccupation of scientists and engineers alike because it helps us understand many physical phenomena around us as well as a large number of engineering designs. Examples of such applications include remote sensing, oil prospecting, inverse scattering and imaging, and computer-aided engineering and design.

The algorithms to solve many scattering problems have CPU times that grows as CN^α where N is the number of unknowns in the problem. In such a case, we say that an algorithm has a computational complexity of $O(N^\alpha)$. When α is large, the complexity of the algorithm is bad. It has been long known that when scientific and engineering problems are solved with an algorithm of poor complexity, they can swamp easily the largest and the fastest computers we have. This has been known as the "cruelty of computational complexity" as illustrated in Figure 1.

It is the goal of this research to reduce the computational complexity of a

* This work was supported by the Office of Naval Research under grant N00014-89-1296 and the Army Research Office under grant DAAL03-91-G-0039. The computer time was provided by the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign.

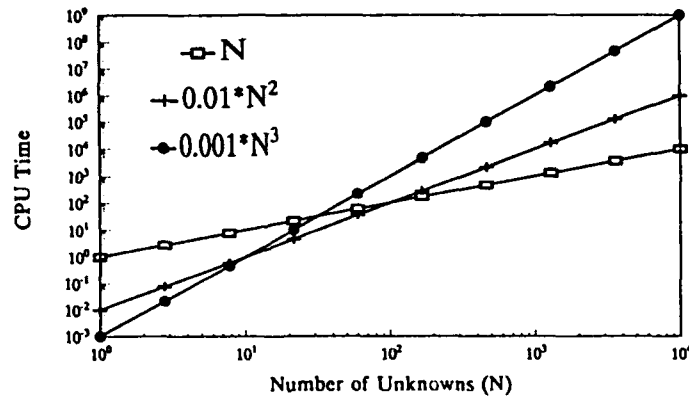


Figure 1. Cruelty of computational complexity. Three algorithms with $\text{CPU} = N$, $\text{CPU} = 0.01N^2$, and $\text{CPU} = 0.001N^3$, with the eventual swamping of the computational resources by the algorithm with the worst complexity.

class of scattering problems. With reduced computational complexity, larger problems can be solved faster and even on smaller computers. The ensuing technological impact can be tremendous.

2. Review of Wave Scattering Solution Techniques

There are two main solution techniques (or solvers) for solving wave scattering problems. One method is to solve the differential equation associated with the wave-scattering problem directly by converting the differential equation by discretization into a sparse matrix. Alternatively, an integral equation can be derived for the induced current in the scatterer or equivalent current on the scatterer. When the unknowns are volume currents in the scatterer, the associated integral equation is called the volume integral equation, whereas when the unknowns are surface current or equivalent surface current on the surface of the scatterer, the associated integral equation is called the surface integral equation (or boundary integral equation). Again, the integral equation can be converted into a matrix equation by discretizing the unknown currents.

2.1. Differential Equation Solvers

Differential equation solvers involve the use of the finite element method (FEM) [1-4] or the finite difference time domain (FDTD) method [5,6]. In the FEM, the space around a scatterer is discretized into many tiny finite elements with about 10 to 20 elements per wavelength. By so doing, the differential equation is converted into a matrix equation with N unknowns. The matrix equation is then solved with the conjugate gradient (CG) method [7], the preconditioned conjugate gradient method, the banded matrix solver, the nested dissection method [4], or the multigrid method [3]. In these ap-

proaches, the CPU time grows as N for multigrid applied to elliptic equations, $N^{1.25}$ for preconditioned conjugate gradient method, $N^{1.5}$ for conjugate gradient method and the nested dissection method, and N^2 for a banded matrix factorization [1, p. 388]. Though faster than a banded matrix factorization, the disadvantage of the conjugate gradient method is that it is valid for one excitation source only. Moreover, it is not often easy to find a good preconditioner for the preconditioned conjugate gradient method.

Recently, the finite difference time domain method has become quite popular for solving scattering problems [6]. In this method, the differential equation is converted into a finite difference equation. Then, time-stepping and causality are used to solve the equations from the initial values of the field at some initial time. No matrix inversion is involved, and the memory storage requirement is proportional to N , where N is the number of grid points used to model the problem. The CPU time for such a method grows as $N^{1.5}$ [8, p. 241]. The disadvantage is that like the conjugate gradient methods, the solution is valid for only one excitation source.

Differential equation solvers are usually implemented in a finite region of space. For an open-space problem, absorbing boundary conditions have to be imposed at the edge of the simulated space to approximate an open space problem [8, p. 246, 9-11].

2.2. Integral Equation Solvers

In the volume integral equation solver, a volume integral equation [8,12,13] is solved for the field instead. The scatterer, rather than the entire space, is discretized into many tiny elements with about 10 to 20 elements per wavelength. Then the integral equation is converted into a matrix equation. A straightforward method for solving such a matrix equation will entail a computational complexity of N^3 where N is the number of elements or unknowns characterizing the scatterer. Alternatively, the conjugate gradient (CG) method can be applied to solve the matrix equation iteratively. This reduces the complexity to N^2 per iteration of the CG method. Hence this method has a computational complexity of $N^{2+\alpha}$ where N^α is the total number of iterations that yields an answer with a desired accuracy, and $\alpha < 1$. By noticing the convolutional property of the integral in the integral equation, the conjugate gradient method together with fast Fourier transform (CG-FFT) [14-19] can solve the resultant matrix equation rapidly with complexity $N^{1+\alpha} \log N$.

Integral equation solvers have the attractive features that the radiation condition at infinity is automatically satisfied by the solutions. Moreover, less unknowns are needed to solve a problem of a given size compared to the differential equation solvers because a differential equation solver needs to discretize the space around the scatterer as well as the scatterer itself. The disadvantage of the integral equation solvers is that the CPU time grows faster than that of the differential equation solvers when the size of the scatterer increases (except for the CG-FFT method). As a result, the differential equa-

tion solvers have been quite popular compared to integral equation solvers. Therefore, unless fast algorithms are developed to solve integral equations, they may not become as popular as differential equation solvers.

On the other hand, integral equation solvers for surface integral equations have become quite popular in the recent past because the number of unknowns that are associated with integral equation solvers is usually quite small compared to volume solvers (be it differential equation solvers or volume integral equation solvers) [20,21]. However, not all problems can be formulated in terms of surface integral equations.

As expected, surface integral equations can also be solved by the conjugate gradient method. However, each iteration of the conjugate gradient method entails two matrix-vector multiplications which has a complexity of $O(N^2)$. Recently, a fast multipole method (FMM) [22] has been proposed to enhance the speed of this matrix-vector multiplication. Such solution technique, however, is valid only for one illumination of the scatterer with one excitation source.

Recently, we have developed several integral equation solvers [23-36] providing a full solution valid for all excitation sources (or all incident angles), and yet competitive in computational complexity to the banded matrix factorization for the finite element method. The advantage of our solution techniques over that of the differential equation solvers, however, is that it generally has less unknowns compared to them, and a lot less book-keeping is required for instance, when compared to the finite element method. These algorithms will be discussed in the next section.

3. Fast Algorithms We Have Developed

We will summarize here a chronological development by our research group of fast algorithms for solving integral equations. The main theme of our work is that it supplies a solution valid for all excitation sources. The algorithm is in general recursive and noniterative—i.e. the solution is sought in N steps or less. The use of recursive algorithm is not new. It has been used in solving for the reflection coefficients of layered media [8]. More recently, it has been generalized to solving for the solution of inhomogeneous media [37,38], or inhomogeneous strips in layered media [39,40].

As mentioned before, a full solution to an integration equation can be obtained with a computational complexity of $O(N^3)$. The following algorithms described will have computational complexity less than N^3 . Before we proceed on the description of these algorithms, we shall list them as follows:

- **RTMA**—Recursive T Matrix Algorithm
(1989), Computational Complexity $< N^{2.5}$ [23,24,29].
- **RATMA**—Recursive Aggregate T Matrix Algorithm
(1990), Computational Complexity $< N^2$ [26-29].
- **GRATMA**—Generalized Recursive Aggregate T Matrix Algorithm
(1991), Computational Complexity $< N^2$ [30,31].
- **WRATMA**—Windowed Recursive Aggregate T Matrix Algorithm

(1992), Computational Complexity $< N^2$ [32,33].

- **RATMA-PWB**—RATMA with plane wave basis (1992), Computational Complexity $\sim N \log^2 N$ [34,35].
- **NEPAL**—Nested Equivalence Principle Algorithm (1992), Computational Complexity $\sim N^{1.5}$ [36].

3.1. Description of Fast Algorithms We Have Developed

- **RTMA**—Recursive T Matrix Algorithm (1989, 1990)

In this algorithm, a large scatterer is first divided into N subscatterers and the T matrix [41] of each subscatterer is found. Then, recursively, an n -scatterer problem is used to solve the $(n + 1)$ -scatterer problem yielding a solution in N steps. Addition theorem is used to represent the translation operators that translate the wave harmonics from one scattering center to another scattering center. In this method, all interactions between the scatterers are accounted for. Moreover, the computational complexity can be shown to be $< N^{2.5}$. The validation of RTMA against the method of moments (MOM) [13] with Gaussian elimination is shown in Figure 2.

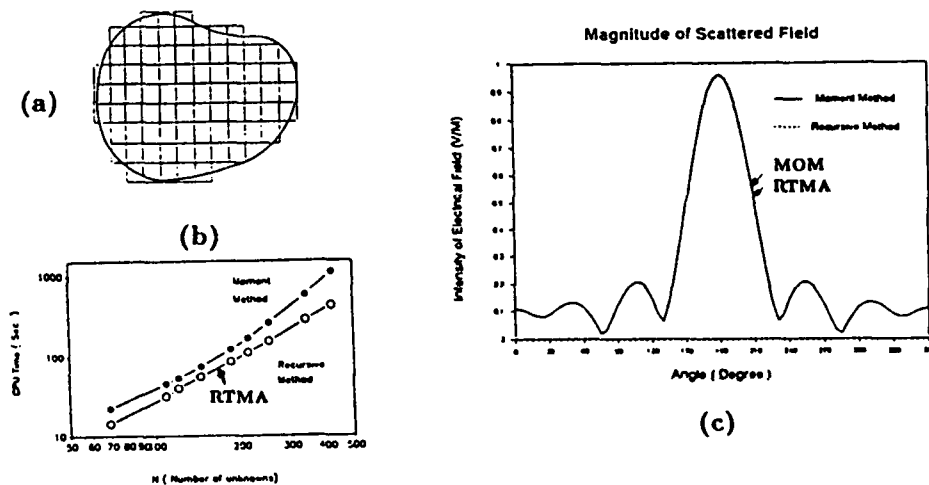


Figure 2. (a) Discretization of an inhomogeneous scatterer. (b) Comparison of RTMA against the traditional method of moments (MOM) with Gaussian elimination in CPU time versus the number of unknowns. (c) Validation of RTMA against the solution of MOM.

- **RATMA**—Recursive Aggregate T Matrix Algorithm (1990)

Again, in this method, an n -scatterer solution is used to solve the $(n + 1)$ -scatterer solution. However, differing from RTMA, an aggregation of the previous n scatterers into one scatterer is performed at each stage so that only a two-scatterer problem is considered at each stage. A conservative estimate of the computational complexity shows that it is an $O(N^2)$ algorithm. However, if grid dispersion error is taken into account, its computational complexity can be as low as $O(N^{1.66})$. In Figure 3, we show the comparison of RTMA,

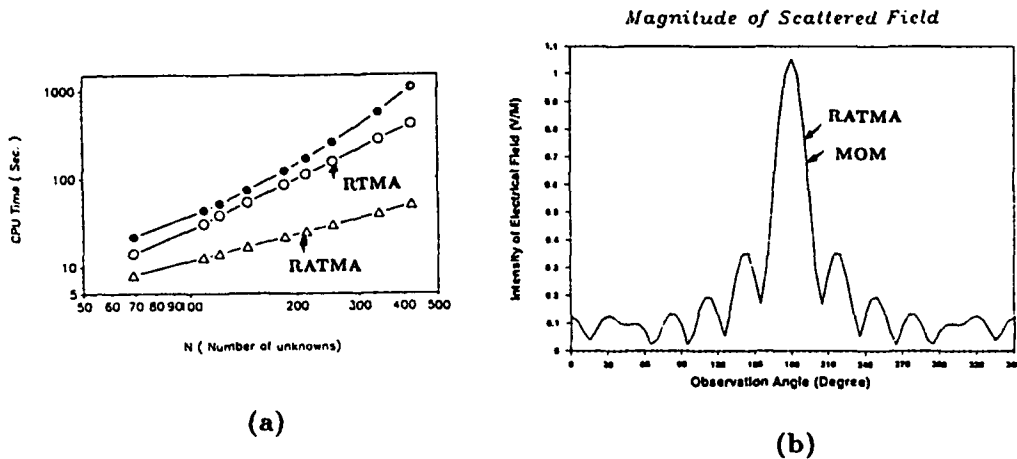


Figure 3. (a) Comparison of RATMA against RTMA and MOM in CPU time versus the number of unknowns. (b) Validation of RATMA against the solution of MOM.

RATMA, and MOM. The result of aggregation is a dramatic saving in computational time.

To test the stability of the algorithm, RATMA has been used to solve for the solution of a large scatterer with 12,000 unknowns. A 12,000 unknown problem will require 20 hours on a CRAY using Gaussian elimination but only 30 seconds with RATMA as shown in Figure 4.

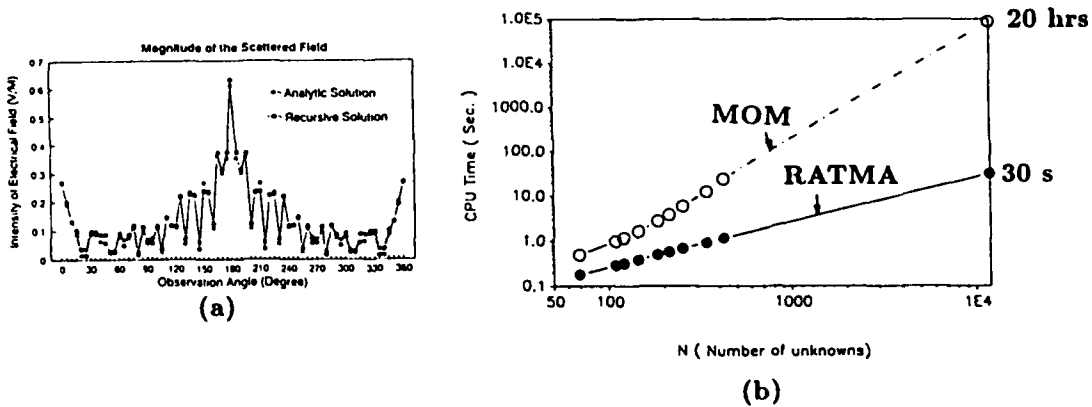


Figure 4. (a) Comparison of RATMA against analytic closed form solution of the amplitude of the scattered field for a scatterer discretized into 12,000 unknowns. (b) Comparison of CPU time of RATMA (solid dots) versus MOM (white dots).

To show that RATMA can be applied to objects of oblong shapes, we have

applied it to coated metallic scatterers of oblong shapes and compared the solution with the method of moments [42] as shown in Figure 5. The method of moment solution has to be solved on the CRAY, while the RATMA solution is solved on an IBM PC.

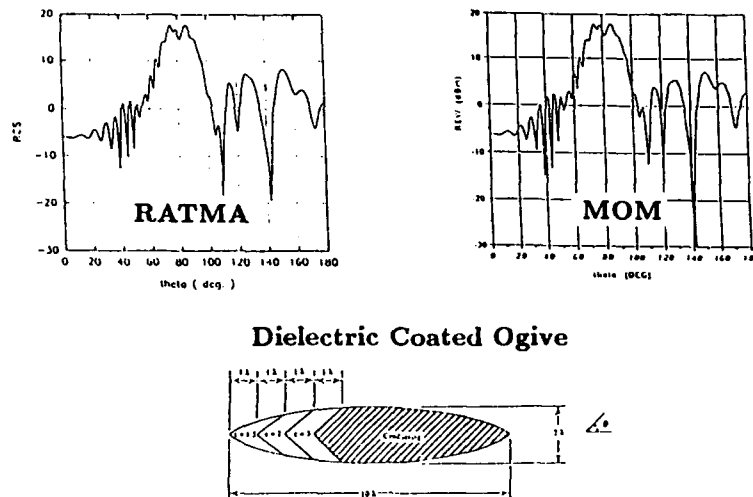


Figure 5. Comparison of Radar Cross Section (RCS) for oblong shaped coated metallic scatterers using RATMA and MOM. The RATMA solutions are obtained on an IBM PC while the MOM solutions are obtained on a CRAY.

• GRATMA—Generalized RATMA (1991)

RATMA works very well when applied to E_z polarized waves. However, when applied to H_z -polarized waves, it fails. This has been traced to the more singular nature of the integral equation associated with H_z -polarized waves. As a result, the addition theorem, which has a dichotomous regions of validity, could be badly violated. To overcome this problem, a generalized RATMA has been developed.

In this generalized algorithm, an n -scatterer solution is used to solve for the $(n + n')$ -scatterer solution to overcome the singularity problem for H_z -polarized wave scattering. Such a generalization does not increase the computational complexity of the algorithm which is still $< N^2$. Figure 6 shows the result of GRATMA on the scattering solution of the H_z -polarized wave.

The scattering of electromagnetic waves by a three-dimensional inhomogeneous object is inherently a vector problem with a singular Green's function. As a result, the addition theorem is always violated. However, GRATMA can be used to solve such a scattering problem. Figure 7 shows the result of GRATMA used for the simulation of three-dimension electromagnetic scattering by a cluster of spheres [43].

• WRATMA—Windowed RATMA (1992)

Yet another way of mitigating the effect of the violation of the addition theorem is to use a windowed addition theorem in RATMA. In this manner,

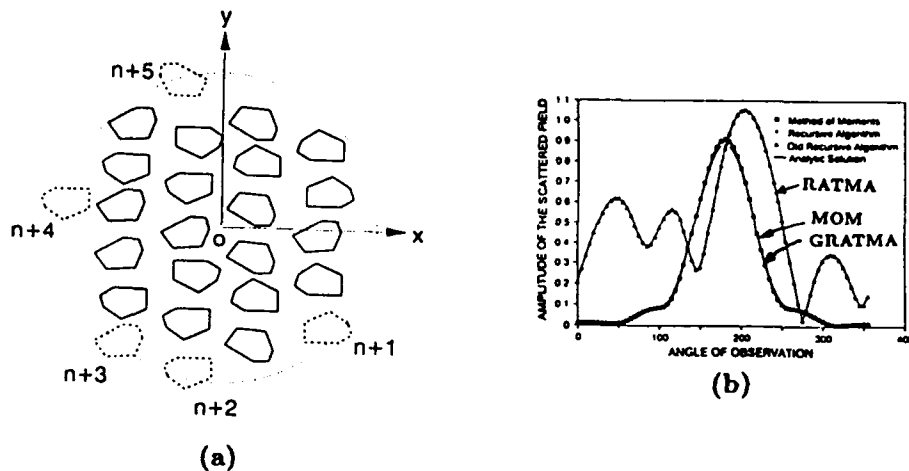


Figure 6. (a) In GRATMA, n' subscatterers are added to the previous n subscatterers all at once. (b) GRATMA rectifies the problem faced by RATMA in the H_z scattering problem.

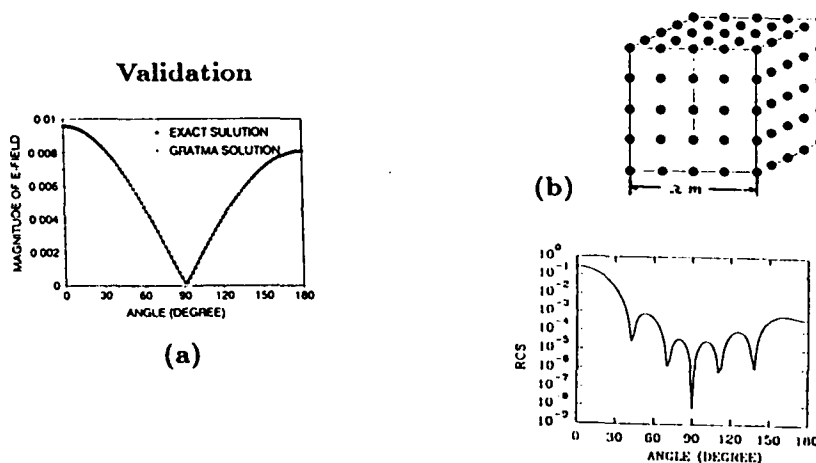


Figure 7. (a) The solution is validated by packing a cluster of spheres into a bigger sphere. The scattering solution is computed using both GRATMA and a closed-form solution. (b) The RCS of a finite cluster of an array of 125 spheres calculated by using GRATMA.

the solution is made to converge well. This can be done without increasing the computational complexity of RATMA. Figure 8 shows how windowing can rectify the problem faced by RATMA for solving the scattering by H_z -polarized waves.

• RATMA-PWB—RATMA Using Plane-Wave Basis (1992)

The aforementioned fast algorithms have all used cylinder harmonics or spherical harmonics as basis for the expansion of the incident field and the scattered field. Alternatively, the incident field and the scattered field can

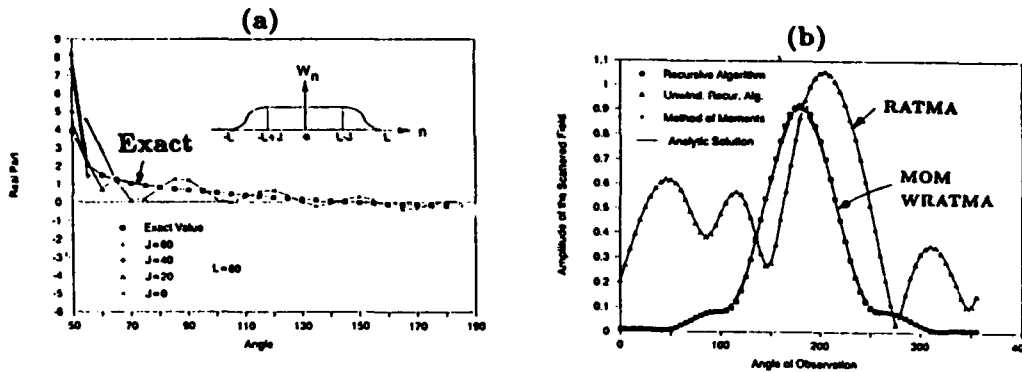


Figure 8. (a) The removal of the Gibbs' phenomenon in the addition theorem by windowing. (b) The divergence of the solution is removed after the windowed addition theorem is used in RATMA (WRATMA).

be expanded in terms of plane waves. It turns out that for a coplanar array of strips (see Figure 9a), the plane-wave basis is a more suitable basis than cylindrical harmonics or spherical harmonics. Moreover, the plane wave expansion of the scattered field can be made very efficient by choosing the plane-wave basis along a branch cut in the complex plane (see Figure 9b). A variant of RATMA can be derived using this idea, resulting in an algorithm with $N \log^2 N$ complexity for one incident plane wave excitation, and $N^2 \log N$ complexity for N incident plane wave excitations.

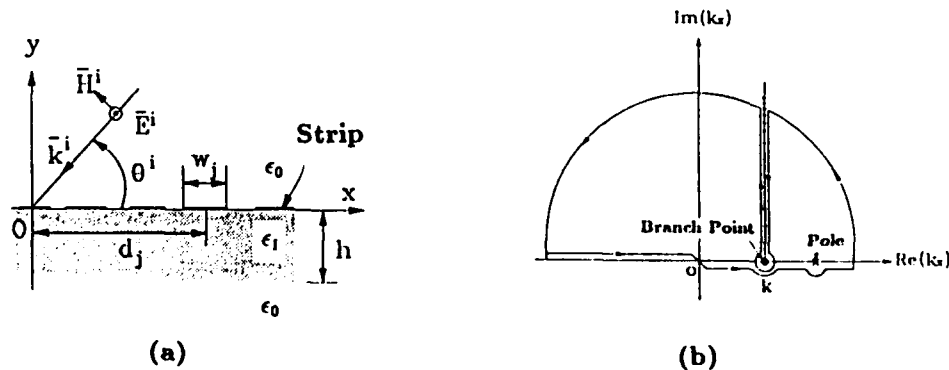


Figure 9. (a) An array of coplanar strips supported by a dielectric slab. (b) The choice of the plane-wave basis along a vertical branch cut on the complex plane.

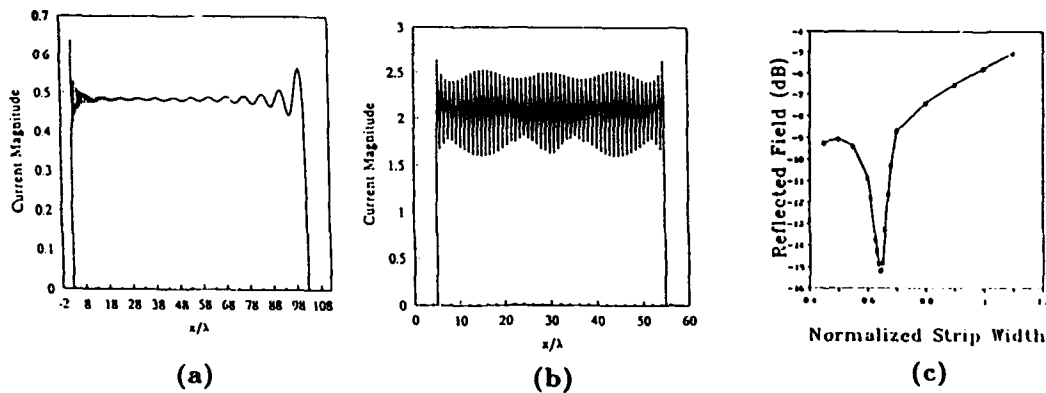


Figure 10. RATMA-PWB can be used to calculate for the scattering solution of a wide array of strips. (a) The induced current on a 100λ strip illuminated by an H_z polarized wave. (b) Induced current on a wide strip backed by a dielectric slab. (c) RATMA-PWB can be used to calculate the reflection coefficient of a finite array of strips in frequency-selective surface applications.

The resulting algorithm can be used to calculate the scattering of electromagnetic waves by a wide strip of several hundred wavelength wide. Interesting interference phenomenon due to the presence of surface waves is observed when a large scatterer is illuminated by an H_z -polarized wave. For a wide strip backed by a dielectric slab, the interference phenomenon is due to the simultaneous presence of a slab guided wave and a surface wave.

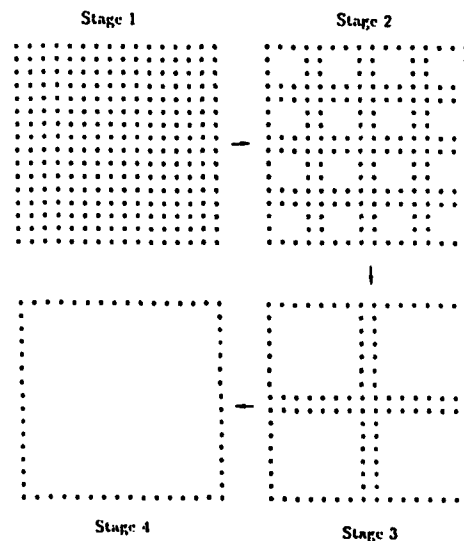


Figure 11. The Nested Equivalence Principle Algorithm involves nesting a smaller problem within a larger problem to reduce computational complexity.

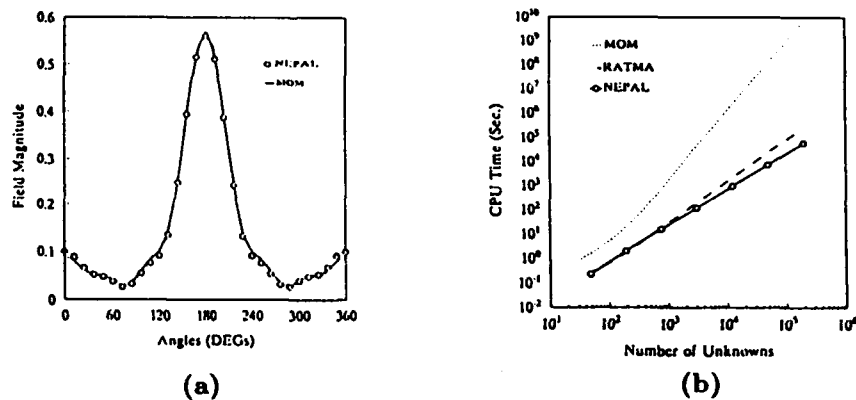


Figure 12. (a) Validation of NEPAL against the MOM. (b) Comparison of the CPU time versus the number of unknowns for NEPAL, RATMA, and MOM.

• NEPAL—Nested Equivalence Principle Algorithm (1992)

Recently, we have developed a new method of solving volume integral equations called the Nested Equivalence Principle Algorithm (NEPAL). In this approach, the scatterer is first subdivided into N subscatterers. Then, the N subscatterers are divided into subgroups, each consisting, for instance, of 16 subscatterers. The scattering solution from each of the subgroups is solved separately. Then the scattering centers interior to a subgroup of subscatterers are removed using Huygens' equivalence principle (see Figure 11). Hence, in Stage 2, the number of unknowns is less than that in Stage 1. By the same token, four subgroups in Stage 2 are grouped together into larger subgroups, whose scattering solutions are found separately. After that, the interior scattering centers of the larger subgroups are removed to go to Stage 3, and the process is repeated until all the interior scattering centers are replaced by surface scattering centers. The total replacement by surface scatterers requires $\log_4 N$ steps. It can be shown that this algorithm has a computational complexity of $N^{1.5}$. In Figure 12, we show the validation of the result of NEPAL by MOM, and the comparison of the CPU time of NEPAL against those of RATMA and MOM. The result shows that NEPAL has an asymptotic computational complexity of $O(N^{1.5})$. NEPAL is reminiscent of the nested dissection method [4].

4. Mathematical Description of RATMA*

For those who are mathematically inclined, we shall present the derivation of RATMA here. In the RATMA, an inhomogeneous scatterer is first divided into N subscatterers. The solution to the N subscatterer problem is

* This section could be skipped without loss of continuity.

constructed recursively. In other words, the n -subscatterer solution is used to construct the $(n + 1)$ -subscatterer solution. In this manner, the N subscatterer solution is obtained recursively from a one subscatterer solution.

When the n -subscatterer solution is known, the total field outside the n subscatterers can be expressed as

$$\phi(\mathbf{r}) = \Re g \psi^i(k_0, \mathbf{r}_0) \cdot \mathbf{a} + \psi^i(k_0, \mathbf{r}_0) \cdot \bar{\mathbf{T}}_{(n)} \cdot \mathbf{a}, \quad (1)$$

where $\bar{\mathbf{T}}_{(n)}$ is the aggregated $\bar{\mathbf{T}}$ matrix [46] for n subscatterers which is assumed known at this point. $\psi(k_0, \mathbf{r}_0)$ is a column vector whose elements are the cylindrical wave functions consisting of Hankel functions: $H_m^{(1)}(k_0 \rho) e^{im\phi}$. “ $\Re g$ ” stands for “the regular part of” and hence, $\Re g \psi(k_0, \mathbf{r}_0)$ contains Bessel cylindrical wave functions. In the above, the first term is the incident wave expanded in terms of Bessel functions, or standing cylindrical wave functions. The amplitudes of the Bessel wave functions are given by the elements of the column vector $\mathbf{a}$. The second term in (1) is the scattered field from the n -subscatterers, which emanates from the origin O of the global coordinate system. When the $(n + 1)$ -st subscatterer is added to the aggregated n subscatterers, the total field outside can be written as

$$\phi(\mathbf{r}) = \Re g \psi^i(k_0, \mathbf{r}_0) \cdot \mathbf{a} + \psi^i(k_0, \mathbf{r}_0) \cdot \bar{\mathbf{T}}_{(n+1)} \cdot \mathbf{a}, \quad (2)$$

where $\bar{\mathbf{T}}_{(n+1)}$ is the aggregated $\bar{\mathbf{T}}$ matrix for $(n + 1)$ subscatterers.

The recursion relation for $\bar{\mathbf{T}}_{(n+1)}$ in terms of $\bar{\mathbf{T}}_{(n)}$ has been previously derived [8,26]. The result is

$$\bar{\mathbf{T}}_{(n+1)} = \bar{\mathbf{T}}_{(n)} + [\bar{\boldsymbol{\beta}}_{0,n+1} \bar{\mathbf{T}}_{(n)} \cdot \bar{\boldsymbol{\alpha}}_{0,n+1}] \cdot \bar{\mathbf{T}}_{n+1(n+1)} \cdot \bar{\boldsymbol{\beta}}_{n+1,0}, \quad (3)$$

where

$$\bar{\mathbf{T}}_{n+1(n+1)} \cdot \bar{\boldsymbol{\beta}}_{n+1,0} = [1 - \bar{\mathbf{T}}_{n+1(1)} \cdot \bar{\boldsymbol{\alpha}}_{n+1,0}]^{-1} \cdot \bar{\mathbf{T}}_{n+1(1)} \cdot [\bar{\boldsymbol{\beta}}_{n+1,0} + \bar{\boldsymbol{\alpha}}_{n+1,0} \cdot \bar{\mathbf{T}}_{(n)}], \quad (4)$$

and $\bar{\boldsymbol{\alpha}}$ and $\bar{\boldsymbol{\beta}}$ are appropriate translation matrices [8, p. 464]. Note that (3) and (4) can be combined into one single recursion formula if we so wish, and hence, $\bar{\mathbf{T}}_{n+1(n+1)} \cdot \bar{\boldsymbol{\beta}}_{n+1,0}$ can be regarded as an intermediate unknown at each recursion. This algorithm has been shown to have order N^2 complexity [31]. Moreover, it is exact, and no nearest-neighbor approximation is made.

The reduction in computational complexity can be traced to the fact that the translation matrices are a representation of the translation group [44]. Alternatively, one can view that the reason for the reduction in computational complexity is that the integral equation of scattering has inherent symmetry because the Green’s operator is translationally invariant. The symmetry in the integral equation of scattering yields a special class of matrices obtained by the discretization of such an integral equation (e.g. by MOM) which can be inverted efficiently in $O(N^2)$ operation counts.

The use of addition theorem to construct translation matrices exploits this translational symmetry to reduce redundancy in the inversion of the matrix equation and hence the computational complexity of a scattering problem.

5. Significance of Research

Fast wave-scattering solvers can result in tremendous technological impact in a number of fields. They will reduce the turn-around time in which computer models can be run.

For instance, in remote sensing, these solvers can be used to study scattering from ensemble of particles and to study scintillation phenomenon. These physical models can be used to determine particle size distribution of particles in the remote sensing of the atmosphere. In the past, such problems were thought to be too complex to solve precisely, and hence, they have been addressed with approximate methods.

The speed at which a wave-scattering problem can be solved implies that it can be used to develop imaging algorithms based on physical wave models [45-47]. In oil prospecting, these imaging algorithms will help us "see" the interior of the earth better for better detection of hydrocarbon. In bioengineering and nondestructive testing, they can help in improving imaging with ultrasounds and microwaves accounting for multiple-scattering effect.

Fast wave-scattering algorithms can be used in the computer-aided design and engineering of aircraft with minimum radar cross-section. They can help in the design of antennas, and monolithic microwave integrated circuits. Cost can be reduced in cut-and-try method if the computer model or "experiment" can be performed quickly to ascertain the necessary parameter, shape, or dimension for meeting the specifications of an engineering design.

6. Conclusions

We reviewed a number of fast algorithms for wave scattering developed at the Electromagnetics Laboratory, University of Illinois. This work represents a symbiotic marriage between computer science, applied mathematics, and electromagnetic theory and portends the creation of a new and important technology not previously available. Even though the field of electromagnetic theory itself is matured, the field on the computational theory of electromagnetic field is still very much alive, and in need of quite basic contributions.

Bibliography

- [1] O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation*, Academic Press, New York, 1984.
- [2] P. P. Silvester and R. L. Ferrari, *Finite Elements for Electrical Engineers*, Cambridge University Press, New York, 1983.
- [3] K. Stuben, *Multigrid Methods*, Springer-Verlag, 1985.

- [4] A. George and J. W. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1981.
- [5] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302-307, 1966.
- [6] A. Taflovie and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 623-630, 1975.
- [7] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Stand., Sect. B*, vol. 49, pp. 409-436, 1952.
- [8] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, 1990.
- [9] B. Engquist and A. Majda, "Radiation boundary conditions for acoustic and elastic wave calculations," *Comm. Pure Appl. Math.*, vol. 32, pp. 314-357, 1979.
- [10] E. L. Lindman, "Free space boundary conditions for the time dependent wave equation," *J. Comp. Phys.*, vol. 18, pp. 66-78, 1975.
- [11] Z. P. Liao, H. L. Wong, B. P. Yang, and Y. F. Yuan, "A transmitting boundary for transient wave analysis," *Scientia Sinica, (series A.)*, vol. 27, no. 10, pp. 1063-1076, 1984.
- [12] M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, New York, First Edition, 1959, Sixth Edition, 1980.
- [13] Harrington, R., *Field Computation by Moment Method*. Malabar, FL: Krieger, 1983 (First printing, 1968).
- [14] D. T. Borup and O. P. Gandhi, "Fast-Fourier transform method for calculation of SAR distributions in finely discretized inhomogeneous models of biological bodies," *IEEE Trans. Micro. Theory Tech.*, vol. 32, no. 4, pp. 355-360, 1984.
- [15] T. K. Sarkar, E. Arvas, and S. M. Rao, "Application of FFT and the conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 635-639, 1986.
- [16] M. F. Cátedra, "Solution to some electromagnetic problems using fast Fourier transform with conjugate gradient method," *Electron. Lett.*, vol. 22, pp. 1049-1051, Sept. 25, 1986.
- [17] D. T. Borup, D. M. Sullivan, and O. P. Gandhi, "Comparison of two FFT conjugate gradient method and the finite-difference time-domain method for the 2-D absorption problem," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, no. 4, pp. 383-395, April, 1987.
- [18] C. C. Su, "Calculation of electromagnetic scattering from a dielectric cylinder using the conjugate gradient method and FFT," *IEEE Trans. Antennas Propag.*, vol. AP-35, no. 12, pp. 1418-1425, December, 1987.
- [19] C. Y. Shen, K. J. Glover, M. I. Sancer, and A. D. Varvatsis, "The dis-

- crete Fourier transform method of solving differential-integral equations in scattering theory," *IEEE Trans. Ant. Propag.*, vol. AP-37, no. 8, pp. 1032-1041, August, 1989.
- [20] A. J. Poggio and E. K. Miller, "Integral equation solution of three-dimensional scattering problems," *Computer Techniques for Electromagnetics*, ed. R. Mittra, Pergamon Press, New York, 1973.
 - [21] A. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antennas Propagat.*, vol. AP-28, no. 5, pp. 593-506, 1980.
 - [22] V. Rokhlin, "Rapid solution of integral equations of scattering theory in two dimensions," *J. Comput. Phys.*, vol. 86, pp. 414-439, 1990.
 - [23] W. C. Chew, "An N^2 algorithm for the multiple scattering solution of N scatterers," *Micro. Optical Tech. Letter*, vol. 2, no. 11, pp. 380-383, 1989.
 - [24] W. C. Chew, J. Friedrich, and R. Geiger, "A multiple scattering solution for the effective permittivity of a sphere mixture," *IEEE Geoscience and Remote Sensing*, vol. 28, no. 2, p. 207, 1990.
 - [25] Y. M. Wang and W. C. Chew, "An efficient algorithm for solution of a scattering problem," *Micro. Opt. Tech. Lett.*, vol. 3, no. 3, pp. 102-106, March, 1990.
 - [26] W. C. Chew and Y. M. Wang, "A fast algorithm for solution of a scattering problem using a recursive aggregate tau matrix method," *Microwave Optical Technology Letters*, vol. 3, no. 5, pp. 164-169, May 1990.
 - [27] Y. M. Wang and W. C. Chew, "Solution of large 2-D inhomogeneous scatterers for TM waves by fast recursive algorithm," *Micro. Opt. Tech. Lett.*, vol. 4, no. 3, March 1991.
 - [28] J. H. Lin and W. C. Chew, "Solution of oblong, dielectric coated, metallic scatterers by the recursive algorithm," *Electron. Lett.*, vol. 28, no. 2, pp. 185-186, January 1992.
 - [29] L. Gurel and W. C. Chew, "A recursive T-matrix algorithm for strips and patches," *Radio Science*, vol. 27, no. 3, pp. 387-401, 1992.
 - [30] L. Gurel and W. C. Chew, "Scattering solution of three-dimensional array of patches using the recursive T-matrix algorithms," *Micro. Guided Wave Lett.*, vol. 2, no. 5, 1992.
 - [31] W. C. Chew, L. Gurel, Y. M. Wang, G. Otto, R. Wagner, and Q. H. Liu, "A generalized recursive algorithm for wave-scattering solutions in two dimensions," *IEEE Trans. Micro. Theory Tech.*, vol. 40, no. 4, pp. 716-723, April 1992.
 - [32] W. C. Chew, Y. M. Wang, and L. Gurel, "A recursive algorithm for wave-scattering using windowed addition theorem," *J. Electromag. Waves Appl.*, accepted for publication.
 - [33] W. C. Chew, "A review of wave-scattering solution techniques," *Asia-Pacific Engg. J.*, vol. 1, no. 2, pp. 125-150, December 1991.
 - [34] W. C. Chew and C. C. Lu, "A recursive algorithm to compute the wave-scattering solution of a finite-strip array using an efficient plane-wave

- basis," *Micro. Opt. Tech. Lett.*, vol. 5, no. 3, pp.146-148, 1992.
- [35] C. C. Lu and W. C. Chew, "Electromagnetic scattering of finite strip array on a dielectric slab," *IEEE Trans. Micro. Theory Tech.*, accepted for publication.
 - [36] W. C. Chew and C. C. Lu, "NEPAL—An $N^{1.5}$ algorithm for solving volume integral equation," *IEEE Ant. Propagat. Soc. Int. Sym. Dig.*, vol 1, pp. 184-187, July 1992.
 - [37] W. C. Chew and B. Anderson, "Propagation of electromagnetic waves through geological beds in a geophysical probing environment," *Radio Science*, vol. 20, no. 3, pp. 611-621, May-June 1985.
 - [38] W. C. Chew, Z. Nie, Q. Liu, and B. Anderson, "An efficient solution of well logging tools in a complex environment," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, no. 2, pp. 308-313, 1991.
 - [39] W. C. Chew and L. Gurel, "Reflection and transmission operators for strips or disks embedded in homogeneous and layered media," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, no. 11, pp. 1488-1497, 1988.
 - [40] L. Gurel and W. C. Chew, "Guidance and resonance conditions for strips or disks embedded in homogeneous and layered media," *IEEE Trans. Microwave Theory Tech.*, vol. 36, no. 11, pp. 1498-1506, 1988.
 - [41] P. C. Waterman, "New formulation of acoustic scattering," *J. Acoust. Soc. Am.*, Vol. 45, pp. 1417-1429, 1969.
 - [42] N. J. Vilenkin, *Special Functions and the Theory of Group Representations*, American Mathematical Society, Providence, Rhode Island, 1968.
 - [43] Y. M. Wang and W. C. Chew, "An iterative solution of the two-dimensional electromagnetic inverse scattering problem," *International J. Imaging Systems Tech.*, vol. 1, pp.100-108, 1989.
 - [44] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Medical Imag.*, vol. 9, no. 2, pp 218-225, 1990.
 - [45] D. T. Borup, S. A. Johnson, W. W. Kim, and M. J. Berggren, "Nonperturbative diffraction tomography via Gauss-Newton iteration applied to the scattering integral equation," *Ultrasonic Imaging*, vol. 14, pp. 69-85, 1992.

ON THE PROBLEM OF SYNTHESIS OF OFFSET DUAL SHAPED REFLECTOR ANTENNAS

Vladimir Oliker

Department of Mathematics and Computer Science
Emory University
Atlanta, Georgia 30322
and
Matis, Inc.
120 Parkwood Lane, Decatur, Georgia 30030

Abstract

The geometric optics approximation may be used in design of offset single and dual nonaxially symmetric reflector antennas when it is required to control the field amplitude and/or phase on the far field or on the output aperture. In principle, the formulation of the problem is straight forward and consists of two steps: (a) using Snell's law one derives the ray tracing equations and (b) using the energy conservation law for the energy flow along each differential tube of rays one derives the relation between the primary feed and output power patterns. These relations are supplemented by boundary or initial like conditions. In describing step (b), analytically, one needs to compute the Jacobian of the map established by the ray tracing equations. This can be done in different ways which, while expressing the same physical principles, lead to different differential equations. Then, when solving the problem numerically, one may get different results with different sources of errors. There have been, essentially, two approaches to the problem in case of nonradially symmetric and noncoaxial reflectors. The approach of B. Kinber in the former USSR and V. Galindo-Israel et al. in USA is based on formulating the problem as a nonlinear system of first order partial differential equations describing the ray tracing equations and the Jacobian. Another approach, due to F. Brickell, B. Westcott, and their collaborators in England, is based on formulating the problem as one second order nonlinear partial differential equation. In this paper, after a brief discussion of these two techniques, we present an alternative approach that allows to avoid some of the theoretical and computational difficulties present in the first two approaches.

1 Introduction

The problem of synthesis of offset single and dual reflector antennas using geometric optics approximation has been a subject of investigation by many authors [1], [2], [3], [4], [5], [6], [7], [8], [9]. The objectives of design are to build high efficiency antennas for which it is possible to control the field amplitude and/or phase on the output aperture, and, at the same time minimize energy losses due to spillovers, reduce blockage, achieve low sidelobes, etc.; see [10].

In this paper we discuss only systems with two reflectors. However, much of that discussion applies also to single reflector systems; see [11]. In principle, the formulation of the synthesis problem is straight forward and consists of two steps: (a) using Snell's law one derives the ray tracing equations and (b) using the energy conservation law one derives the relation between the primary feed and output power pattern. These relations are supplemented by boundary conditions relating the boundaries of the input pupil and output aperture. In step (b) one needs to compute the Jacobian of the map between the input pupil and output aperture. This can be done in different ways, which, while expressing the same physical principles, lead to different differential equations. Then, when solving the problem numerically, one may get different results with different sources of errors.

The case of coaxial and rotationally symmetric antennas reduces to a boundary value problem in ordinary differential equations and is well understood; see, for example, [10]. In case of offset and not rotationally symmetric antennas, there have been, essentially, two approaches to the problem. The approach of B. Kinber [1], [2] in the former USSR and V. Galindo-Israel et

al. [3], [4], [5] (see also other references there) in USA is based on formulating the problem as a nonlinear system of first order partial differential equations (PDE's) describing (a) and (b). This system is solvable only if certain integrability conditions are satisfied. These conditions may not always hold and in such circumstances the problem does not admit a solution. Nevertheless, as far as numerical computations are concerned, one still may proceed and construct a numerical surrogate "solution". In general, this solution will not satisfy the energy equation and/or boundary conditions and the errors could be substantial [3], [2], p.24. Strictly speaking, if the integrability conditions are not satisfied one has to allow solutions that are multivalued function, - a notion, physically, not well motivated in this context.

Another approach, due to F. Brickell, B. Westcott and their collaborators in England [6], [7], [8], is based on formulating the problem in complex form as one second order nonlinear PDE for a complex valued function (like a potential) defined on the output aperture. The subreflector and the main reflector are recovered via a formula involving that potential and its first derivatives. In this approach the corresponding boundary value problem is formulated on the output aperture. This adds certain "rigidity" to the problem because if it is necessary to consider and test different output apertures with the same input pupil then the problem has to be solved anew each time on a new domain with a new grid and new boundary condition. This becomes especially involved in case of elliptic apertures [8]. For antennas with large diameter this may result in a substantial additional computational effort.

Unfortunately, in either approach theoretical results concerning existence and uniqueness of solutions to the problem are still lacking. Some progress in

this direction in case of the second approach and single reflectors was made in [12], [13], and [11]. However, many questions, computational and theoretical, still remain open. Evidently, if one hopes to have a reliable and user friendly computer code solving the synthesis problem, then more theoretical and numerical work needs to be done.

The purpose of this paper is to describe an alternative approach to formulating the synthesis problem analytically and solving it numerically. In this approach we derive a new second order PDE describing simultaneously the steps (a) and (b). In contrast with the approach of F. Brickell and B. Westcott this equation is obtained in real valued form and it is an equation directly for the polar radius of the subreflector. In important practical cases this information is sufficient for recovering the coordinates of the main reflector. Another novel feature is that the equation and the boundary condition are defined on the input pupil and the shape of the output aperture may be used as a control parameter. Its adjustment introduces only a change in the boundary condition, while the domain on which the problem is solved remains the same. It should be noted that in case of single reflectors the framework for deriving such an equation was formulated by J. S. Schruben [14].

The equation that we obtain is also nonlinear and currently we are investigating several algorithms for solving it numerically. The numerical algorithm presented here is based on the classical Newton type scheme. Several tests that we have ran so far, show fast convergence and allow easy handling of elliptic input and output pupils. Our actual implementation is a modification of the well known elliptic PDE's solver, ELLPACK [15].

The rest of the paper is organized as follows. In section 2 we present the

new equation relating the input and output power densities and the boundary condition. In section 3 we discuss the properties of the corresponding boundary value problem and in section 4 we outline the algorithm for solving the synthesis problem numerically.

2 The equations of the problem

On figures 1 and 2 we schematically represent a dual offset antenna system of Cassegrain type. The input pupil $\bar{\Omega}$ is specified and is described by a vector function $\mathbf{m}(u^1, u^2)$ in some coordinates u^1 and u^2 on the unit sphere S . Usually, $u^1 = \theta$, $u^2 = \phi$ are the spherical coordinates on S , and in these coordinates $\mathbf{m}(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$. However, spherical coordinates are not always convenient and occasionally we need to describe $\bar{\Omega}$ in some other coordinates. The bar over Ω indicates that the boundary $\partial\Omega$ of the region Ω is included; thus, $\bar{\Omega} = \Omega \cup \partial\Omega$. Our other notation is as follows:

The input power pattern $I(\mathbf{m})$ is specified for the point source O as a function of the input direction $\mathbf{m} \in \Omega$. The position vector of the subreflector, that is to be determined, is denoted by the vector function $\mathbf{r}(\mathbf{m}) = \rho(\mathbf{m})\mathbf{m}$, $\mathbf{m} \in \bar{\Omega}$, where ρ is the polar radius of F . The normal vector $\mathbf{n}$ on F can be described in terms of ρ as

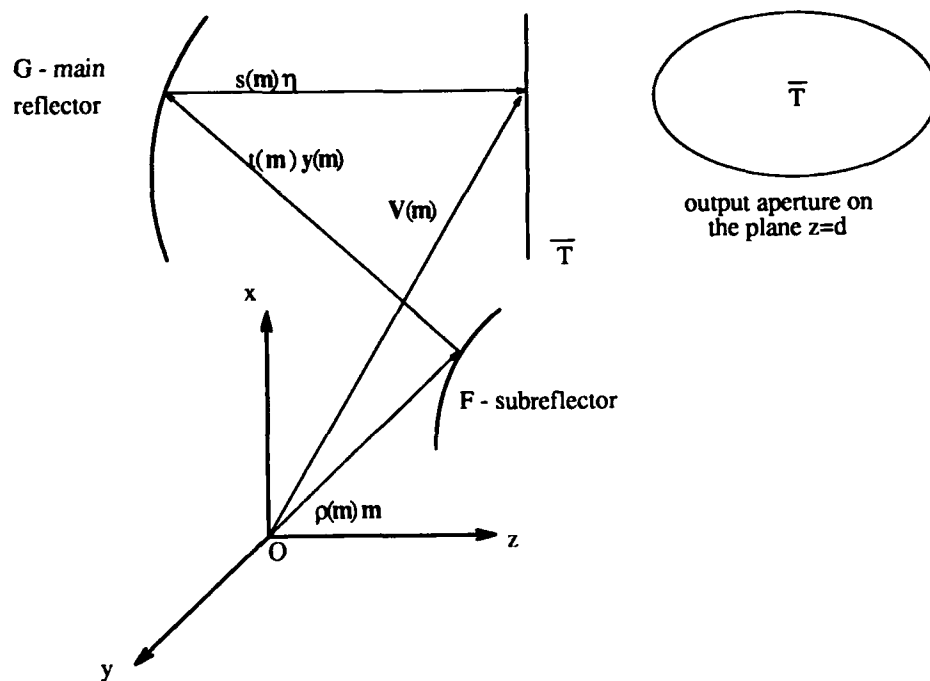


Figure 1:

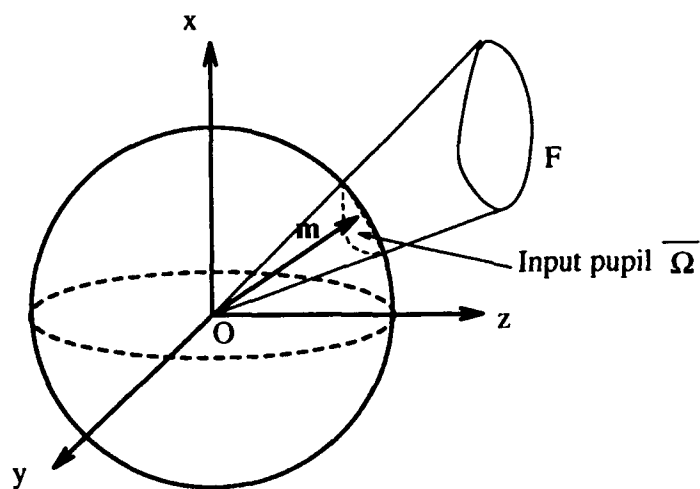


Figure 2:

$$\mathbf{n} = \frac{1}{\sqrt{\rho^2 + |\nabla\rho|^2}}(\rho\mathbf{m} - \nabla\rho), \quad (1)$$

where

$$\nabla\rho = \sum_{i,j=1}^2 e^{ij} \frac{\partial\rho}{\partial u^i} \frac{\partial\mathbf{m}}{\partial u^j},$$

with (e^{ij}) being the inverse matrix of the matrix

$$e_{ij} = \frac{\partial\mathbf{m}}{\partial u^i} \cdot \frac{\partial\mathbf{m}}{\partial u^j}, \quad i, j = 1, 2;$$

$\cdot$ denotes the scalar product, and

$$|\nabla\rho|^2 = \nabla\rho \cdot \nabla\rho.$$

The matrix (e_{ij}) is the matrix of coefficients of the first fundamental form of the sphere S (see [16], p. 68).

The unit vector $\mathbf{y}(\mathbf{m})$ in the direction of the ray reflected off F for the incidence direction $\mathbf{m}$ is determined by Snell's law as

$$\mathbf{y}(\mathbf{m}) = \mathbf{m} - 2(\mathbf{m} \cdot \mathbf{n}(\mathbf{m}))\mathbf{n}(\mathbf{m}). \quad (2)$$

It also can be expressed in terms of ρ and its derivatives if we substitute (1) into (2).

We denote by $t(\mathbf{m})$ the function giving the distance from F to the main reflector G along the direction $\mathbf{y}(\mathbf{m})$ (see figure 1).

The output aperture $\bar{T}$ is specified as a region on the plane $z = d$; its points are described as (x, y, d) . In the problem under consideration $\bar{T}$ is the reflected wave front, that is, it is a (plane) surface of constant phase.

We denote by $\boldsymbol{\eta}$ the unit vector in the direction of reflection off G ; in our case $\boldsymbol{\eta}$ is a constant vector in the positive direction of the z -axis.

We denote by $L(x, y)$, $(x, y) \in T$, the prespecified output power pattern on T .

The total optical path length is denoted by l ; thus,

$$l(\mathbf{m}) = \rho(\mathbf{m}) + t(\mathbf{m}) + s(\mathbf{m}), \quad (3)$$

where $s(\mathbf{m})$ is the distance from G to $\bar{T}$ in the direction $\boldsymbol{\eta}$. The fact that on $\bar{T}$ the phase is constant implies that

$$l(\mathbf{m}) = \text{const} \equiv l. \quad (4)$$

It is assumed that l is specified.

Finally, the vector function

$$\mathbf{V}(\mathbf{m}) = \rho(\mathbf{m})\mathbf{m} + t(\mathbf{m})\mathbf{y}(\mathbf{m}) + (l - \rho(\mathbf{m}) - t(\mathbf{m}))\boldsymbol{\eta}, \quad \mathbf{m} \in \bar{\Omega}, \quad (5)$$

gives a representation of $\bar{T}$ after the two reflections.

The fact that the output aperture $\bar{T}$ and the total path length are specified allow to express $t(\mathbf{m})$ entirely in terms of ρ and input data. Namely, projecting $\mathbf{V}(\mathbf{m})$ on $\boldsymbol{\eta}$, resolving for $t(\mathbf{m})$ and noting that $\mathbf{V}(\mathbf{m}) \cdot \boldsymbol{\eta} = d$, we obtain

$$t(\mathbf{m}) = \frac{l - d + \rho(\mathbf{m})(\mathbf{m} \cdot \boldsymbol{\eta} - 1)}{1 - \mathbf{y} \cdot \boldsymbol{\eta}}. \quad (6)$$

The energy conservation law implies that we must have

$$\int_{\Omega} I(\mathbf{m}) d\sigma = C \int_T L(x, y) dx dy \quad (7)$$

for some positive constant C . Here $d\sigma$ is the area element on S .

The equation relating the input and output power patterns pointwise was derived in [17]. We will not reproduce here its derivation. The main equation is as follows:

$$CL(\mathbf{V}(\mathbf{m})) \frac{\det(M_{ij})}{\det(e_{ij}(\mathbf{m}))} = I(\mathbf{m}), \quad \mathbf{m} \in \Omega, \quad (8)$$

where

$$M_{ij} = \frac{2t(\mathbf{m})}{\rho(\mathbf{m})} b_{ij}(\mathbf{m}) (\mathbf{m} \cdot \mathbf{n}(\mathbf{m}) + (\rho(\mathbf{m}) + t(\mathbf{m})) e_{ij}(\mathbf{m})), \quad (9)$$

$$b_{ij}(\mathbf{m}) = \left(\frac{\partial^2 \mathbf{r}(\mathbf{m})}{\partial u^i \partial u^j} \cdot \mathbf{n}(\mathbf{m}) \right), \quad i, j = 1, 2. \quad (10)$$

The matrix (b_{ij}) is the matrix of coefficients of the second fundamental form (see [16], p. 78) of the subreflector F . These coefficients can be expressed completely in terms of ρ : (see [18]):

$$b_{ij} = \frac{\rho \nabla_{ij} \rho - 2\rho_i \rho_j - \rho^2 e_{ij}}{\sqrt{\rho^2 + |\nabla \rho|^2}}, \quad (11)$$

where

$$\rho_i = \frac{\partial \rho}{\partial u^i}, \quad \nabla_{ij} \rho = \frac{\partial^2 \rho}{\partial u^i \partial u^j} - \sum_{k=1}^2 \Gamma_{ij}^k \rho_k,$$

with Γ_{ij}^k 's being the Christoffel symbols of the second kind of (e_{ij}) (see [16], p. 127).

Now we describe the boundary condition. It is required that the boundary $\partial\Omega$ of the input pupil $\bar{\Omega}$ be mapped by $\mathbf{V}(\mathbf{m})$ (see (5)) onto the boundary ∂T of the output aperture $\bar{T}$. In order to express this condition analytically, assume that ∂T is given on the plane $z = d$ by the equation

$$\Phi(x, y) = 0, \quad (12)$$

where Φ is a given function. Put

$$\mathbf{V}(\mathbf{m}) = V_x(\mathbf{m})\mathbf{i} + V_y(\mathbf{m})\mathbf{j} + V_z(\mathbf{m})\mathbf{k},$$

where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors in directions of coordinate axes. Then the above boundary condition analytically is equivalent to the requirement that

$$\Phi(V_x(\mathbf{m}), V_y(\mathbf{m})) = 0, \quad \mathbf{m} \in \partial\Omega. \quad (13)$$

This is a restriction on the V_x and V_y components of V . There is no need to impose a restriction on the component V_z , because by (5) such a restriction would be

$$\rho(\mathbf{m})(1 - \mathbf{m} \cdot \boldsymbol{\eta}) + t(\mathbf{m})(1 - \mathbf{y}(\mathbf{m}) \cdot \boldsymbol{\eta}) = l - d, \quad \mathbf{m} \in \partial\Omega,$$

and it is already taken into account by means of defining t in (6) for all points in $\bar{\Omega}$.

If the equation (8) is solved for ρ then $\rho(\mathbf{m})\mathbf{m}$ gives the coordinates of the subreflector. Using (1) and (6) we construct $\mathbf{y}$ and t and then the vector function $\rho(\mathbf{m})\mathbf{m} + t(\mathbf{m})\mathbf{y}(\mathbf{m})$ gives the coordinates of the main reflector. Thus, the equation (8) with t as in (6) and boundary condition (13) constitute the analytic form of the synthesis problem for determining the reflectors F and G .

3 Remarks on equation (8)

Put

$$D(\rho) = \frac{\det(M_{ij})}{\det(e_{ij})}, \quad (14)$$

where (M_{ij}) and (e_{ij}) are as in (8). It was shown in [17] that

$$D(\rho) = |\text{Jacobian of } \mathbf{V}|,$$

where $\mathbf{V}$ is as in (5) and considered as a map from $\bar{\Omega} \rightarrow \bar{T}$. It is clear from the expressions for M_{ij} and (11) that $D(\rho)$ is a second order differential operator in ρ . This operator is positively (negatively) elliptic on a given ρ if the quadratic form

$$\sum_{i,j=1}^2 \frac{M^{ij}}{\det(e_{ij})} \xi_i \xi_j, \quad (15)$$

is positive (negative) definite for any $\xi = (\xi_1, \xi_2)$ such that $|\xi| \neq 0$. Here (M^{ij}) denotes the cofactor matrix of the matrix (M_{ij}) . In order to explain the physical meaning of the difference between positive and negative ellipticity we consider two special cases. In the classical offset Cassegrain (figure 3) and Gregorian (figure 4) geometries the subreflector F is either one sheet of a two sheeted hyperboloid or an ellipsoid.

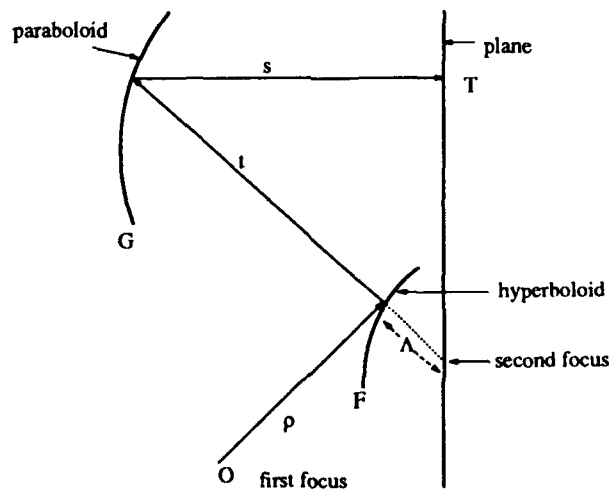


Figure 3: Cassegrain geometry

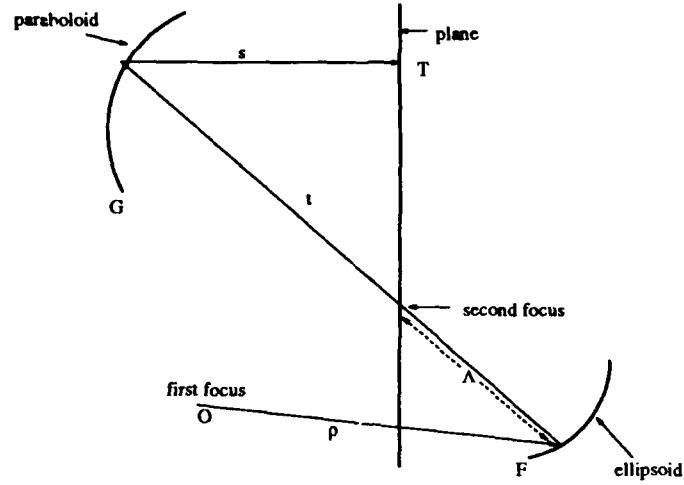


Figure 4: Gregorian geometry

It follows from a computation in [19], section 3.1, and formula (9) that for the case when F is a hyperboloid or ellipsoid we have

$$M_{ij} = \rho(\pm \frac{t}{\Lambda} + 1)e_{ij}, \quad (16)$$

where “+” should be taken in case of hyperboloid and “-” should be taken in case of ellipsoid; here Λ is the distance from the point of reflection on F to the second focus. Consequently, (15) becomes

$$\sum_{i,j=1}^2 M^{ij} \xi_i \xi_j = \rho(\pm \frac{t}{\Lambda} + 1) \sum_{i,j=1}^2 e^{ij} \xi_i \xi_j. \quad (17)$$

The expression $\sum_{i,j=1}^2 e^{ij} \xi_i \xi_j > 0$ for any $\xi \neq 0$ (see [16], p. 70) and, therefore, the operator D is positively elliptic for the Cassegrain geometry and negatively elliptic for the Gregorian geometry provided $t > \Lambda$. Note also that in these two cases

$$D(\rho) = \rho^2(\pm \frac{t}{\Lambda} + 1)^2. \quad (18)$$

4 Numerical algorithm and simulations

For the convenience of the reader we summarize here the equations and relations for the dual synthesis problem. All the notations are the same as in previous sections.

We need to solve for the unknown function ρ the equation

$$CL(\mathbf{V}(\mathbf{m})) \frac{\det(M_{ij})}{\det(e_{ij}(\mathbf{m}))} = I(\mathbf{m}), \quad \mathbf{m} \in \Omega, \quad (19)$$

subject to the conditions:

$$t(\mathbf{m}) = \frac{l - d + \rho(\mathbf{m})(\mathbf{m} \cdot \boldsymbol{\eta} - 1)}{1 - \mathbf{y}(\mathbf{m}) \cdot \boldsymbol{\eta}}, \quad \mathbf{m} \in \bar{\Omega}, \quad (20)$$

$$\Phi(V_x(\mathbf{m}), V_y(\mathbf{m})) = 0, \quad \mathbf{m} \in \partial\Omega. \quad (21)$$

If the domains Ω , T , and functions I and L are arbitrary the problem (19)-(21), in general, may not have a solution. Fortunately, in practical applications, as in [3]-[5] and [6]-[8], the problem has to be solved in cases when a good initial approximation is available for the probable solution. The discussion below is pertaining only to such cases. Since the synthesis equation (19) was found very recently, we have, so far, only preliminary results regarding numerical solutions of (19)-(21).

1. We begin by considering the equation (19) on the classical rotationally symmetric Cassegrain paraboloid/hyperboloid system. The polar equation of the hyperboloid is given by

$$\rho = \frac{f}{1 - \epsilon(\mathbf{m} \cdot \hat{\mathbf{k}})}, \quad \mathbf{m} \in \bar{\Omega}, \quad (22)$$

where f is the latus rectum, ϵ the eccentricity, and $\hat{\mathbf{k}}$ the unit vector in the direction of the axis of revolution of the hyperboloid. In the first example below we take the following data:

$\bar{\Omega}$ is a disk on the sphere S with center $(0,0,1)$ and angular radius 14° ; $\epsilon = 2$; $f = -1$; the optical path length $l = 10$; the distance to the aperture plane $d = 4$. The output aperture disk $\bar{T}$ lies on the plane $z = d$ and is centered at $(0,0,4)$. Using (22) and (6), we construct $W(\mathbf{m}) = \rho(\mathbf{m})\mathbf{m} + t(\mathbf{m})\mathbf{y}(\mathbf{m})$ and determine the radius A of the disk $\bar{T}$ by computing the length of the vector $W^*(\mathbf{m}) = W_x\mathbf{i} + W_y\mathbf{j}$, where W_x and W_y are the x and y components of the vector W . With 6 significant digits we determined that $A = 2.45569$. Next, we use (19) and (18) to compute $D(\rho(\mathbf{m}))$ in Ω . The computed $D(\rho(\mathbf{m}))$ is taken as $I(\mathbf{m})$ in (19). The output power pattern $L(\mathbf{V}(\mathbf{m}))$ is required in this case to be uniform and identically equal to 1. Consequently, the constant C in (19) is also equal to 1. The condition (21) assumes the form:

$$\begin{aligned} [\rho(\mathbf{m})\mathbf{m} + t(\mathbf{m})\mathbf{y}(\mathbf{m})]_x^2 + [\rho(\mathbf{m})\mathbf{m} + t(\mathbf{m})\mathbf{y}(\mathbf{m})]_y^2 &= A^2 \\ \text{when } (\mathbf{m} \cdot (0,0,1)) &= \cos 14^\circ. \end{aligned} \quad (23)$$

The problem (19)-(21) in this case is solved for radially symmetric solution $\hat{\rho}(\mathbf{m})$, $\mathbf{m} \in \bar{\Omega}$. We solve it with the standard routine IVPAG from the IMSL package. The program IVPAG solves an initial value problem for systems of ordinary differential equations. In order to apply this program to our problem we set up the following initial data: at $\mathbf{m}_0 = (0,0,\cos \theta_0)$, with $\theta_0 = 1.0 \times 10^{-5}$ grad $\hat{\rho} = 0$ and $\hat{\rho}(\mathbf{m}_0) = 1$. The initial conditions are set up slightly off the origin $(0,0,1)$ because the equation (19) has a singularity at this point due to the use of spherical coordinates. The table below shows the maximal differences between the computed values of $\hat{\rho}$, $\hat{t}$, $\hat{L}$, and $\hat{A}$ and the corresponding

true values (without the over hat).

$$\begin{array}{cccc} \max |\hat{\rho} - \rho| & \max |\hat{t} - t| & \max |\hat{L} - 1| & |\hat{A} - A| \\ 2.80155 \times 10^{-10} & 2.06280 \times 10^{-9} & 0 & 5.68 \times 10^{-9} \end{array}$$

In the next example we solved (19)-(21) for a radially symmetric solution with the following data: $\bar{\Omega}$, $\bar{T}$, l , and d are as before and $I(\mathbf{m}) \equiv I(\theta) = \cos^{229} \theta$, $0 \leq \theta \leq 14^\circ$. This corresponds to $-30dB$ taper. Again the objective was to achieve uniform ($L \equiv 1$) energy density distribution on T . The balancing constant C was determined from the equation (7);

$$\begin{array}{ccc} \text{Balancing constant } C & |\hat{A} - A| & |\hat{L}(\hat{\mathbf{V}}) - 1| \\ .00144 & 1.5 \times 10^{-8} & 3.10862 \times 10^{-15} \end{array}$$

2. Next we consider the problem (19)-(21) in the nonrotationally symmetric case. We used a linearization procedure based on the Newton scheme [15]. The latter requires knowledge of the Frechet derivatives of the operators in (19)-(21). The corresponding computation is quite involved. First we rewrite the equation (19) (using (9) and (11)) as follows. Put

$$M_{ij} = tE_{ij} + \rho e_{ij}, \quad (24)$$

where

$$E_{ij} = 2 \frac{b_{ij}(\mathbf{m} \cdot \mathbf{n})}{\rho} + e_{ij}. \quad (25)$$

We make a change of the unknown function by setting $\rho = e^{-P}$. Then

$$E_{ij} = - \frac{\nabla_{ij} P + \partial_i P \partial_j P + (1 - R)e_{ij}}{R}, \quad (26)$$

where

$$R = \frac{1}{2}(1 + |\nabla P|^2), \quad \partial_i = \frac{\partial}{\partial u^i}.$$

Put

$$Q(P) \equiv CL(P) \frac{\det(M_{ij}(P))}{\det(e_{ij})} - I. \quad (27)$$

Let h be an arbitrary smooth function defined in $\bar{\Omega}$ and λ a small real number. The Frechet derivative of $Q(P)$ is a second order differential expression. We denote this derivative symbolically as $Q'(P)h$. It is computed as follows.

$$\begin{aligned} Q'(P)h &\equiv \left. \frac{d}{d\lambda} Q(P + \lambda h) \right|_{\lambda=0} \\ &= C \left. \frac{dL}{d\lambda} \frac{\det(M_{ij}(P))}{\det(e_{ij})} \right|_{\lambda=0} + CL(P) \sum_{i,j=1}^2 \frac{M^{ij}(P)}{\det(e_{ij})} \left. \frac{dM_{ij}}{d\lambda} \right|_{\lambda=0}. \end{aligned} \quad (28)$$

where (M^{ij}) is the cofactor matrix of (M_{ij}) . In order to get an explicit expression for $Q'(P)h$ we need additional notation. Put

$$\begin{aligned} a &= l - d, \quad b = \mathbf{m} \cdot \boldsymbol{\eta} - 1, \quad \mathbf{c} = \mathbf{m} + \nabla P, \\ g &= 1 - \mathbf{y} \cdot \boldsymbol{\eta}, \\ \mathbf{F}^i &= \frac{1}{R} \sum_{j=1}^2 e^{ij} \left(\frac{\mathbf{c}}{R} \frac{\partial P}{\partial u^j} - \frac{\partial \mathbf{m}}{\partial u^j} \right), \quad i = 1, 2, \\ F^i &= \mathbf{F}^i \cdot \boldsymbol{\eta}, \quad i = 1, 2, \\ P^i &= \sum_{j=1}^2 e^{ij} \frac{\partial P}{\partial u^j}, \\ \gamma_{ij}^k &= (E_{ij} - e_{ij})P^k + 2 \frac{\partial P}{\partial u^i} \delta_j^k - \Gamma_{ij}^k, \quad i, j, k = 1, 2, \end{aligned}$$

where $\delta_j^k = 0$ if $k \neq j$ and $\delta_j^k = 1$ if $k = j$.

With this notation we have:

$$\left. \frac{dE_{ij}}{d\lambda} \right|_{\lambda=0} = -\frac{1}{R} \frac{\partial^2 h}{\partial u^i \partial u^j} - \frac{1}{R} \sum_{k=1}^2 \gamma_{ij}^k \frac{\partial h}{\partial u^k}, \quad (29)$$

$$\left. \frac{dt}{d\lambda} \right|_{\lambda=0} = \frac{t}{g} \sum_{i=1}^2 F^i \frac{\partial h}{\partial u^i} - \frac{b}{g} e^{-P} h, \quad (30)$$

$$\left. \frac{dM_{ij}}{d\lambda} \right|_{\lambda=0} = \left. \frac{dt}{d\lambda} \right|_{\lambda=0} E_{ij} - e^{-P} e_{ij} h. \quad (31)$$

Substituting (29)-(31) into (28) we obtain the expression for $Q'(P)h$.

Next we compute the derivative of the boundary operator in (21). The left hand side of (21) we write as $\Phi(P)$ and compute the Frechet derivative $\Phi'(P)h$ as

$$\Phi'(P)h = \left. \frac{d\Phi(P + \lambda h)}{d\lambda} \right|_{\lambda=0} \quad (32)$$

So far the particular form of the function Φ has not been specified. In applications the important cases are the ones when $\Phi(x, y) = 0$ is a circle or an ellipse on the plane $z = d$. Let

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} = 1 \quad (33)$$

be an ellipse with center at (x_0, y_0, d) and semiaxes α and β . As before, let

$$\mathbf{W}(P) = e^{-P}\mathbf{m} + t(P)\mathbf{y}(P), \quad (34)$$

and let $\mathbf{X}_0 = (x_0, y_0, d)$ be the center of the required output aperture ellipse. Then the boundary condition (21) can be written as

$$\beta^2(\mathbf{W}(P) - \mathbf{X}_0)_x^2 + \alpha^2(\mathbf{W}(P) - \mathbf{X}_0)_y^2 - \alpha^2\beta^2 = 0, \quad (35)$$

where $()_x$ and $()_y$ denote the x - and y - components of corresponding vectors.

Put

$$\begin{aligned} \mathbf{H}^0 &= e^{-P}(\mathbf{m} + \frac{b}{g}\mathbf{y}), \\ \mathbf{H}^i &= t\left(\frac{F^i}{g}\mathbf{y} + \mathbf{F}^i\right), \quad i = 1, 2, \\ B^k &= 2\beta^2(\mathbf{W} - \mathbf{X}_0)_x H_x^k + 2\alpha^2(\mathbf{W} - \mathbf{X}_0)_y H_y^k, \quad k = 0, 1, 2. \end{aligned}$$

Then

$$\Phi'(P)h = \sum_{i=1}^2 B^i \frac{\partial h}{\partial u^i} - B^0 h. \quad (36)$$

In order to solve the problem (19)-(21) we start with an initial guess P_0 and compute the increments h_n by solving for $n = 0, 1, 2, \dots$ the linear problems

$$Q'(P_n)h_{n+1} = -Q(P_n) \text{ in } \Omega, \quad (37)$$

$$\Phi'(P_n)h_{n+1} = -\Phi(P_n) \text{ on } \partial\Omega. \quad (38)$$

The update is found by setting

$$P_{n+1} = P_n + h_{n+1}. \quad (39)$$

It can be shown that after a normalization of the function ρ in (19)-(21) the homogeneous problem corresponding to (37)-(38) with right hand sides replaced by zeroes admits $h = \text{const}$ as a solution. Consequently, an additional condition should be required of $Q(P_n)$; this condition is equivalent to the requirement that the balance equation be satisfied for all P_n . The increments are determined so that at some fixed point all h_n are the same. A theoretical analysis of this situation will be presented elsewhere.

Numerically, the above scheme was implemented using a linear PDE solver ELLPACK [15]. The Newton scheme suggested in [15], p. 101, had to be modified to eliminate computation of $Q'(P_n)P_n$. The operators $Q(P)$ and $Q'(P)$ are sensitive to small errors in second derivatives (due to quadratic nonlinearities in second derivatives in the operator Q) and additional errors are introduced if the term $Q'(P_n)P_n$ is included (as required by the scheme suggested on p. 102 of [15]). In addition, in ELLPACK the coefficients of $Q'(P_n)$ and $Q(P_n)$ are evaluated using bicubic polynomial approximation to P . Again, because of the strong nonlinearities, one may need to use an interpolant of a higher degree. Currently, we are implementing an algorithm that incorporates the

needed modifications. It should be noted that the difficulty can be overcome by introducing into (37), (38) parameters $\mu_n, \nu_n, 0 < \mu_n \leq 1, 0 < \nu_n \leq 1$,

$$Q'(P_n)h_{n+1} = -\mu_n Q(P_n), \quad (40)$$

$$\Phi'(P_n)h_{n+1} = -\nu_n \Phi(P_n). \quad (41)$$

These parameters also help to control the errors if the initial guess is not well chosen. However, when the parameters are small the convergence may be slow. Note also, that the scheme (40), (41) can be interpreted as a nonstationary process with $\mu_n = \nu_n = \text{time step}$.

Within the above limitations the scheme produced reasonably good results. As an example for testing the feasibility of the algorithm and numerical scheme we present the results of calculations for the data in [7], pp. 108-109. A dual offset paraboloid/hyperboloid system is considered with the following parameters: the hyperboloid (subreflector) is given by (22) with its axis tilted at 10° ; eccentricity = 2.26, $f = -.2$; the input pupil $\bar{\Omega}$ is a disk centered at the point $(\sin 35.5^\circ, 0, \cos 35.5^\circ)$; its angular radius is 14° ; the optical path length $l = 10$; the distance from the source to aperture plane $d = 9$; the output aperture $\bar{T}$ is a disk on the plane $z = d$ with center at $(.569, 0, 9)$; its radius is .341. The output energy pattern L on T is required to be uniform, that is, $L \equiv 1$. As in the example in the beginning of this section we computed $D(\rho(\mathbf{m}))$ for the hyperboloid and used it as the input $I(\mathbf{m})$ in (19). The two tables below show the maximal absolute change in computed increments h_n per iteration, respectively, for 23×23 and 39×39 grids. The parameters $\mu_n = \nu_n = .2$.

h_1	h_2	h_3	h_4	h_5
2.5389×10^{-3}	2.3444×10^{-3}	1.9430×10^{-3}	1.7380×10^{-3}	1.3867×10^{-3}

h_1	h_2	h_3	h_4	h_5
7.9931×10^{-4}	3.7788×10^{-4}	3.2376×10^{-4}	1.2729×10^{-4}	1.1831×10^{-4}

The CPU time on a SUN per iteration in the case of 23×23 grid was 15 seconds and in the case of 39×39 grid was 1 minute. It should be emphasized that these computations were carried out with the original ELLPACK modules.

5 Conclusions

In this paper we consider the problem of synthesizing a dual reflector system which transforms a given input power pattern into a prespecified power pattern on the output aperture. We describe here a new approach to formulating the problem analytically and present a numerical scheme for computing solutions. In this approach we are able to formulate precisely the energy conservation law and boundary conditions directly in terms of the polar radius describing the reflector. This formulation allows also to deal efficiently with different shapes of output apertures and, in fact, the latter can be used as an easily adjustable control parameter. We also describe a numerical scheme for solving synthesis equations. The scheme is based on Newton's method and its first implementation was made with the use of standard PDE's solver ELLPACK. Numerical simulations demonstrate clearly the feasibility of the method.

Acknowledgements Research sponsored by the AFOSR (AFSC) under Contracts F49620-91-C-0001 and F49620-92-C-0009. The United States Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation hereon.

References

- [1] B. E. Kinber. On two reflector antennas. *Radio Eng. Electron. Phys.*, 7(6):973–979, 1962.
- [2] B. E. Kinber. Inverse problems of the reflector antennas theory - geometric optics approximation. *preprint No. 38*, pages 1–48, 1984. in Russian.
- [3] V. Galindo Israel, R. Mittra, and A. G. Cha. Aperture amplitude and phase control of offset dual reflectors. *IEEE Transactions of Antennas and Propagation*, AP-27:154–164, 1979.
- [4] V. Galindo-Israel, W. A. Imbriale, and R. Mittra. On the theory of the synthesis of single and dual offset shaped reflector antennas. *IEEE Transactions of Antennas and Propagation*, AP-35(8):887–896, 1987.
- [5] V. Galindo Israel, W. A. Imbriale, R. Mittra, and K. Shogen. On the theory of the synthesis of offset dual-shpaed reflectors – case examples. *IEEE Transactions of Antennas and Propagation*, 39(5):620–626, May 1991.
- [6] B. S. Westcott, F. A. Stevens, and F. Brickell. Go synthesis of offset dual reflectors. *IEE Proceedings : H*, 128(1):11–18, February 1981.
- [7] B. S. Westcott. *Shaped Reflector Antenna Design*. Research Studies Press, Letchworth, UK, 1983.
- [8] B.S. Westcott, R. K. Graham, and I. C. Wolton. Synthesis of dual-offset shaped reflectors for arbitrary aperture shapes using continuous domain deformation. *IEE Proceedings:H*, 133(1):57–64, February 1986.

- [9] C. J. Sletten. Reflector antennas and surface shaping techniques. In C. J. Sletten, editor, *Reflector and Lens Antennas*. Artech House, Norwood, MA, 1988.
- [10] W. V. T. Rusch. Quasioptical antenna design (section 3.4). In A. W. Rudge, K. Milne, A. D. Olver, and P. Knight, editors, *The handbook of antenna design, Volumes 1 and 2*. Peter Peregrins Ltd., London, UK, 1986.
- [11] V. I. Oliner, E. J. Newman, and L. Prussner. Differential-geometric methods in design of single and dual reflector antennas. *preprint*, 1991.
- [12] L. Marder. Uniqueness in reflector mappings and the Monge-Ampere equation. *Proc. R. Soc.*, A(378):529–537, 1981.
- [13] V. I. Oliner. Near radially symmetric solutions of an inverse problem in geometric optics. *Inverse Problems*, 3:743–756, 1987.
- [14] J. S. Schruben. Formulation of reflector-design problem for a lighting fixture. *J. of the Optical Society of America*, 62(12):1498–1501, 1972.
- [15] J. R. Rice and R. F. Boisvert. Solving elliptic problems using ellpack. In *Springer series in computational mathematics*, volume 2. Springer-Verlag, New York, 1985.
- [16] E. Kreyszig. *Introduction to Differential Geometry and Riemannian Geometry*. Univ. of Toronto Press, 1975.
- [17] V. I. Oliner. Ray tracing and energy conservation equations for systems with two reflectors. *preprint*, 1992.

- [18] V. I. Oliker. Hypersurfaces in R^{n+1} with prescribed Gaussian curvature and related equations of Monge-Ampere. *Comm. in PDE's*, 9(8):807–838, 1984.
- [19] V. I. Oliker, E. J. Newman, and L. Prussner. On the reflector design problem. *preprint*, 1992.

EVALUATION OF FREQUENCY SELECTIVE REFLECTOR ANTENNA SYSTEMS

T.K. Wu
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

S.W. Lee
Dept. of Electrical Engineering
University of Illinois
Urbana, IL 61801-2991

M.L. Zimmerman
Analex Corporation
Brook Park, OH 44142

ABSTRACT

The performance of a four-frequency band (S/X/Ku/Ka) frequency selective surface (FSS) with double-ring element is evaluated both for a flat panel development and for a curved subreflector in a dual reflector antenna. Good agreement is observed between the measured and computed data for the flat panel FSS. To accurately predict the FSS effects on a dual reflector antenna's radiation pattern, we take into account the FSS subreflector's transmitted/reflected field variation as functions of the polarization and incident angles w.r.t. the local coordinates. An FSS transmission/reflection coefficient lookup table is computed for TE and TM polarization as well as various incident angles based on the planar FSS model. Next, the hybrid Geometric Optics (GO) and Physical Optics (PO) technique is implemented with the lookup table to determine the FSS effects in the dual reflector antenna.

INTRODUCTION

Frequency selective surfaces (FSS) are often employed in the reflector antenna system of a communication satellite¹⁻³ or a deep space exploration vehicle⁴⁻⁷ for multi-frequency operations. Typically, an FSS consists of one or more grids supported by multiple layers of dielectrics. The grids are thin conducting periodic elements etched on thin substrates, such as Kapton sheets. Because these grids can be designed to resonant at the specified frequencies, the FSS will reflect waves at these frequencies and pass waves at other frequencies. Hence only one single main reflector is required for the multi-frequency operation, if the FSS is employed for the subreflector and the different frequency feeds are optimized independently and placed at the real and virtual foci of the subreflector. For example, the Voyager FSS was designed to duplex S and X bands⁴. Here the S band feed is placed at the prime focus of the main reflector, and the X band feed is placed at the Cassegrain focal point. Note that only one main reflector is required for these two band operation. Thus tremendous reduction in weight, volume and most important the cost of the antenna system are achieved with the FSS subreflector.

In the past, the FSS subreflector effect in a dual reflector antenna system¹ was determined by a simple gain reduction based on the reflection/transmission loss of an ideal flat panel analysis. The surface curvature, different incident angles/polarizations and the phase front distortion of the FSS subreflector were not considered. In fact, with FSS, the radiation pattern of a dual reflector exhibits higher sidelobe and cross-polarization levels. However, these FSS effects can not be predicted by the simple gain reduction approach.

To improve the accuracy, it is necessary to incorporate in the pattern computational model the FSS subreflector's transmitted/reflected field variation in polarization and incident angle w.r.t. the local coordinates. It is also assumed that the local transmitted/reflected fields of the subreflector are available based on the planar FSS model. An FSS transmission/reflection coefficient lookup table is generated with TE and TM polarization as well as various incident angles. Two methods, i.e., the PO-PO method⁸ and the hybrid GO and PO technique⁹ may next be implemented with this lookup table to compute the dual reflector radiation pattern.

Recently, NASA's Cassini project requires the use of multiple RF frequencies at S,X,Ku and Ka-bands for science investigations and data communication links. A single high gain antenna (HGA) with an FSS subreflector, as illustrated in Figure 1, is proposed to allow a Cassegrain configuration at X (7.2 and 8.4 GHz) and Ka (32 and 34.5 GHz) bands, and a prime focus configuration at S (2.3 GHz) and Ku (13.8 GHz) bands⁵. In this paper, the design and performance of a four-frequency FSS (S/X/Ku/Ka bands) with double-ring element are demonstrated. The selection of the concentric double-ring elements is based on the fact that the ring's geometry is particularly conforming to the circular polarization requirement of the Cassini project and the ease of manufacture. Next, the hybrid GO and PO approach is employed to assess a dual reflector antenna with an FSS subreflector.

FLAT PANEL FSS WITH DOUBLE-RING ELEMENT

A double screen design approach, as depicted in Figure 2, was considered for the

four-frequency band (S/X/Ku/Ka) flat panel FSS development. The top FSS screen is called the Ka-add-on FSS, which passes the S/X/Ku band signals but reflects the Ka-band signal. The bottom FSS screen is called the three-frequency FSS, which passes the S and Ku band signals but reflects the X-band signal. The two screens are separated by a 1.75" thick Kevlar honeycomb sandwich or foam spacer. In operation, assuming multiple band signals incident from the top of the FSS, the S and Ku band signals will pass this double screen FSS. The Ka-band signal will be reflected by the top screen. The X-band signal will pass through the top screen, then it is reflected by the bottom screen and back out through the top screen.

Figure 3 shows the double-ring element design for the Ka-add-on FSS screen. The double-ring patch element array is etched on a 0.002" thick Kapton sheet and then bonded to the Kevlar honeycomb sandwich. The transmission characteristics of this FSS are also plotted in Figure 3. The computed results are based on the theory described earlier⁷. Only representative measured data at 30° is given here to demonstrate the good agreement between the computed and measured data. Similarly, the three-frequency FSS design and performance are given in Figure 4. The double screen FSS performance is summarized in TABLE 1.

A DUAL REFLECTOR ANTENNA WITH FSS SUBREFLECTOR

To accurately predict the FSS effects in the dual reflector antenna system, we take into account the subreflector's transmitted/reflected field variation with the polarization and the various incident angles w.r.t. the local coordinates. Based on the planar FSS model of

the previous section, an FSS transmission/reflection coefficient lookup table is generated for TE and TM polarizations as well as the various incident angle (θ_i, ϕ_i). The hybrid GO and PO technique is next implemented with lookup table to compute the antenna patterns⁹.

Consider the Cassegrain reflector antenna of Figure 1. Here the main dish is a 12' diameter paraboloid with $F/D = 0.3377$ and the FSS subreflector is a 17.25" diameter hyperboloid with eccentricity = 1.5. All the feeds are assumed to have a $\cos^{22.5} \theta$ pattern. The gain patterns with/without FSS are given in Figures 5-10 for frequencies at 2.3, 7.2, 8.4, 13.8, 32 and 34.5 GHz, respectively. As expected that the sidelobe and cross-polarization levels are higher with the presence of the four-frequency FSS. The peak gain is reduced with the FSS and is summarized in TABLE 2. The table shows the peak gain reduction due to FSS insertion loss is less than 1 dB for all the in-band frequencies except 34.5 GHz. It is interesting to note that at 13.8 GHz the transmission loss of the planar FSS model varies from 0.54 to 0.94 dB for incident angles (θ_i, ϕ_i) from ($0^\circ, 0^\circ$) to ($45^\circ, 90^\circ$). However, the dual reflector's gain reduction with FSS is only 0.2 dB. The sidelobe level is about the same with/without the FSS. This is mainly due to the higher transmission/reflection loss occurs only at large incident angles, namely near the edge of the subreflector. This also indicates that the reflector's aperture field variation caused by the FSS yields good directivity and sidelobe performance. It should be pointed out that this hybrid GO-PO method has the capability of including the diffracted fields from the subreflector rim. In the cases considered here, including this edge diffracted fields only cause a 0.1 dB difference in the peak gain.

CONCLUDING REMARKS

The performance of a four-frequency FSS with double-ring element is evaluated for the flat panel development and for a curved subreflector in a dual reflector antenna system. Good agreement is obtained between the measured and computed data for the planar FSS model. The FSS effects in a four-frequency Cassegrain reflector antenna are accurately evaluated by taking into account the surface curvature and the FSS subreflector's transmitted/reflected fields variation as functions of the polarization and the incident angles w.r.t the local coordinates. Although only a Cassegrain reflector is considered here, the technique is very robust and can be applied to any dual shaped reflector antenna system.

TABLE 1. Loss (dB) Summary of the Planar FSS Model.

frequency (GHz)	$(\theta_i, \phi_i) = (0^\circ, 0^\circ)$	$(30^\circ, 0^\circ)$		$(45^\circ, 0^\circ)$	
		TE	TM	TE	TM
2.3	0.71	0.71	0.5	0.8	0.3
7.2	0.43	0.33	0.42	0.37	0.88
8.4	0.13	0.11	0.15	0.11	0.18
13.8	0.54	0.82	0.57	0.94	0.39
32	0.69	0.62	0.71	0.52	0.76
34.5	1.09	1.24	1.82	1.29	2.1

TABLE 2. FSS Reflector Peak Gain Summary

frequency (GHz)	peak gain (dB)	FSS loss (dB)
2.3	37.94	0.72
7.2	47.61	0.41
8.4	48.95	0.16
13.8	53.5	0.2
32	60.57	0.68
34.5	61.22	1.09

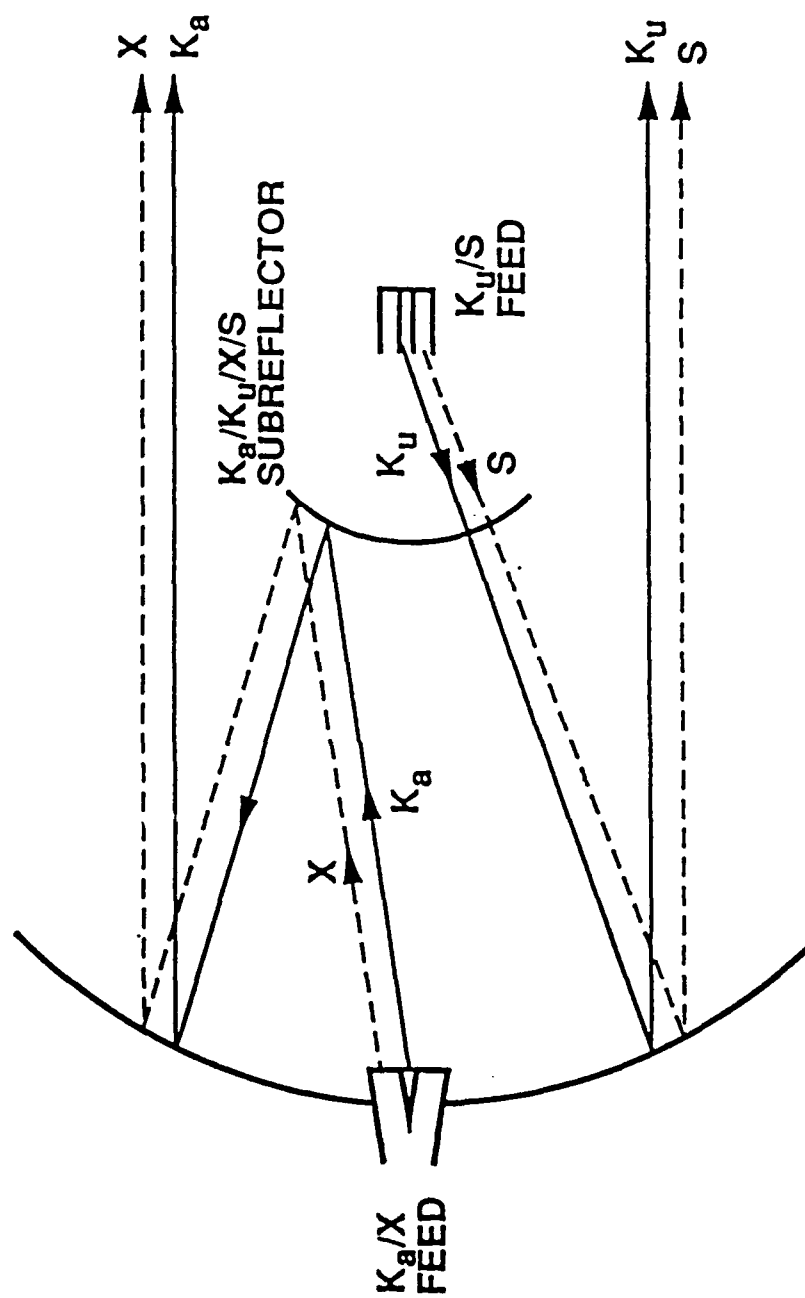
ACKNOWLEDGMENT

The work described in this article was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The authors wish to thank Dr. Ken Woo for his support and Mr. Cosme Chavez for taking the FSS measurement.

REFERENCES

1. Agrawal, V.D. and Imbriale, W. A. (1979) "Design of a dichroic Cassegrain subreflector," IEEE Trans., vol. TAP-27, no.4, pp. 466-473.
2. Lee, S.W. et al (1991) "Designs for the ATDRSS Tri-band reflector antenna," paper presented at the 1991 Int. IEEE AP-S Symposium, Ontario, Canada, pp. 666-669.
3. Ueno, K. et al (1991) "Characteristics of FSS for a multi-band communication satellite", paper presented at the 1991 Int. IEEE AP-S Symposium, Ontario, Canada.
4. Schennum, G.H. (1975) "Frequency-selective surfaces for multiple frequency antennas," Microwave Journal, vol. 16, pp.55-57.
5. Wu, T.K. (1991) "Double-square-loop FSS for multiplexing four (S/X/Ku/Ka) bands", paper presented at the 1991 Int. IEEE AP-S Symposium Digest, Ontario, Canada, pp.1885-1888.
6. Huang, J. and Lee, S.W. (1991) "Tri-band FSS with circular ring elements," paper presented at the 1991 Int. IEEE AP-S Symposium Digest, Ontario, Canada.
7. Wu, T.K. et al (1992) "Multi-ring element FSS for multi-band applications," paper presented at the 1992 Int. IEEE AP-S Symposium, Chicago, IL.
8. Bresciani, D. and Contu, S. "Scattering analysis of dichroic subreflectors," Electromagnetics, v. 5, no. 4, pp. 375-407, 1985.
9. Zimmerman, M.L. and Lee, S.W. (1991) "Integrated reflector antenna design and analysis," Univ. of Illinois, Urbana-Champaign, TR MZ1-91, sec. 6.

FIGURE 1. PROPOSED CASSINI HIGH GAIN ANTENNA



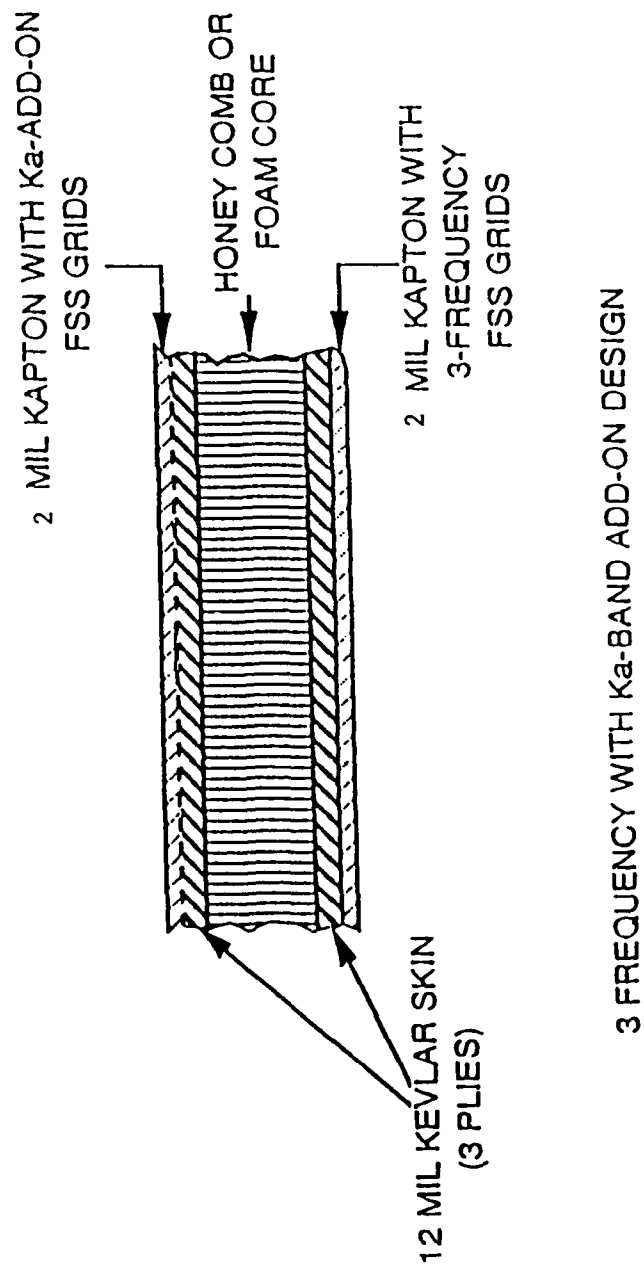
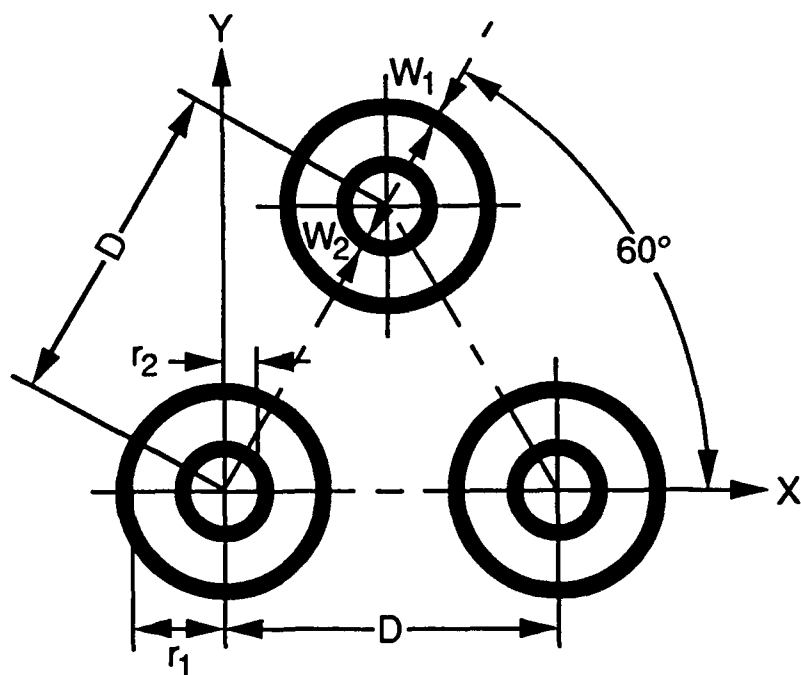


Figure 2. Flat panel FSS Design Configurations



$$\begin{aligned}
 D &= .169'' \\
 r_1 &= .047'' \\
 r_2 &= .028'' \\
 w_1 = w_2 &= .01''
 \end{aligned}$$

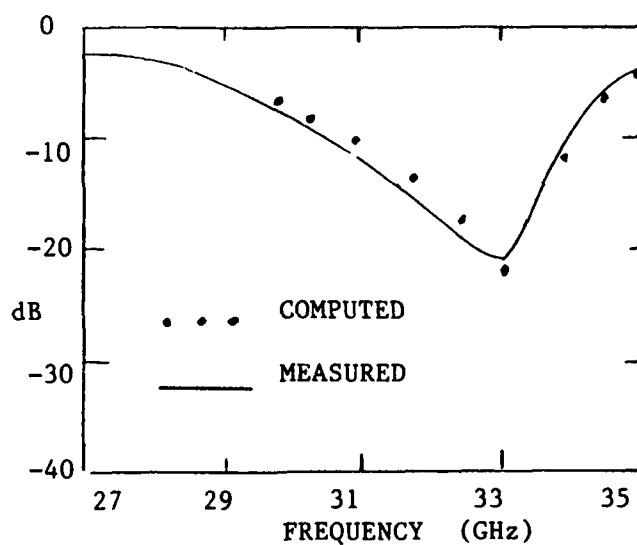


Figure 3. Design and transmission performance of the Ka-add-on FSS with double-ring element.

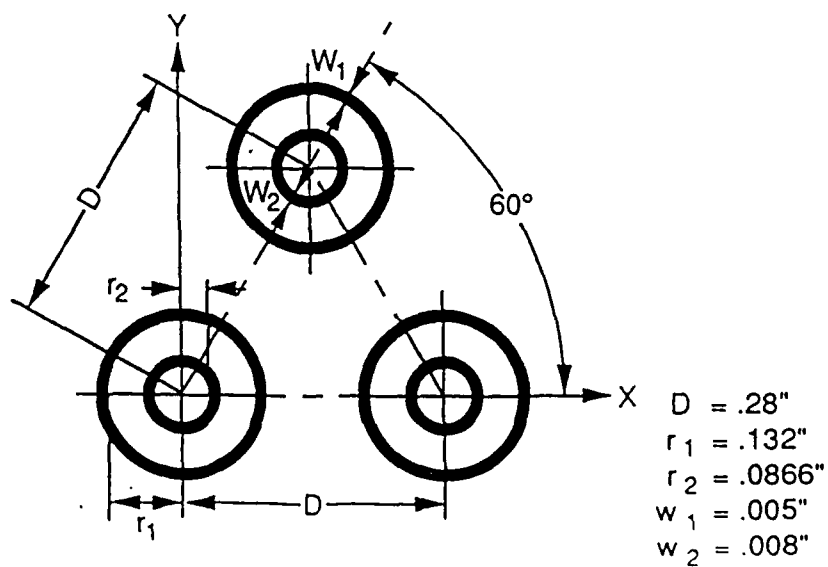
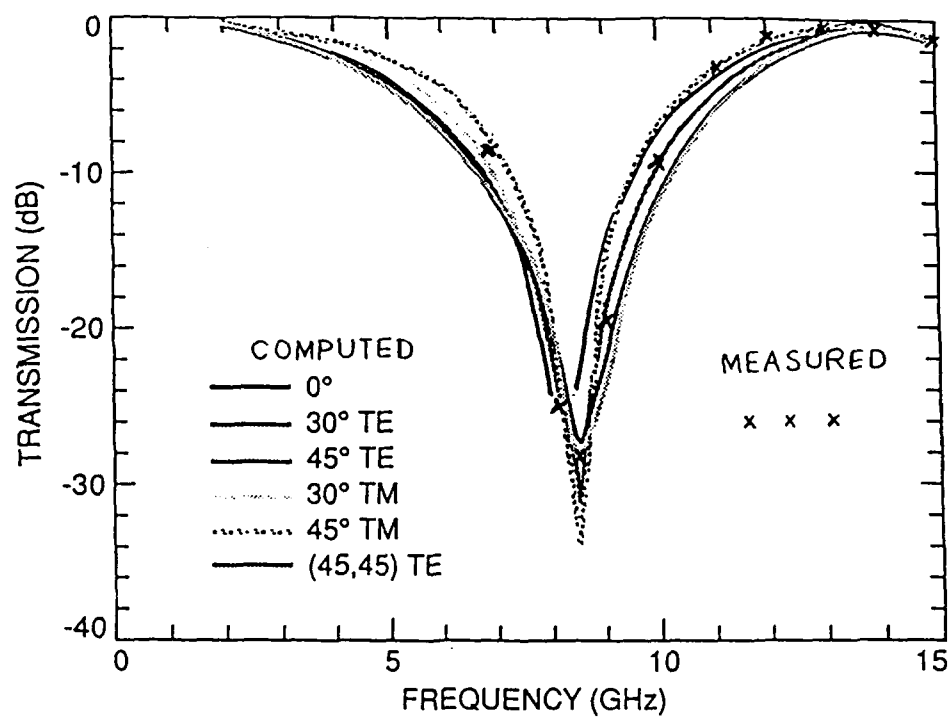


Figure 4. Design and transmission performance of the three-frequency FSS with double-ring element.

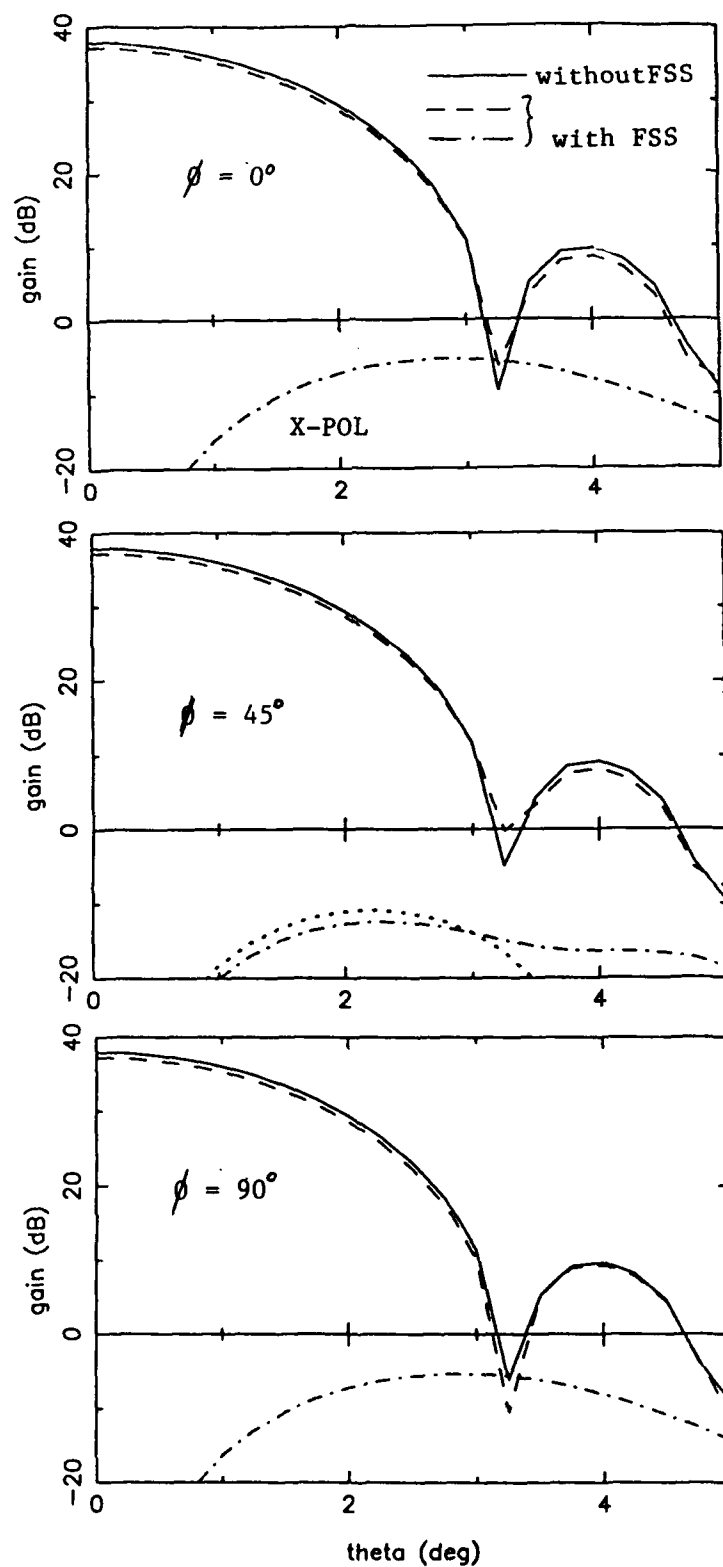


Figure 5. Gain pattern of the Cassegrain reflector antenna at 2.3 GHz.

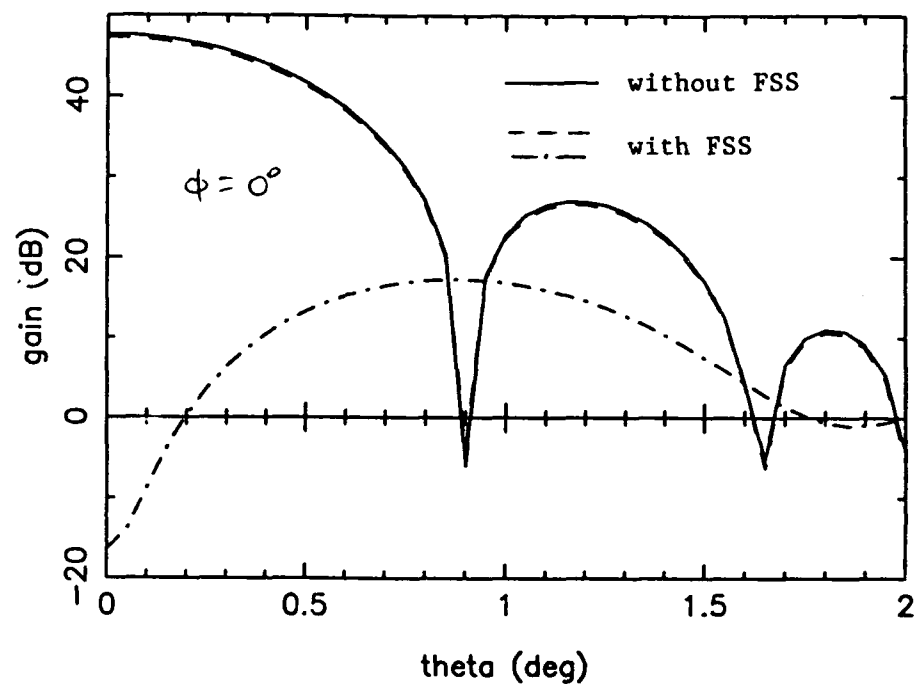


FIGURE 6. GAIN PATTERN OF THE CASSEGRAIN ANTENNA AT 7.2 GHZ.

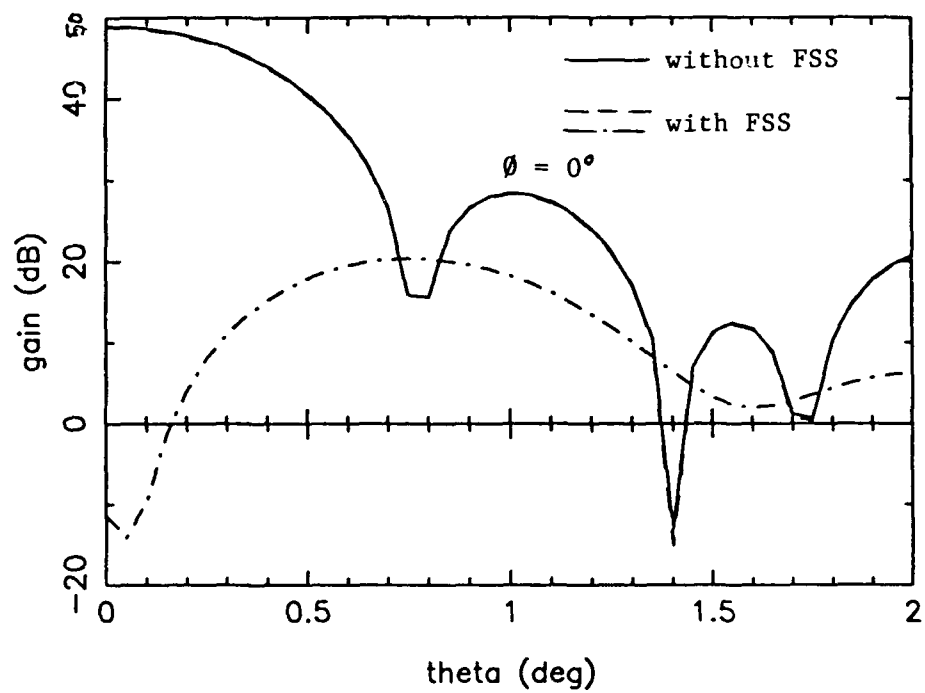


FIGURE 7. GAIN PATTERN OF THE CASSEGRAIN REFLECTOR ANTENNA AT 8.4 GHZ.

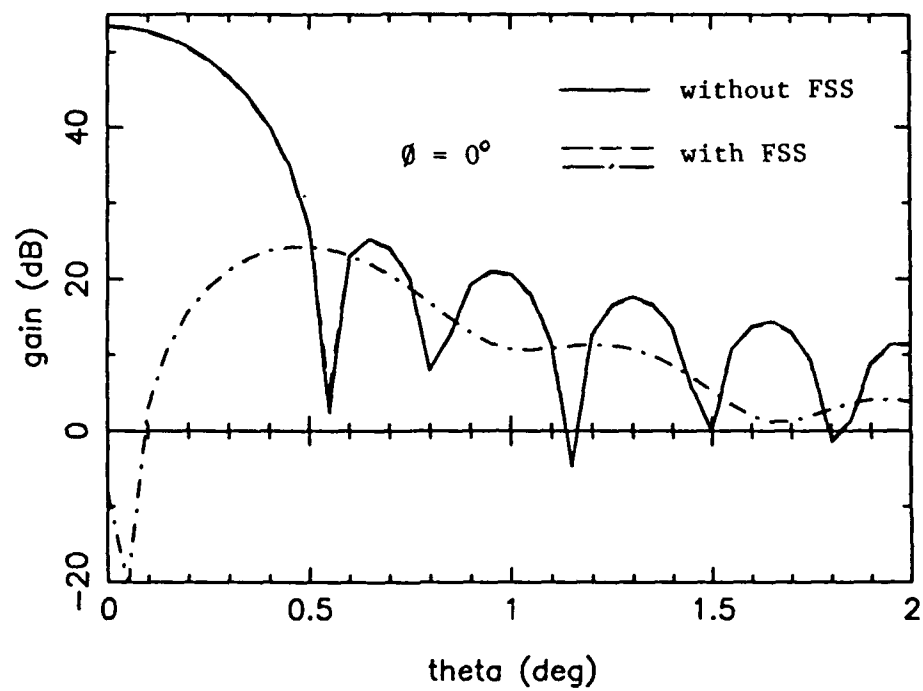


FIGURE 8. GAIN PATTERN OF THE CASSEGRAIN REFLECTOR ANTENNA AT 13.8 GHZ.

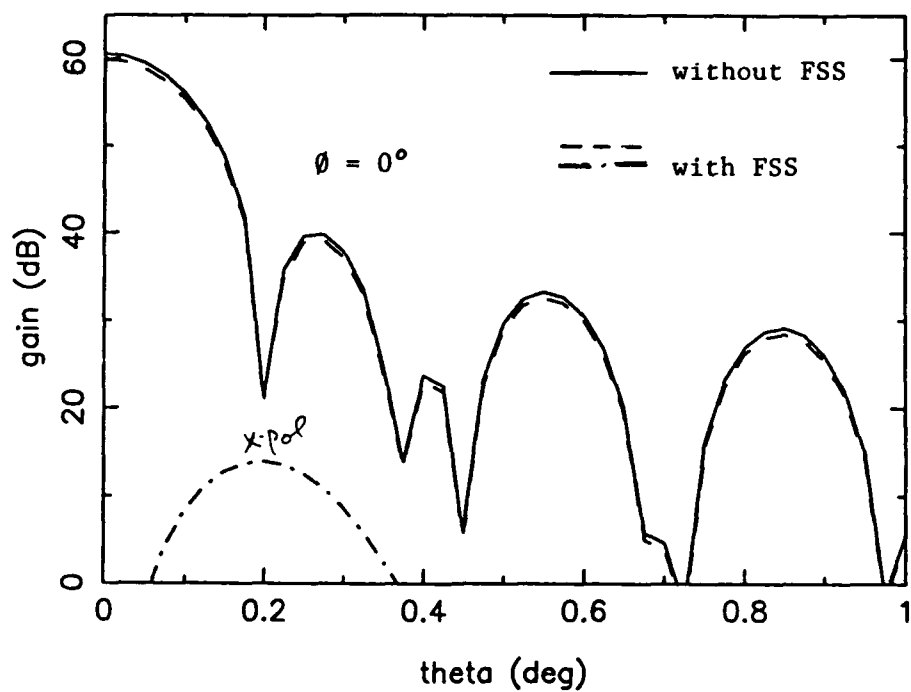


FIGURE 9. GAIN PATTERN OF THE CASSEGRAIN REFLECTOR ANTENNA AT 32 GHZ.

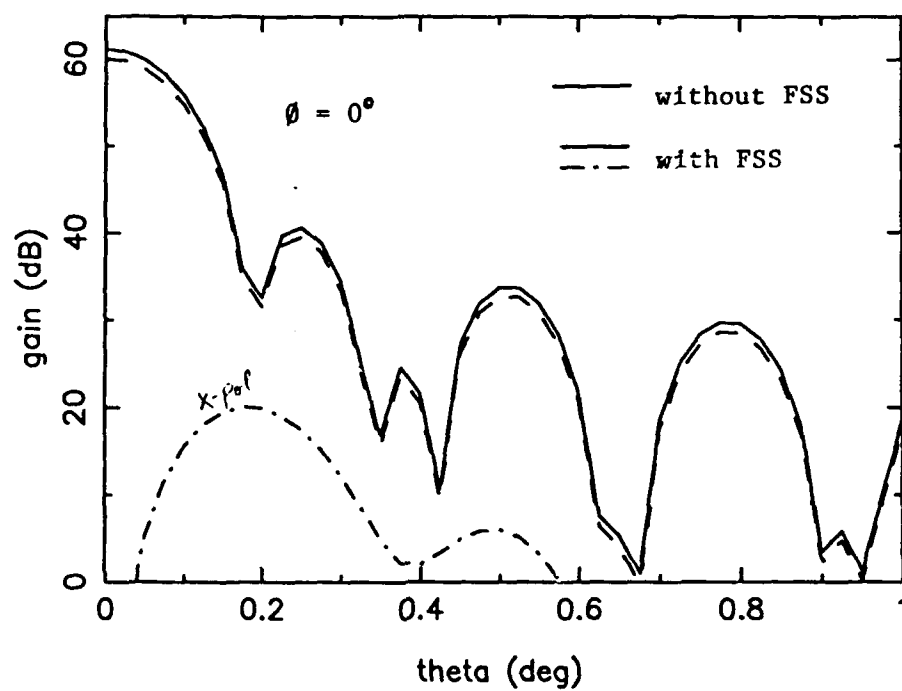


FIGURE 10. GAIN PATTERN OF THE CASSEGRAIN REFLECTOR ANTENNA AT 34.5 GHZ.

Productivity Gains Using Multiple-Channel, Multiple-Frequency Measurement Techniques for Testing the E2C Antenna

James Scherer, Senior Vice President of
Technology

Robert Magatagan, Senior Scientist
John Miller, Production Test Manager
Gary Amaral, Production Test Supervisor

Loral Randtron Systems
130 Constitution Drive
Menlo Park, CA 94025

John Swanstrom

Hewlett-Packard
Santa Rosa Systems Division
1400 Fountaingrove Parkway
Santa Rosa, CA 95403

Abstract

This paper presents a feasibility study of the productivity improvements that are possible for the production test of the E2C antenna, using multiple-parameter, multiple-frequency measurement techniques. The measurement requirements for the antenna are presented along with the current measurement times. A multiple-channel, multiple-frequency measurement technique is described which will greatly reduce the measurement times. The new measurement times are calculated, and used to determine if the productivity improvements are justified financially.

An economic analysis is included also, which examines the financial impact of the improved productivity, and compares this to the cost of implementing the new measurement system. The financial analysis calculates the payback period, return on investment, net present value, and internal rate of return.

Introduction

The tendency is to leave successful endeavors alone, but not at Randtron. A progressive company, Randtron is concerned with maintaining its productivity and its competitive edge. At Randtron, we produce the E2C antennas, a very successful program for us.

Because the acceptance testing of the E2C antenna consumes a significant portion of test time (and money), we performed a feasibility study, in conjunction with Hewlett-Packard, to determine ways to reduce test time. Reducing the time reduces the cost of producing the antennas, keeps us competitive, and improves the profitability of our company. This paper presents the results of the feasibility study.

The E2C Antenna Program

The E2C antenna is used for airborne early warning and command for the U.S. Navy. It provides electronic surveillance and protection for the U.S. Naval fleet worldwide. The antenna is a large, 24-foot diameter disk that is mounted atop the E2C Hawkeye aircraft, which can be either land or carrier based. Randtron has produced over two hundred E2C antennas since 1972. Our current production averages 20 units per year. We were selected as the sole supplier of E2C antennas because of our ability to produce quality antennas in large volumes at low cost.

Required Antenna Patterns for E2C Antenna

Like all critical flight hardware, the E2C antenna undergoes a complete acceptance test procedure before delivery. The procedure consists of a complete set of antenna patterns, absolute gain measurements, VSWR measurements, a power handling verification, and a measurement of isolation between test channels. The focus of this paper is on the productivity gains that can be made by making improvements in the antenna pattern measurements. Table 1 summarizes the required antenna patterns. There are 185 azimuth patterns, 26 elevation patterns, and 3 relative-phase measurements.

Present Measurement Techniques and Times

Our measurement techniques for the antenna patterns have been the same since 1972. We utilize older measurement instrumentation; a single channel receiver (SA 1742), and a rectangular paper pattern recorder (SA 1520). One pattern (at a single frequency) is produced for each 360 degree rotation of the antenna.

Because of the inertia and momentum associated with large antenna size, the maximum positioner velocity is one revolution per minute. Therefore it takes one minute to acquire the pattern, and one minute to rewind the positioner. We average one azimuth pattern approximately every four minutes. Data acquisition time is 12 hours 20 minutes (185 patterns x 4 minutes/pattern) for the azimuth patterns. For the elevation patterns, the maximum positioner speed is 0.2 degrees/second or 7 minutes per elevation cut. It takes four hours to measure all of the elevation patterns. The absolute gain measurements take four hours, and the relative phase measurements take 1.5 hours to complete. When the measurement overhead is considered, the acceptance testing and data analysis takes 3.5 days to complete. We utilize a four-person crew for the duration of the test.

Technique for Reducing Measurement Time

One of the reasons the data acquisition takes so long is that only one pattern is acquired for each complete rotation of the antenna. If several patterns at different test ports and frequencies could be taken for each rotation of the antenna, this would greatly increase the amount of data acquired in one rotation of the antenna, and significantly reduce the acquisition time. Thus, a multiple-channel, multiple-frequency test technique is indicated. With this technique, at each angular increment of the test antenna, all test ports can be measured at all frequencies. Since the instrumentation is very fast, the data acquisition can be accomplished well before the next angular increment needs to be measured.

Recommended Measurement Technique

Figure 1 shows the instrumentation configuration for measuring the E2C antenna. Key parts of the measurement system are the PIN switches, which allow rapid switching between the test channels, and the fast frequency agility of the system. The PIN switch controller orchestrates all of the data acquisition, and sequences the switching of the PIN switches.

For the E2C antenna, the test sequence is:

- Establish a gain and phase reference at each frequency. This allows each of the subsequently acquired antenna patterns to have a gain and phase reference.
- Measure the sum, difference, auxiliary 1, and auxiliary 2 test channels at their four different operating frequencies at seven different elevation angles.
- Measure the IFF test channel at the required nine different frequencies at seven elevation angles.
- Measure the cross-polarized IFF channel antenna patterns at seven elevation angles.

With this technique, all 185 azimuth patterns are acquired in twenty one rotations of the antenna.

Predicted Data-Acquisition Times

The details for computing the data-acquisition times follow. Frequency switching speeds range from 5-18 msec per frequency depending on the configuration and the size of the frequency step. For this example, telephone line modems are used to communicate frequency changes to the microwave source, and the frequency switching speed is 18 msec. Switching between test channels requires 0.4 msec per channel.

First, measure the sum, difference, auxiliary 1, and auxiliary 2 test channels at four frequencies in one rotation of the antenna.

Acquisition time per angular increment = 78.4 msec.

$$\left[\frac{18 \text{ msec.}}{\text{Freq.}} + \left(\frac{0.4 \text{ msec.}}{\text{channel}} \right) \times 4 \text{ chan.} \right] \times 4 \text{ freq.}$$

Maximum positioner velocity:

$$\frac{0.25 \text{ deg}}{78.4 \text{ msec}} = 3.2 \text{ deg/sec}$$

Time to measure four channels at four frequencies at one elevation angle:

$$\frac{360 \text{ deg}}{3.2 \text{ deg/sec}} = 112.5 \text{ sec}$$

To measure the IFF test channel at nine frequencies at one elevation angle, the acquisition time per angular increment is 165.6 msec, calculated as follows:

$$\left[\frac{18 \text{ msec}}{\text{freq.}} + \left(\frac{0.4 \text{ msec.}}{\text{channel}} \right) \times 1 \text{ chan.} \right] \times 9 \text{ freq.}$$

Maximum positioner velocity:

$$\frac{0.25 \text{ deg}}{0.165 \text{ sec.}} = 1.5 \text{ deg/sec.}$$

Time to measure one channel (IFF) at 9 frequencies at one elevation angle:

$$\frac{360 \text{ deg}}{1.5 \text{ deg/sec.}} = 240 \text{ sec.}$$

Total data acquisition time for azimuth patterns is: 112.5 sec./rotation x 7 elevation angles + 240 sec./rotation x 7 elevation angles x 2 polarizations = 69 minutes. This would require twenty one rotations of the test antenna.

For elevation patterns, the data acquisition time will be limited by how fast the positioner can move in elevation, which is 0.2 degrees per second due to the gear reduction required in the positioner.

Elevation patterns: Measure all five channels at 13 different frequencies.

Acquisition time per angular increment = 260 msec calculated as follows:

$$\left[\frac{18 \text{ msec}}{\text{freq.}} + \left(\frac{0.4 \text{ msec}}{\text{channel}} \right) \times 5 \text{ chan.} \right] \times 13 \text{ freq.}$$

Maximum positioner velocity (elevation):

$$\frac{0.25 \text{ deg}}{0.260 \text{ sec.}} = 0.96 \text{ deg/sec}$$

Maximum positioner velocity is 0.2 deg./sec., so measurement speed is positioner limited.

Time to measure elevation patterns:

$$\frac{80 \text{ deg}}{0.2 \text{ deg/sec.}} = 400 \text{ sec.} \approx 7 \text{ min.}$$

Total data acquisition times:

Azimuth patterns: 69 minutes

Elevation patterns: 7 minutes

Data-Acquisition Time vs. Measurement Time

There is a difference between the data-acquisition time and total measurement time. Data-acquisition time is the time it takes the instrumentation to acquire the data, while measurement time is the time it takes to perform the complete test, which includes measurement overhead such as boresighting, breaks for personnel, etc.

For our feasibility study to be valid, it is important to accurately estimate the total measurement time, rather than just data-acquisition times. For the proposed system, we calculated the new data-acquisition times, and then, based on our experience, added our estimates of measurement overhead to arrive at total measurement time.

Data Analysis

Our present method of data analysis is a manual process which takes four hours to complete and utilizes four people. With the new system we are considering, the data-acquisition is done through measurement-automation software and is computer controlled. A software analysis program would be used to analyze all the antenna-pattern data. Automating the data-analysis process will result in a tremendous savings in test time and cost. The four people who laboriously analyze the patterns will be free to perform more profitable tasks, and the analysis which now takes hours to perform will be accomplished in minutes by the computer.

To develop our customized measurement analysis software will require a capital investment. The estimated cost will be included

in the financial analysis portion of the feasibility study. The return on this capital investment will be the cost savings realized each time the analysis is completed in minutes instead of hours.

Cost Savings Due To Productivity Improvements

To calculate the cost savings that can be realized, we summarized the factors that contribute to the cost of performing the pattern testing with the current and proposed methods; these cost factors are shown in table 2. The test cost was determined by multiplying the number of test personnel by their hourly cost and then by the test time. In this paper, we used industry-standard labor costs rather than Randtron's labor costs. As can be seen from the table, the proposed new method will save \$4,697 each time the test is performed.

Required Capital Expenditure

Figure 1 shows the block diagram of the proposed measurement instrumentation which will provide the multiple-parameter, multiple-frequency measurement capability. Note that PIN switches and the PIN switch controller are used to rapidly switch between the multiple test ports. The HP 83621A synthesizers provide the fast multiple-frequency switching.

	<u>Current method</u>	<u>Proposed method</u>
Data acquisition times:		
Azimuth patterns	12 hours 20 minutes	.69 minutes
Elevation patterns	3 hours 45 minutes	7 minutes
Gain measurements	3-4 hours	included
Phase measurements	1.5 hours	included
Data analysis	4 hours	25 minutes
Total measurement time	3.5 days	1.5 days
Number of personnel	4	2
Cost of test	\$6,039	\$1,342

Table 2: Factors that contribute to the cost of performing the acceptance test

The total price for the measurement instrumentation is \$300,410. It is estimated that the cost to change over to the new system will be \$10,000, and the cost to write the analysis software will be \$40,000. Thus, the total cost to implement the new measurement system is \$350,410.

Economic Analysis

Table 3 shows the financial calculations used to determine the economic feasibility of this proposal. Assuming a useful life of seven years, the internal rate of return on the project is 18%, with a payback period of less than 4 years. The return on investment is 24%, and the net present value is \$89,000.

Profit Improvement Proposal

The economic analysis clearly indicates that this is a sound financial investment. It has an \$89,000 net present value and an 13% internal return on investment. Beyond the savings from reduced test time, which can be easily quantified, there are other benefits; improved quality of the

measurements, increased technological capacity, better reliability, greater up-time of the system, and much lower maintenance costs for the new system.

Summary

Randtron has shared a feasibility study which examined how modern measurement techniques can be utilized to reduce test times and costs. The benefits to Randtron are lower costs which will insure our competitive position for the future, and improvement in our long term profitability.

Conclusions

The conclusions are clear: to be competitive in the future, companies need to look for better ways of making measurements today.

References

1. John Swanstrom, "Financially Justifying an Antenna/RCS Measurement System", 11th Annual Meeting and Symposium, Antenna Measurement and Techniques Association, Monterey, CA, October 9-13, 1989, pp. 5-27 to 5-30.

<u>Channel</u>	<u>Frequencies</u>	<u>Azimuth Movement</u>	<u>Elevation Movement</u>	<u># of patterns</u>
Sum	F1-F4	360/0.5	-9 to 9 by 3	28
Delta	F1-F4	360/0.5	-9 to 9 by 3	28
IFF	F5-F13	360/0.5	-9 to 9 by 3	45
IFF x-pol.	F5-F13	360/0.5	-9 to 9 by 3	28
Sum + Aux 1	F1-F4	360/0.5	-9 to 9 by 9	24
Sum + Aux 2	F1-F4	360/0.5	-9 to 9 by 9	24
Alignmemnt	F1-F4	360/0.5	0	8
Sum	F1-F4	0	-40 to +40	4
Sum	F1-F4	0	-18 to +18	4
IFF	F5-F13	0	-40 to +40	9
IFF	F5-F13	0	-18 to +18	9

Table 1: Required antenna patterns for E2C acceptance test

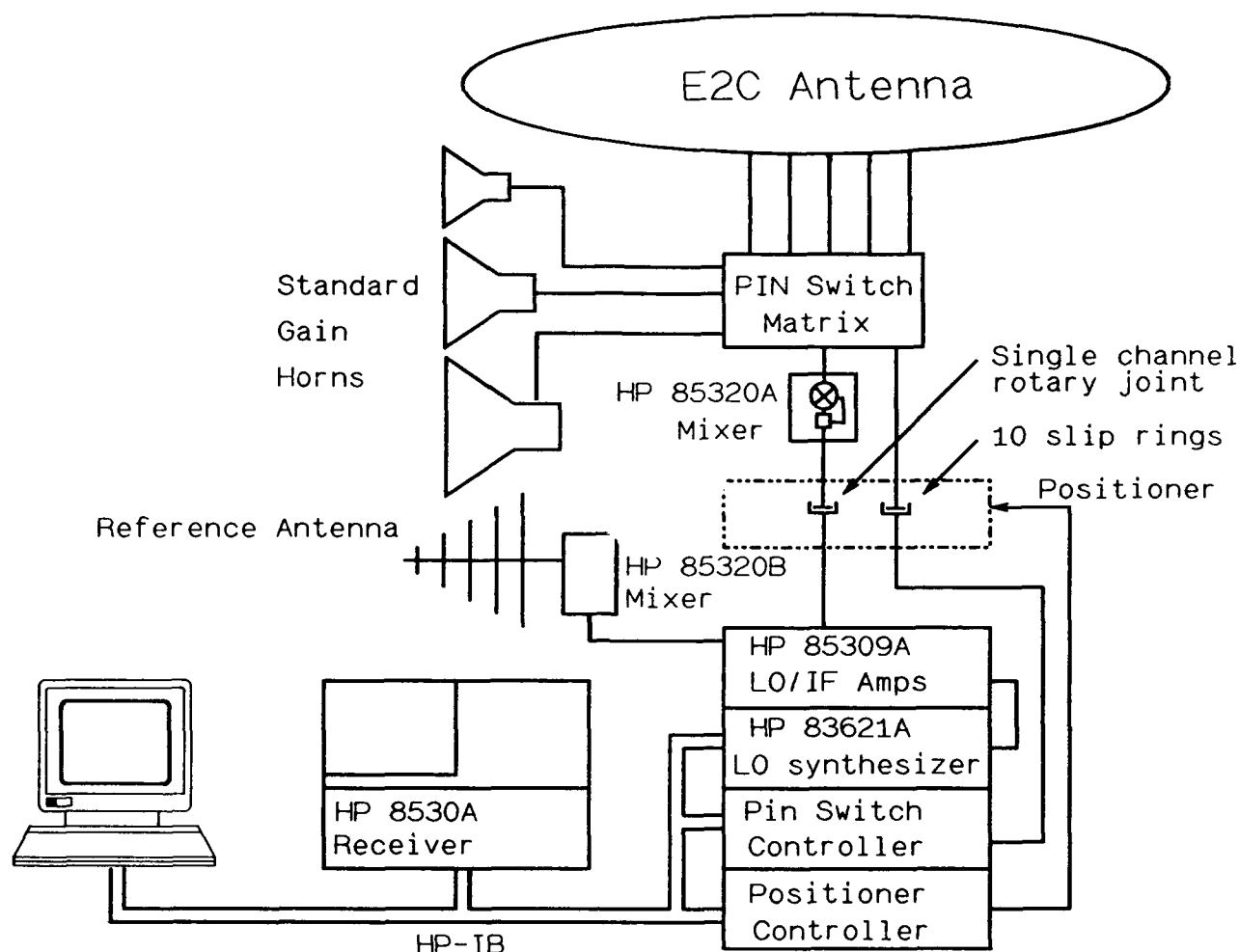


Figure 1: Proposed instrumentation configuration for measuring E2C antennas

COST OF SAVINGS SUMMARY								
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	
Direct Labor Productivity (4% per year inflation)	93942	97700	101608	105672	109899	114295	118867	
Cost savings due to maintenance and repair	0	0	0	0	0	0	0	
Annual Cost Savings	93942	97700	101608	105672	109899	114295	118867	
<u>Depreciation of new equipment</u>								
Cost of Instrument	300410							
Plus Installation	10000							
TOTAL	310410							
Useful life	3 years							
Method:	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Total
Sum of Year's Digits	155205	103470	51735					310410
<u>Straight Line</u>	<u>103470</u>	<u>103470</u>	<u>103470</u>					<u>310410</u>
Difference	51735	0	-51735	0	0	0	0	0
<u>Cash Flow Analysis - SYD Depreciation</u>								
Initial expense								
1. New Measurement System	-300410							
Installation	-10000							
2. Software (non-depreciable)	-40000							
3. Annual Cash Savings	93942	97700	101608	105672	109899	114295	118867	741962
4. Depreciation	155205	103470	51735	0	0	0	0	310410
5. Profit Improvement (3-4)	-61263	-5770	49873	105672	109899	114295	118867	431572
6. Tax (34%)	-13600	-20829	-1962	16957	35926	37366	36880	146734
7. Net Profit Improvement (5-6)	-40434	-3808	32916	69744	72533	75435	79452	264837
8. Cash Flow (3-6)	-336810	114771	99662	84651	69744	72533	75435	595247
Payback period	3.7 years							
Return on Investment	0.24							
Net Present Value @ 10% interest	89004							
Internal Rate of Return	18.44%							

Table 3: Financial calculations to determine economic feasibility

CALIBRATION OF MISMATCH ERRORS IN ANTENNA GAIN MEASUREMENTS

JOHN MCLAUGHLIN
ROBERT SHOULDERS

HEWLETT-PACKARD COMPANY
SANTA ROSA SYSTEMS DIVISION
1400 FOUNTAINGROVE PKWY
SANTA ROSA, CA 95403

I. ABSTRACT

This paper describes a calibration technique for reducing the errors due to mismatch between the measurement receiver and the antenna in microwave antenna relative gain measurements. In addition, this technique provides an accurate method for measuring the input return loss of the antenna under test.

In this technique, a microwave reflectometer is mounted between the measurement receiver and the antenna test port. The reflectometer is calibrated and used to measure the return loss of both the test and calibration antennas. Using this information in conjunction with the HP 8530A antenna gain calibration, the corrected gain of the antenna under test is computed.

Compact range antenna measurements verifying the calibration model and error analysis are presented. Practical implementation considerations are discussed.

Keywords: Calibration, Gain Measurement, Microwave Antennas, Measurement Errors.

II. INTRODUCTION

"Mismatch between the receiver (mixer or frequency converter and cables, adapters, etc.) and the test antenna can be a significant cause of antenna gain measurement inaccuracy" [1]. To illustrate this point, consider the flowgraph for the connection between an antenna and receiver (Figure 1). In traditional gain measurements, the gain of the antenna is calculated as

$$G_{AUT}[\text{Measured}] = \frac{P_R[\text{AUT}] * G_S}{P_R[\text{Standard}]}$$

Equation 1

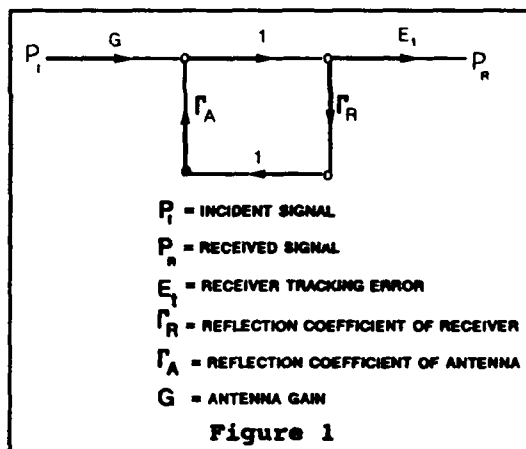
where

G_{AUT}, G_S

are the gain of the antenna under test and the known gain of the standard antenna, and

$P_R(x)$

is the signal level measured by the receiver with antenna 'x' attached.



From the flow graph, we observe that

$$\frac{P_R}{P_i} = \frac{G \cdot E_t}{1 - \Gamma_A \Gamma_R}$$

Equation 2

Substituting Equation 2 for both test and standard antennas in Equation 1, we find that

$$G_{AUT}[\text{Measured}] = G_{AUT} \left(\frac{1 - \Gamma_{GS} \Gamma_R}{1 - \Gamma_{AUT} \Gamma_R} \right)$$

Equation 3

where

$\Gamma_{GS}, \Gamma_{AUT}$
$$\text{Error (dB)} = 20 \log(1 \pm \Gamma_{AUT} \Gamma_R) + 20 \log(1 \pm \Gamma_s \Gamma_R)$$

Equation 4

III. REDUCING MISMATCH ERRORS

This paper presents a practical calibration technique which mathematically reduces the effect of mismatch without reducing sensitivity. In addition, this calibration provides a method for accurately measuring the return loss of the test antenna while it is mounted in the test range.

IV. A REFLECTOMETER CALIBRATION

$$G_{AUT} = G_S \cdot \frac{P_R(AUT) (1 - \Gamma_{AUT} \Gamma_R)}{P_R(Standard) (1 - \Gamma_{GS} \Gamma_R)}$$

Equation 5

Figure 2 is a schematic diagram of a Compact Antenna Test Range. The diagram shows a Transmit Antenna connected to a Switch, which is connected to a Reflectometer. The Reflectometer is connected to a Test Antenna. The Reflectometer also has a connection to a 60300A unit, which is connected to a 60301 unit. The 60301 unit is connected to a 6300 LCO unit. A 7015A Co-Converter Box is also connected to the 60301 unit. The entire setup is within a Compact Antenna Test Range, which is surrounded by Range Reflectors.

Figure 2

When the signal from the microwave source is switched to the transmit antenna, this system makes standard antenna transmission measurements, with only the slight loss of the coupler main arm and the isolator attenuating the signal from the test antenna. In this case, the flow graph in Figure 1 applies.

559

theory for making accurate complex reflection co-efficient measurements with this type of device is well known [5]. The error model is shown in Figure 3. The AUT is assumed to be attached to port 2 at $z-z'$. E_s is the match of the receiver (source match), E_t is the reflection tracking term, E_d is the directivity error, b_1 is the received signal, and a_1 is the incident signal for measuring return loss.

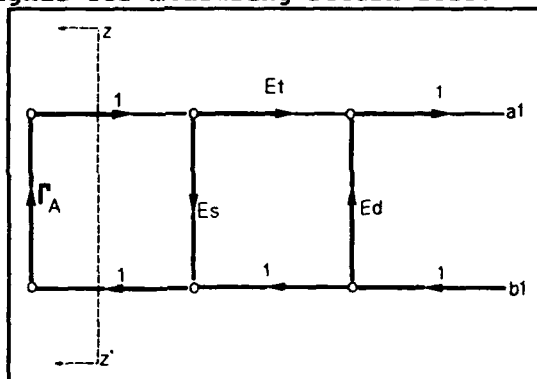


Figure 3

The HP 8530A microwave receiver includes as a calibration option the three term error model shown in figure three, as well as support for various standard connector family open-short-load calibration kits. By measuring three standards of known reflection coefficients connected at port ($z-z$), the three terms of the model are determined by the 8530A and stored internally as a calibration file. This calibration may then be used to measure the complex reflection coefficients of both the gain standard and the antenna under test. In addition, the test port input match term E_s , determined by the 8530 as part of the calibration, is equivalent to the receiver reflection co-efficient in Equation 5. Thus all three reflection coefficients required to solve Equation 5 can be readily determined.

V. PROCEDURE

Error corrected input match and gain measurements have been performed using the described measurement system at Hewlett Packard's compact test range according to the following computer aided test procedure.

1. Perform One Port Calibration

The switch (refer to Figure 2) is set

to drive the reflectometer mounted on the test antenna mount. A one port open-short-load calibration is performed at the port that will later be connected to the standard and test antennas, following the procedure in the 8530A operating manual.

2. Measure Reflection Coefficient of Standard Antenna

An antenna of known gain is mounted and connected to the calibrated port. The corrected input impedance is measured using the calibration of step one. This value is stored by the computer.

3. Perform Standard Gain Horn Calibration

The switch is now set to drive the transmit antenna, and the standard horn is brought to the boresight angle, without disconnecting it from the test port. A standard gain calibration is performed using the 8530A far field calibration. Note that the gain data for the standard antenna must be loaded into the 8530A prior to performing the calibration. The normalization factor computed by the 8530A is entered into the computer.

The source match error term from the one port calibration, the tracking term from the gain calibration, and the measured match of the standard antenna are used to find the intermediate result:

$$E_{t_{\text{corrected}}} = (1 - \Gamma_{GS} \Gamma_R) E_t$$

Equation 6

This result is stored for later use. The standard antenna is removed.

4. Measure Reflection Coefficient of the Antenna Under Test

The AUT is mounted and connected to the calibrated test port. The switch is set to drive the reflectometer, and the reflection co-efficient of the AUT is measured, using the calibration from step 1. Using this value and the intermediate result of step 2, the corrected gain normalization is calculated:

$$E_{t_{AUT} \text{ corrected}} = \frac{E_{t_{GS} \text{ corrected}}}{(1 - \Gamma_{AUT} \Gamma_R)}$$

Equation 7

The result is stored as a normalization calibration in the 8530A. This calibration, applied to measured data, performs the operation of equation 6 above and therefore removes the effect of mismatch error from the antenna gain measurements.

5. Measure Gain of AUT

The AUT gain corrected for mismatch may now be determined using the entire feature set of the hp 8530A. The calibration is valid for the AUT gain and pattern measurements at each frequency at which the calibration was performed. When another AUT is to be measured, Step 4, measurement of the reflection coefficient, is repeated.

VI. COMPACT RANGE MEASUREMENTS

In order to verify the mathematical results in an actual test environment, experimental measurements were performed in the Hewlett Packard compact antenna test range. In these measurements, the gain of a poorly matched test antenna was measured using both the standard gain calibration and the reflectometer calibration. In addition, a second standard calibration and measurement were performed with a well matched 20 dB attenuator attached to the test port. Since both the 20 dB attenuator and the reflectometer calibration should have the effect of reducing the mismatch error in the gain measurement, we expect the results of these two measurements to be similar. Figure 2 shows the measurement system block diagram used for these measurements. A WR-137 standard gain horn was used as the gain calibration standard, and HP 85052 slotless 3.5 mm calibration standards were used to perform the one port calibration. In order to provide a test antenna with a high reflection co-efficient, a WR-187 horn antenna mounted to a waveguide to coaxial adapter whose match had been purposely de-tuned was used. Antenna return loss measurements were performed on antennas mounted to the test mount using the calibrated reflectometer.

All measurements were made over the 5.4 GHz to 8.2 GHz frequency range of the gain standard.

MEASUREMENT RESULTS:

1. Return Loss Calibration

The level of residual errors in the one port calibration of the reflectometer were examined by measuring an offset short and an offset termination following the calibration. Maximum Residual source match errors of approximately -50 dB and residual directivity errors of less than -40 dB were observed. These residual error levels allow measurement of a 20 dB return loss with a maximum error of +/- 1 dB, and were deemed sufficient for the antennas being measured.

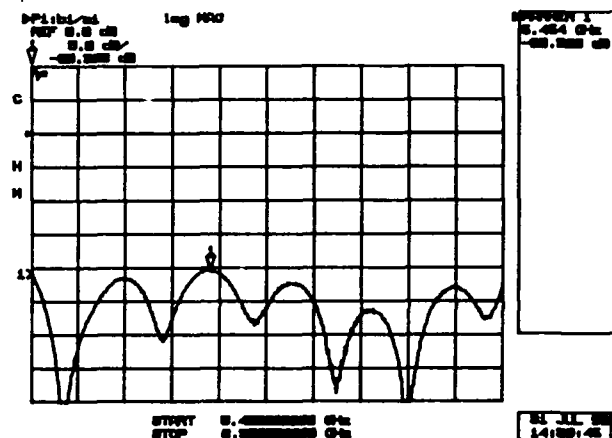


Figure 4

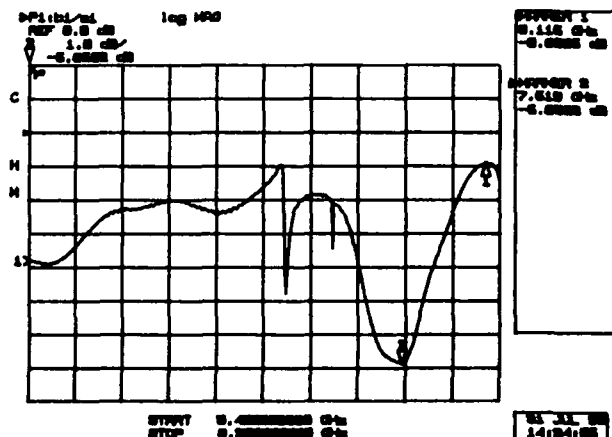


Figure 5

The measured return loss of the standard and test antennas are presented in figures 4 and 5. The gain standard is well matched, with a return loss of better than 20 dB at all

frequencies, while test antenna return loss varies between 7 dB and 1 dB

2. Gain Measurements

The gain measurements are shown in Figures 6 and 7. Figure 6 compares the gain measurements using the standard calibration with the results of the reflectometer calibration. Both measurements were taken with the same test system, and the uncorrected gain data is identical. Note the decrease in the ripple using the reflectometer calibration. This ripple represents measurement errors due to mismatch which are greater than 0.5 dB at some frequencies. In Figure 7, the results obtained by reducing the mismatch error through attenuation and calibration are seen to agree quite well with one another, as predicted.

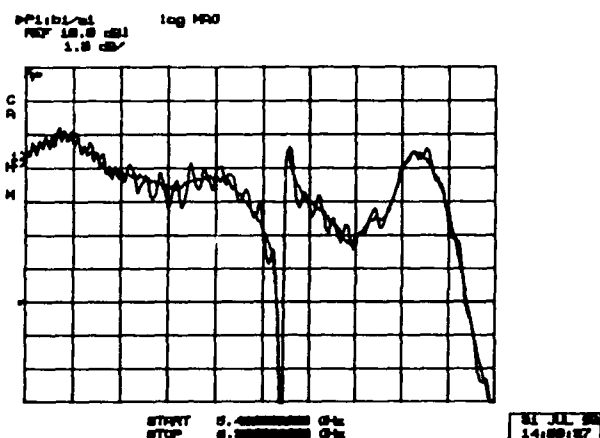


Figure 6

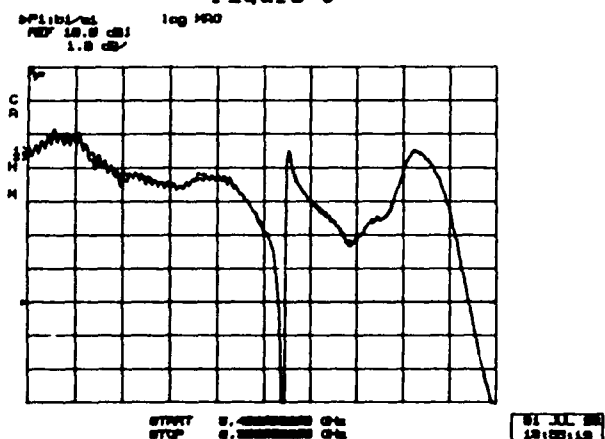


Figure 7

3. Time Domain Response

The HP 8530A time domain feature provides additional insight into the performance of the calibration. Figure

8 compares the result of the standard gain calibration to that of the reflectometer calibration, and Figure 9 shows the results of both the reflectometer calibration and the attenuated standard calibration in time domain.

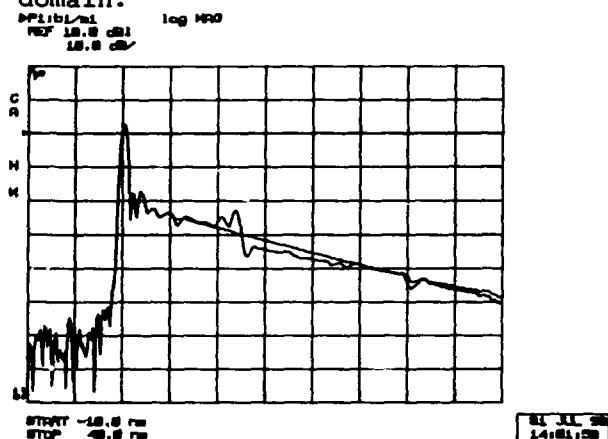


Figure 8

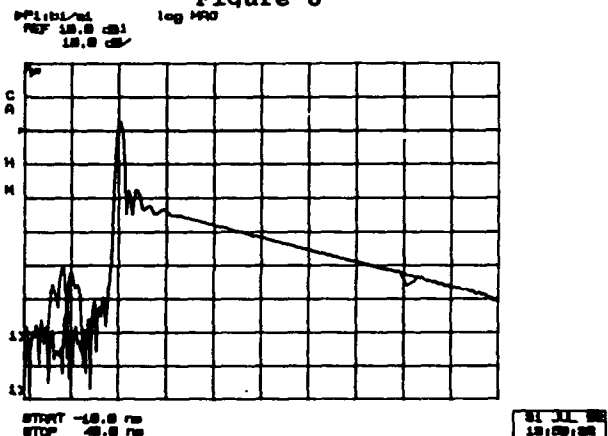


Figure 9

In all three measurements, the large peak at 0 ns represents the direct path signal, and all peaks to the right of this are the result of signals paths which are delayed relative to the direct path response. Since the mismatch error signal is that part of the received signal which is reflected by the receiver input mismatch back to the test antenna, and then back again to the receiver (see Figure 1), it is delayed relative to the direct response and thus occurs to the right of 0 ns. In Figure 8, this delayed mismatch response is observed to occur between 20 ns and 30 ns after the direct path signal in the data corrected with the standard gain calibration, but is not observed in either of the two 'mismatch corrected' measurements. In fact, these measurements are virtually identical

for time > 0 ns. This observation strongly supports the conclusion that the reflectometer calibration has successfully reduced the mismatch error in this measurement.

VII. DISCUSSION

1. Calibration Stability

In order for the calibration to be a useful measurement technique, it must remain stable for the duration of the measurement. In the case of antenna pattern measurements, this implies that the changes in phase and amplitude of the rotary joint and/or cables must not degrade the ability of the calibration to reduce mismatch error. Fortunately, once completed, the mismatch error correction will remain valid as long as the product

$$\Gamma_R \Gamma_{AUT}$$

Equation 8

remains constant. Therefore, if the test antenna and reflectometer are properly connected together, the mismatch correction should remain stable independent of phase changes in range cables. This has been demonstrated experimentally on the compact antenna range. Another related issue is the stability of the system during the calibration. If the receiver error terms are altered by cable movement between measurements of the gain standard and the antenna under test, the calibration will be degraded. Changes in the long range cables due to temperature variation and in cables connected to the test port due to bending are the biggest sources of error. It is important, therefore, to limit cable flexure to a minimum and to maintain a stable temperature during the calibration. The use of precision connectors is recommended as well to ensure good connector repeatability.

2. Extension to Outdoor Ranges

In principle, the reflectometer technique could be used with equal validity on an outdoor test range if a signal source were available to drive the reflectometer.

VIII CONCLUSION

A practical calibration technique based upon the use of a calibrated reflectometer for measuring antenna

return loss and gain has been presented, along with results indicating a clear reduction in mismatch error without sacrifice of sensitivity. This technique uses existing calibration capability of the HP 8530A in conjunction with an external controller to facilitate the calibration.

REFERENCES:

- [1] John W. Boyles, "Determining Measurement Accuracy in Antenna Test," Proc. AMTA 1990, p 6-43.
- [2] Greg McCarter, "Design Your Measurement System For Optimum Throughput," Proc. AMTA 1990, p 14-37.
- [3] Ronald W. Bowman, "Field Strength Above 1 GHz: Measurement Procedures for Standard Antennas," Proceedings of the IEEE, Vol. 55, No. 6, June, 1967.
- [4] E.B. Larsen et. al, "Calibration of Antenna Factor at a Ground Screen Field Site Using an Automatic Network Analyzer," NSEMC 1989, pp. 19-24.
- [5] Doug Rytting "An Analysis of Vector Measurement Accuracy Enhancement Techniques", RF & Microwave Measurement Symposium and Exhibition, March, 1982.

A NOVEL X-BAND, CIRCULARLY POLARIZED FEED
FOR THE I-30 RADAR ANTENNA SYSTEM

F. A. Lauriente and W. E. Hord
Microwave Applications Group
Santa Maria, CA

ABSTRACT

This paper describes a novel application of septum polarizers in the realization of a feed system for I-30 Radar Antenna System. The I-30 antenna is an electronically steerable phased array employing nonreciprocal, circularly polarized ferrite phase shifters in a transmission (bootlace) lens arrangement with a space feed. The feed is required to provide monopulse operation with either sense of circular polarization on receive, as well as the duplexing function between the transmit and receive modes of operation with the sense of the transmit polarization being opposite to that on receive.

The feed uses a square multi-mode pyramidal horn. Flare angle changes are utilized for the generation of higher order waveguide modes to obtain equal E- and H-plane primary patterns for efficient illumination of the phased lens commensurate with low spillover loss. A cruciform arrangement of four stepped septum polarizers excites the throat of the horn. The transmit ports of the polarizers are combined with magic tees to form a transmit sum port, and the receive ports are combined in an off-the-shelf monopulse comparator which forms the receive sum port and two orthogonal difference ports.

1.0 Introduction

The I-30 radar system is a "sophisticated threat" radar simulator. The radar consists of a two-axis electronically scanned phased array antenna, a coherent high-power transmitter, superheterodyne 3-channel receiver and digital signal processor all operating under control of a Concurrent model 7500 radar control computer. The radar features extremely fast beam scan rates, monopulse tracking, real-time simulation capabilities and high throughput digital signal processing.

The I-30 antenna is a transmission (bootlace) lens illuminated by a space feed. Nonreciprocal ferrite phase shifters operating in a circularly polarized mode are contained between an aperture plate and a feed plate. Radiating elements are formed when dielectric transformers on each end of the ferrite phase shifters are inserted into circular cavities bored in the feed and aperture plates. Since the single-bounce target return is desired, the received circular polarization is opposite the transmitted circular polarization and commutation of the phase shifters is not required. Accordingly, the phase shifters are switched at the beam scan rate rather than at twice the radar pulse repetition frequency which minimizes power supply requirements.

2.0 Functional Description of the I-30 Antenna System

The I-30 antenna consists of a transmission lens, a three-channel monopulse feed, an r-f switch for selection of either two or three-channel monopulse operation, a beam steering computer, the power supplies necessary to power the system and a structure

for mounting and/or support of the various subsystems. A block diagram is given in Fig. 2-1, and the antenna specifications are given in Table 2-1.

Figure 2-1
System Block Diagram

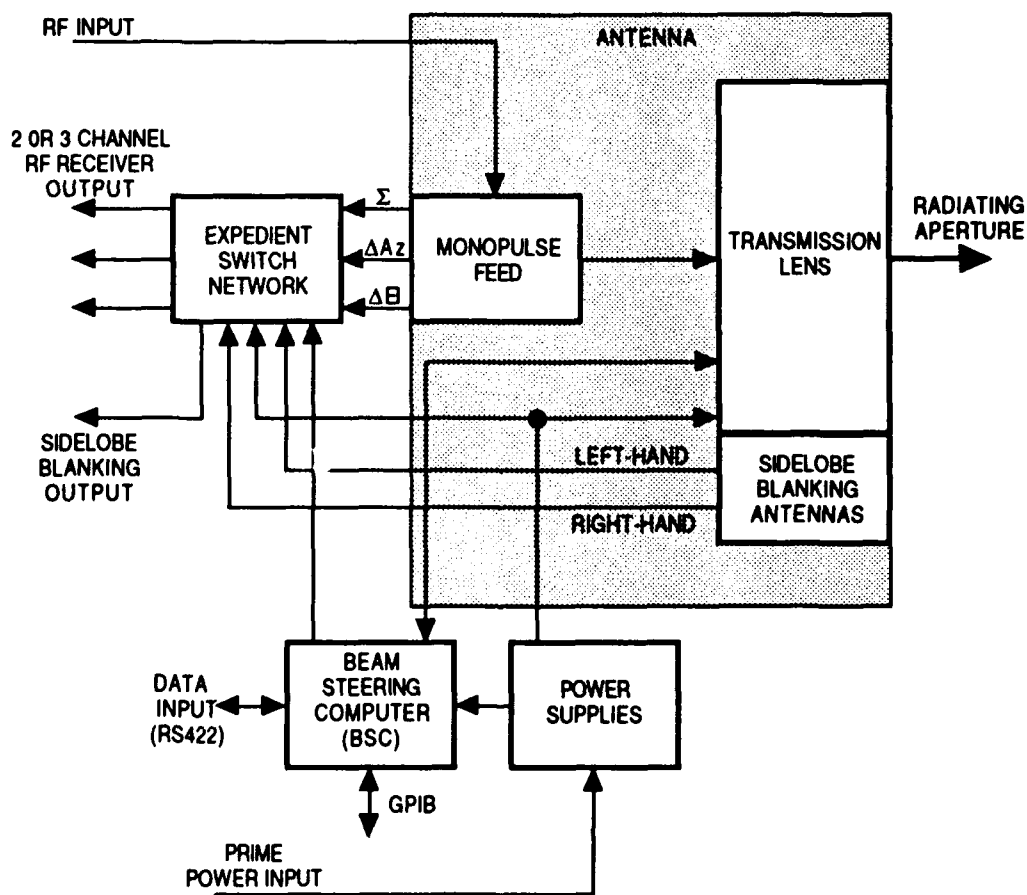


Table 2-1
I-36 Antenna Specifications

Frequency Range	8.3 GHz - 8.9 GHz
Instantaneous Bandwidth	50 MHz
Polarization	Circular
Azimuth Scan	± 20 degrees
Elevation Scan	± 22.5 degrees
VSWR	1.50:1 Max.
Gain (Broadside)	36 dB Max.
Peak Power	100kW
Average Power	8kW
Beamwidth	1.9 Degrees AZ and EL
Beam Pointing Accuracy	0.25 Milliradians Max.
Beam Resolution	0.25 Milliradians
Beam Broadening	0.3 Degrees Max.
Peak Sidelobe Level	-25 dB Max.
Beam Switching Time	100 Microseconds Max.
Load Time	500 Microseconds Max.
Operating Temperature	-15 to 46 Degrees C

The antenna is designed to radiate either sense of circular polarization and to receive the opposite sense circular polarization. All subsystems are pressurized to prevent penetration of the system by foreign matter. Built-in-test is incorporated into the design, so that the beam steering computer acts as a diagnostic/protective subsystem in addition to carrying out its normal control function. Sidelobe blanking antennas—one for each sense circular polarization—are incorporated into the radiating aperture. The final configuration is shown in Fig. 2-2, and a list of features is given in Table 2-2.

Figure 2-2
I-30 Antenna System Configuration

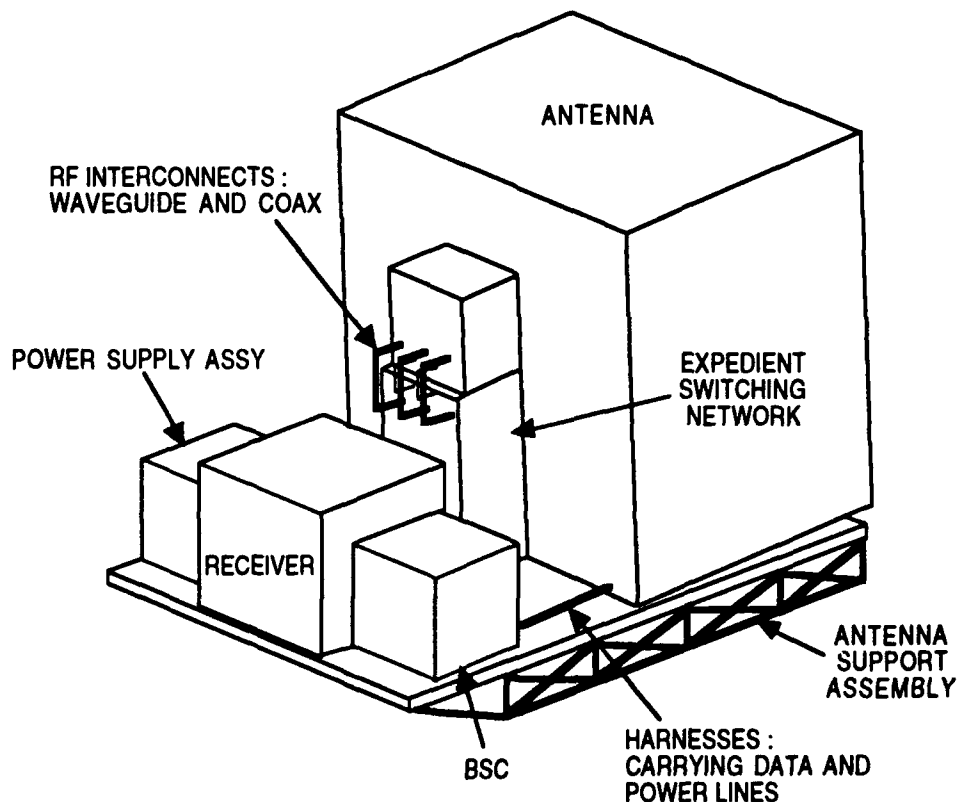


Table 2-2

I-30 Antenna Features

- 2101 Radiating Elements Located on an Isosceles Triangle Grid
- 50 Inch Circular Aperture
- 50 Inch Focal Length
- Circular Polarization
 - Transmit Right Hand, Receive Left-Hand
 - Or Transmit Left-Hand, Receive Right-Hand
- 3-Channel Monopulse Comparator
- Latching Ferrite Phase Shifters
- Sidelobe Blanking Antennas
- Built-in-Test
- Pressurized Assembly
- Heat Pipes for Passive Array Cooling

THE OCCURRENCE OF ANY OF THE CONDITIONS WILL CAUSE THE
BSC TO SET THE FAULT LINE AND DISABLE ANTENNA SWITCHING

- Loss of 208 Volt A-C Power Source
- Loss of 5 Volt Power Supply
- Loss of 15 Volt Power Supply
- Loss of ± 12 Volt Power Supply
- Array Overtemperature Condition
- BSC Overtemperature Condition
- Self Test Failure

MONITORED STATUS INCLUDE THE FOLLOWING

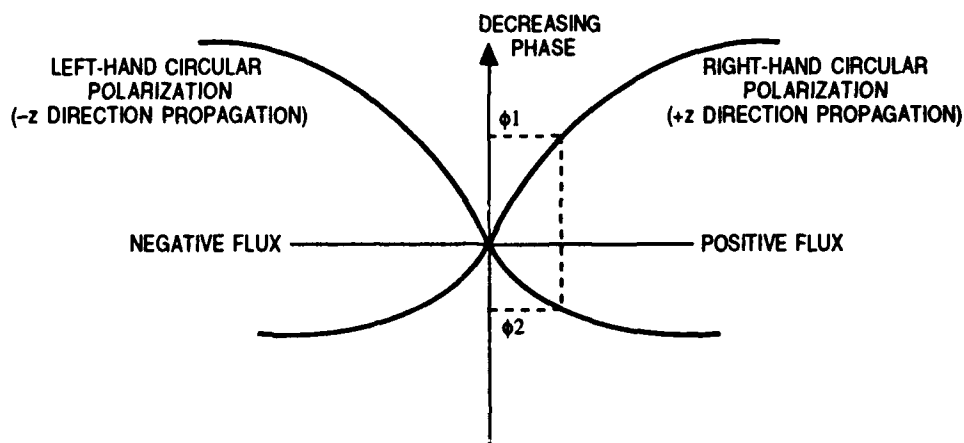
- Power Supply Status
- Cooling Fan Status
- Polarization Status
- I/O Status
- Bit Status
- Self Test Status

Antenna efficiency factors are shown in Table 2-3 leading to a predicted gain of 37 dB.

Table 2-3 Efficiency Factors D = 50.0 IN., N = 2101			
	<i>EFFICIENCY FACTOR (dB)</i>		
	<i>FREQUENCY (GHz)</i>		
<u>Parameter</u>	<u>8.30</u>	<u>8.60</u>	<u>8.90</u>
Illumination Taper	-0.46	-0.52	-0.59
Spillover	-0.51	-0.44	-0.38
Phase Errors ($\sigma = 15^\circ$)	-0.29	-0.29	-0.29
Amplitude Errors ($\sigma = 0.06$)	-0.02	-0.02	-0.02
Cross-Polarization	-0.22	-0.16	-0.22
Diffraction	-0.09	-0.11	-0.20
Phase Shifter Loss (dB)	-1.3	-1.2	-1.2
Phase Shifter Quantization (7-bit)	-0.0	-0.0	-0.0
Feed Loss (dB)	-0.30	-0.30	-0.30
Mismatch Loss (1.75:1)	-0.34	-0.34	-0.34
Total Efficiency (dB)	-3.53	-3.38	-3.54
Maximum Aperture Gain (dB)	40.61	40.90	41.17
Net Antenna Gain			
Boresight (dB)	37.08	37.52	37.63
3 dB Beam Width			
Boresight (Degrees)	1.93	1.88	1.84

The phase characteristics of the nonreciprocal ferrite phase shifters are shown in Fig. 2-3. Note the phase characteristics are symmetrical about the vertical axis. If the phase required of a particular element for antenna collimation is ϕ_1 , the diagram suggests the antenna will be collimated for right-hand circularly polarized waves propagating in the +z direction and left-hand circularly polarized waves propagating in the -z direction. This implies that the antenna is self-duplexing if a means is provided at the feed to separate orthogonal senses of circular polarization.

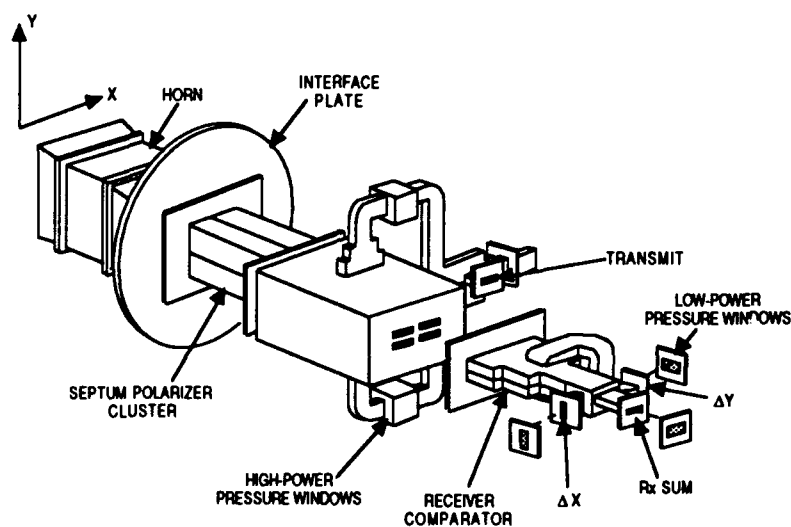
Figure 2-3
Phase Characteristics of I-30 Phase Shifter



3.0 Feed Design

The RF feed assembly is shown pictorially in Fig. 3-1. The feed must efficiently illuminate the lens, provide monopulse operation on receive and provide transmitter-receiver duplexing. Prior art¹ has used four orthomode transducers

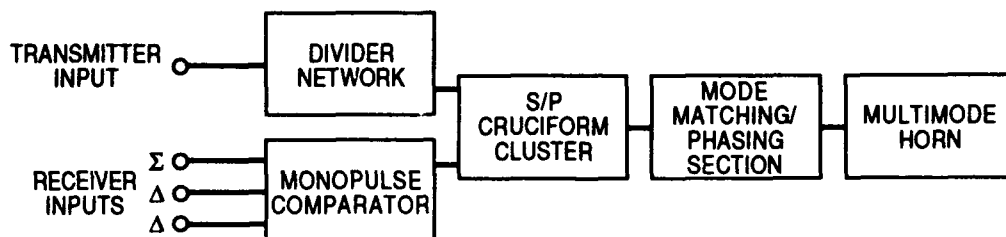
Figure 3-1
RF Feed Assembly



(OMT) in a cruciform with input ports to each OMT fed in phase quadrature to obtain circular polarization. This results in a feed with a costly and complicated external waveguide network. The novelty of the new feed design is in the use of four septum polarizers in the cruciform for the generation of CP. As in prior designs, accommodation of high power is achieved through the factor of four reduction in the input transmitter power by use of the cruciform. Physically the feed is pressurized, and the septum polarizer cluster is removable to allow for

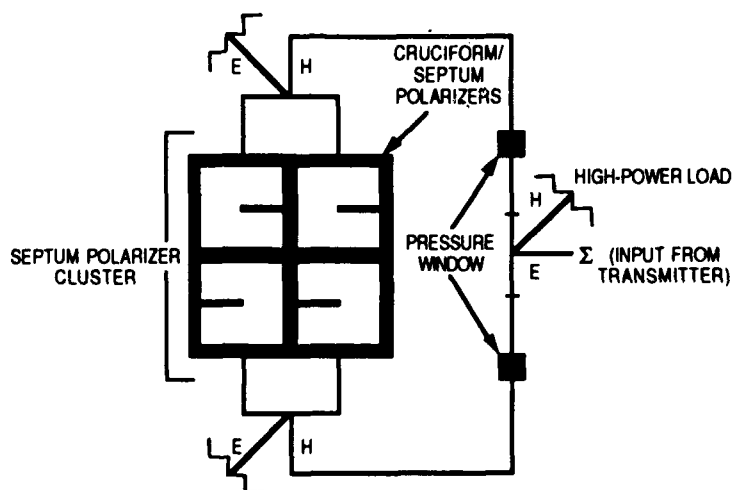
operation with the opposite sense circular polarization. A block diagram of the feed is shown in Fig. 3-2.

Figure 3-2
Multimode Feed Block Diagram



The waveguide network which combines the Tx ports of the septum polarizers to form a single transmitter input is connected as shown in Fig. 3-3. Since the average input power is greater than the average power capacity of off-the-shelf

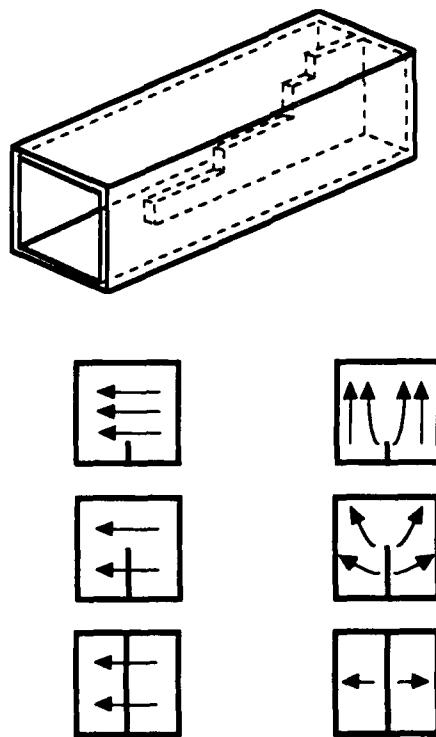
Figure 3-3
Transmit Network Schematic



pressure windows, the windows are inserted into the lines after the first Magic Tee. The power in each line is again divided into equal parts with folded H-plane tees and this results in four equiphase signals for excitation of the septum polarizer cluster on transmit.

The septum polarizer is a four-port waveguide device, and the stepped septum version described by Chen and Tsandoulas² is shown in Fig. 3-4. The square waveguide at one end constitutes two ports since it supports two orthogonal

Figure 3-4
Septum Polarizer



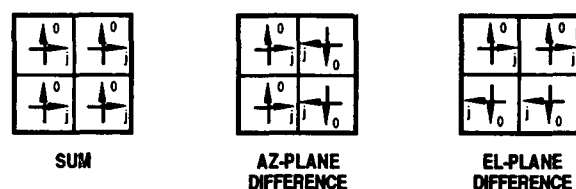
modes. The stepped septum divides the square waveguide into two rectangular waveguides sharing a common wall. If one of the rectangular waveguide ports is excited, the signal is transformed into a circularly polarized signal at the square waveguide. Exciting the other rectangular port results in the opposite sense of circular polarization at the square port.

The operation of the septum polarizer is also illustrated in Fig. 3-4. Assume the square port is excited with an electric field parallel to the septum. This signal transforms into two odd-mode signals at the rectangular ports. An electric field perpendicular to the septum transforms into two even-mode signals as shown. If both input components exist simultaneously, cancellation at one or the other rectangular port occurs if the amplitudes are identical and the phase difference is either zero degrees or 180 degrees. Since the septum is designed to provide 90 degrees of differential phase shift, isolation occurs for circularly polarized input signals.

Monopulse operation on receive is accomplished by means of a standard monopulse comparator connected to the receive ports of the septum polarizer cluster. This is a commercially available off-the-shelf item and provides the excitations to the Rx ports of septum polarizers necessary to develop sum and difference patterns. The excitations and the resulting modes generated in the mode matching section are shown in Fig. 3-5. Since this is a circularly polarized application, six different waveguide modes are launched at the cruciform section. The circularly polarized azimuth difference patterns are formed by the

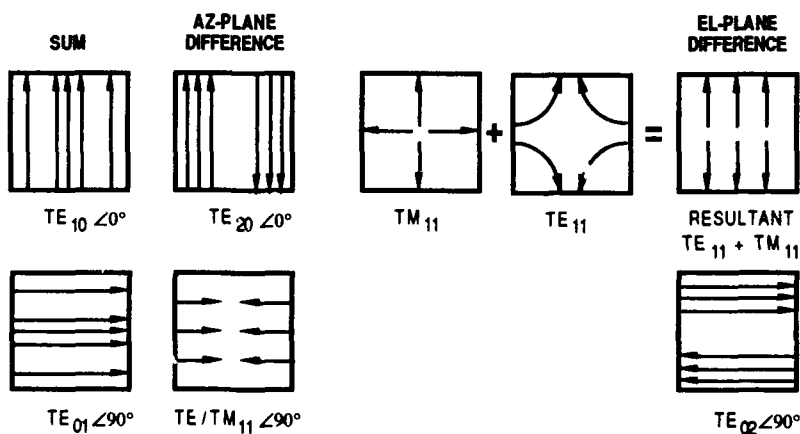
TE_{20} mode and the horizontally polarized hybrid TE/TM_{11} mode phased 90 degrees in time, while the elevation difference pattern results from the vertically

Figure 3-5
Mode Excitation of Multimode Feed
by Monopulse Network



EXCITATION OF CRUCIFORM SECTION

PROPAGATING MODES LAUNCHED IN MULTIMODE MATCHING SECTION



polarized TE/TM_{11} hybrid mode and the TE_{02} mode similarly phased. Since these modes have different phase velocities, a difference mode phasing section is used at the throat of the multi-mode horn to obtain the necessary time quadrature at the aperture of the feed horn.

Efficient illumination of the phased lens requires that the feed horn patterns be rotationally symmetric with low sidelobes. This requires that the dominant TE_{10} uniform E-plane field distribution in the horn be altered to more nearly match that of the H-plane which possesses a cosine distribution. This is accomplished with a square multimode horn which uses flare angle changes³ as the mechanism for generating higher order modes needed for control of the primary sum patterns. The horn geometry is shown in Fig. 3-6 and as can be seen it simple and easy to fabricate. Rotationally symmetric patterns over 10 to 12 percent bandwidths can be obtained with this technique.

Figure 3-6
Multi-Flare Multimode Horn Design

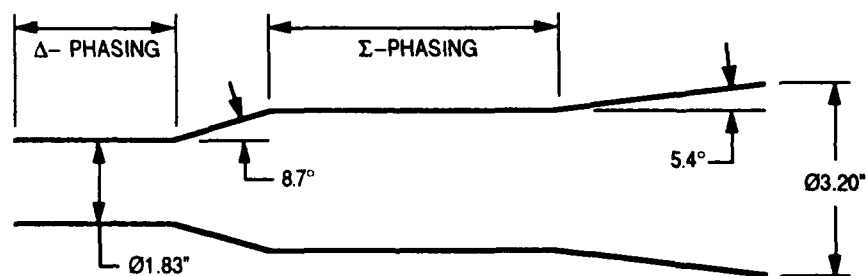
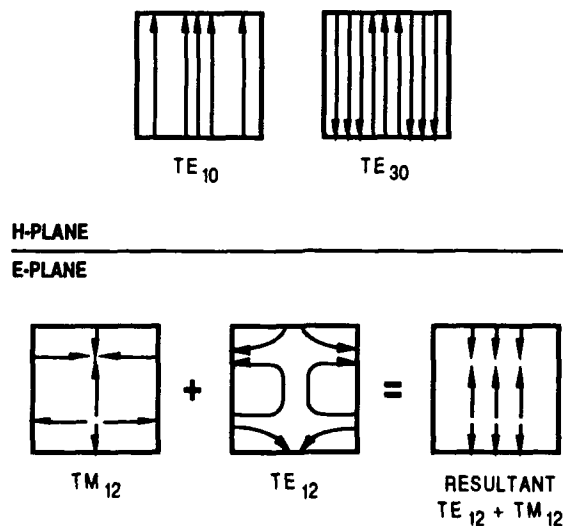


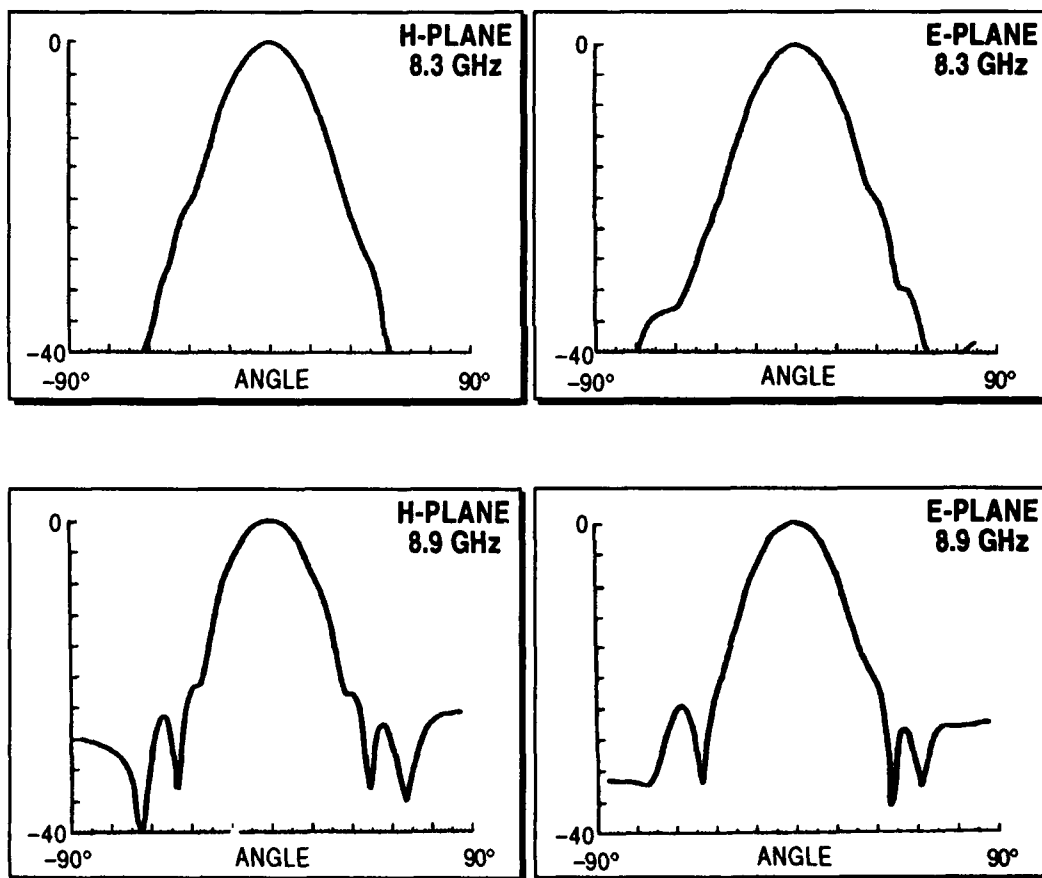
Figure 3-7 shows the modes involved in tapering the E-plane field distribution to nearly match that of the H-plane. Each flare angle change excites symmetric higher order modes with amplitude inversely related to the mode order and directly as the magnitude of the flare angle change. For equal E- and H- plane 10dB beamwidths, it is necessary that the amplitude of the TE/TM_{12} mode relative to the TE_{10} mode be about 0.66 in voltage. (A small amount of TE_{30} mode is also excited, but can be ignored for small flare angle changes.) This determines the magnitude of the flare angle. The length of the square straight section in conjunction with the overall horn length, is chosen so that the hybrid TE/TM_{12} mode arrives at the aperture of the horn in phase with the TE_{10} mode. Note that with circular polarization excitation the equivalent orthogonal waveguide modes are generated in time quadrature.

Figure 3-7
Multimode Horn Pattern Control



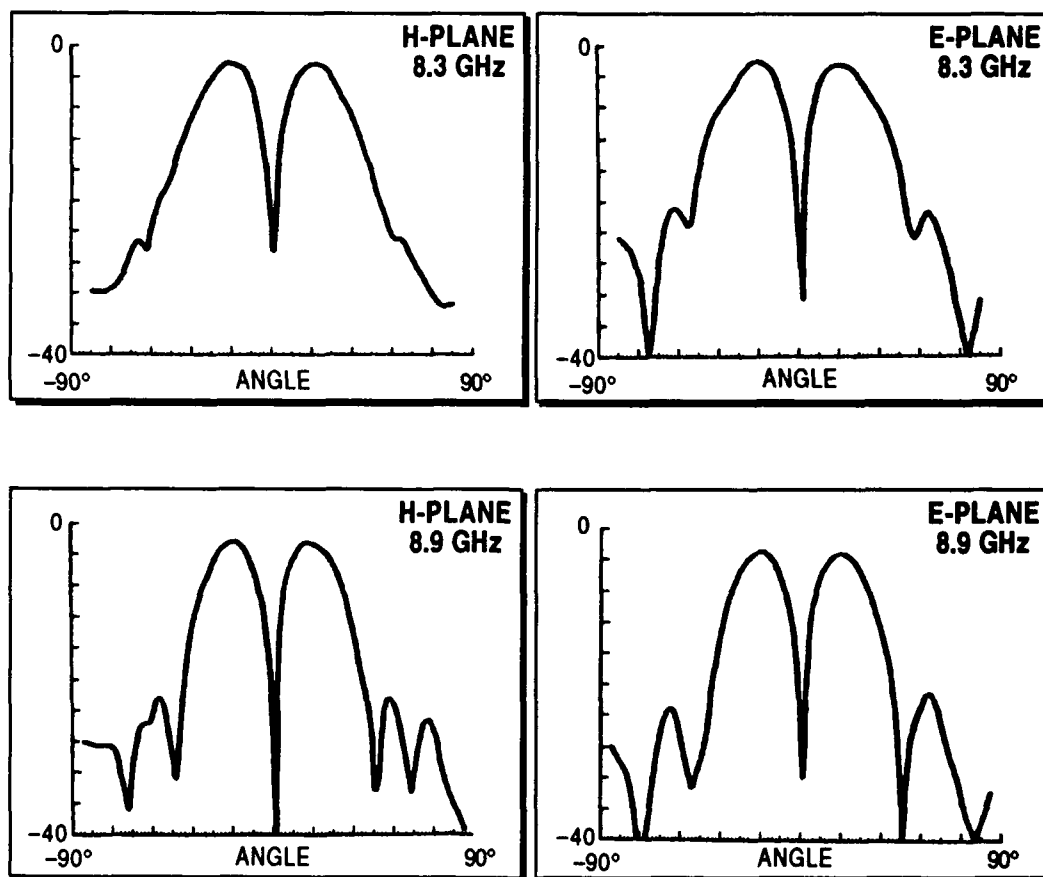
Typical CP sum patterns, measured on a near-field range for the horizontal (H) and elevation (E) planes, are shown in Fig. 3-8. The cross-polarization level over the angle subtended by the lens (27 degrees) was measured in the range of -18

Figure 3-8
RCP Feed Patterns
Receive Sum



to -24dB corresponding to axial ratios of 2.1 to 1.1dB. Difference patterns for both planes are shown in Fig. 3-9. Note that E- and H-plane refer to elevation and horizontal plane respectively.

Figure 3-9
RCP Feed Patterns
Difference



4.0 Secondary Pattern and Gain Measurements

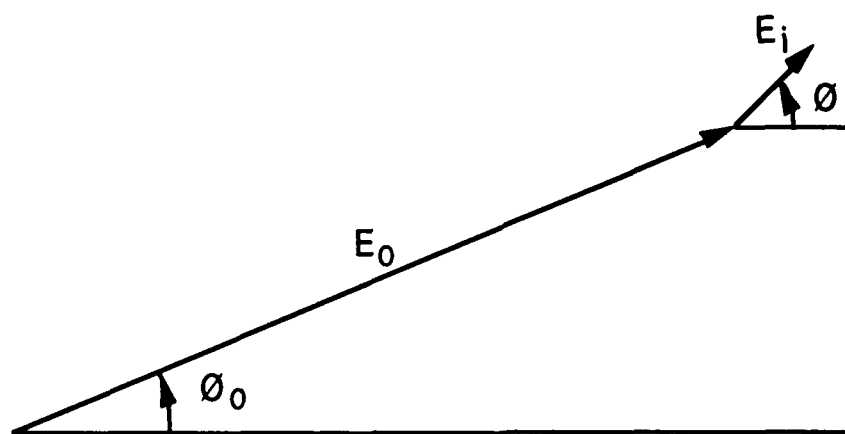
Extensive secondary pattern and gain characteristics were measured on a near-field antenna range. This facility was designed and built by MAG expressly for the secondary pattern measurements on this array. A seven foot by seven foot scanner, shown in Fig. 4-1, was purchased and interfaced to an HP 200 series computer. Software was written for control of the scanner and data acquisition and control of the HP 8510B network analyzer which served as a receiver. The measurement probe is a septum polarizer with axial ratio less than 0.4 dB. Data was taken simultaneously for both senses of CP. The NIST Planar Nearfield Software package was used for transforming near-field data to the far field.

Figure 4-1



Collimation of the array was also performed with the near-field scanner. The probe is positioned sequentially over each element in the array and its phase shifter commanded through all of its phase states while the others are at a fixed phase state. Referring to the diagram in Fig. 4-2, the fixed phasor corresponding to the phase shifters not being cycled is subtracted from the measurement and a "best fit" circle is used to obtain the phase shifter command which aligns the element phasor to the fixed phasor. This procedure is iterated several times and

Figure 4-2
Collimation Procedure
Phasor Diagram

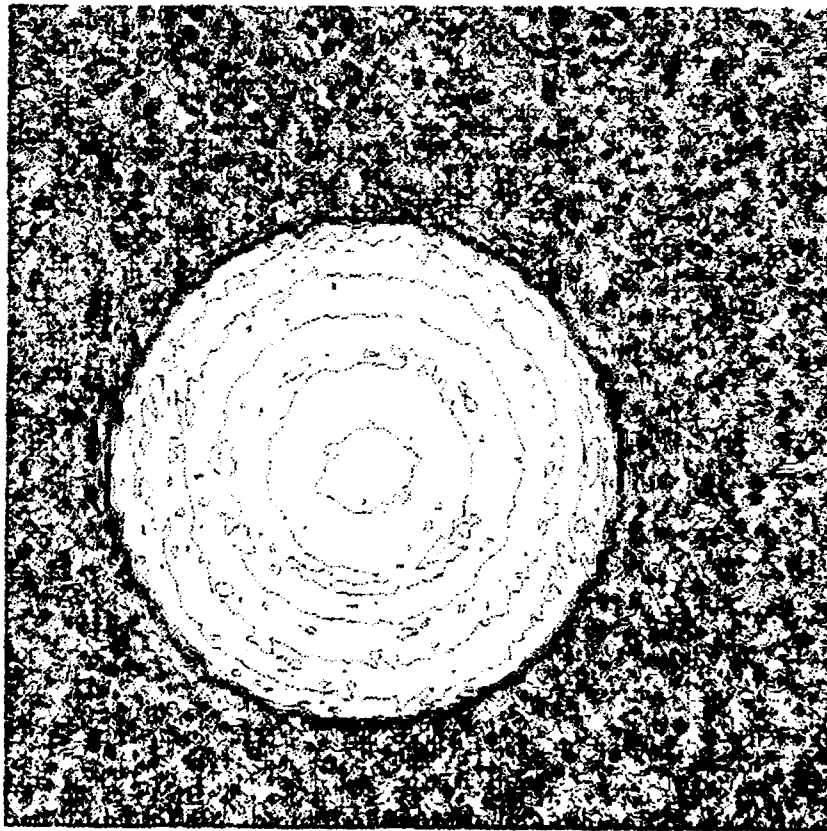


no more than three passes through all of the elements are required for collimation. Fig. 4-3 and Fig. 4-4 shows measured near-field amplitude and

Figure 4-3
Aperture Plane Amplitude
(3dB Contours)

XPD REP AP DIST 8.60 GHZ (A:0 1.0) NEAR FIELD AT 0.00 CM Y_XIN AMPLITUDE

120.0



120.0

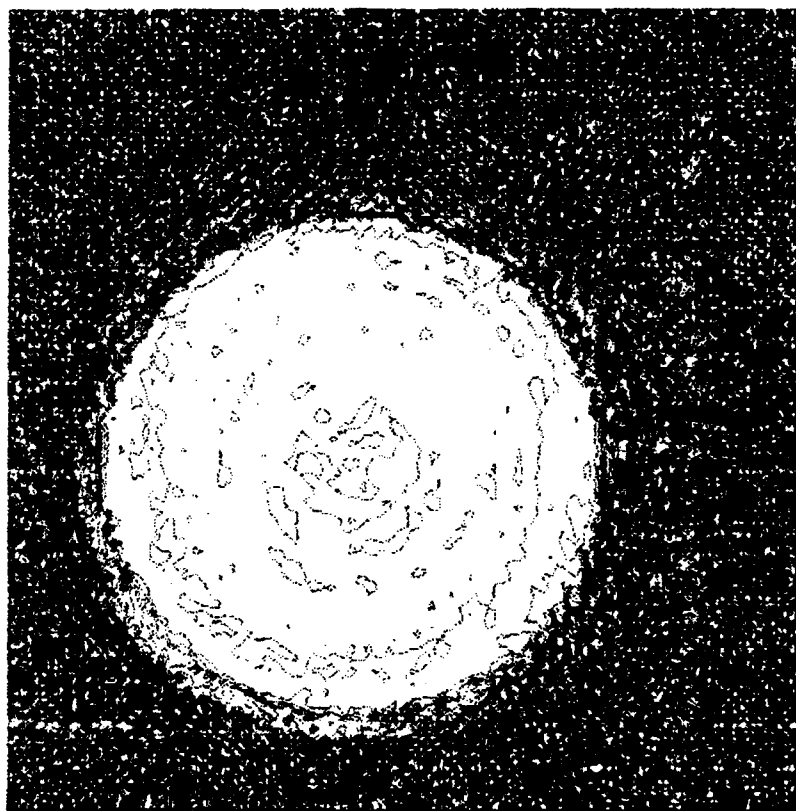
1.0

120.0

Figure 4-4
Aperture Plane Phase
(30 Degree Contours)

XPD RCP AP DIST 8.00 CMZ (A 0 E 0) NEAR FIELD AT 8.00 CM Y_XINU PHASE

128.0



1.0

1.0

128.0

phase data at center band transformed back to the aperture plane after three iterations of the collimation procedure.

Measurements were made for various beam scan angles, polarization and frequency on the receive sum and difference ports and the transmit sum port. Parameters measured included gain, beamwidth and sidelobe levels and are summarized in Table 4-1 for boresight scan. Note that the peak cross polarization in the secondary patterns is lower than that of the primary feed patterns. This arises from the fact that non-reciprocal phase shifters are used in the array.

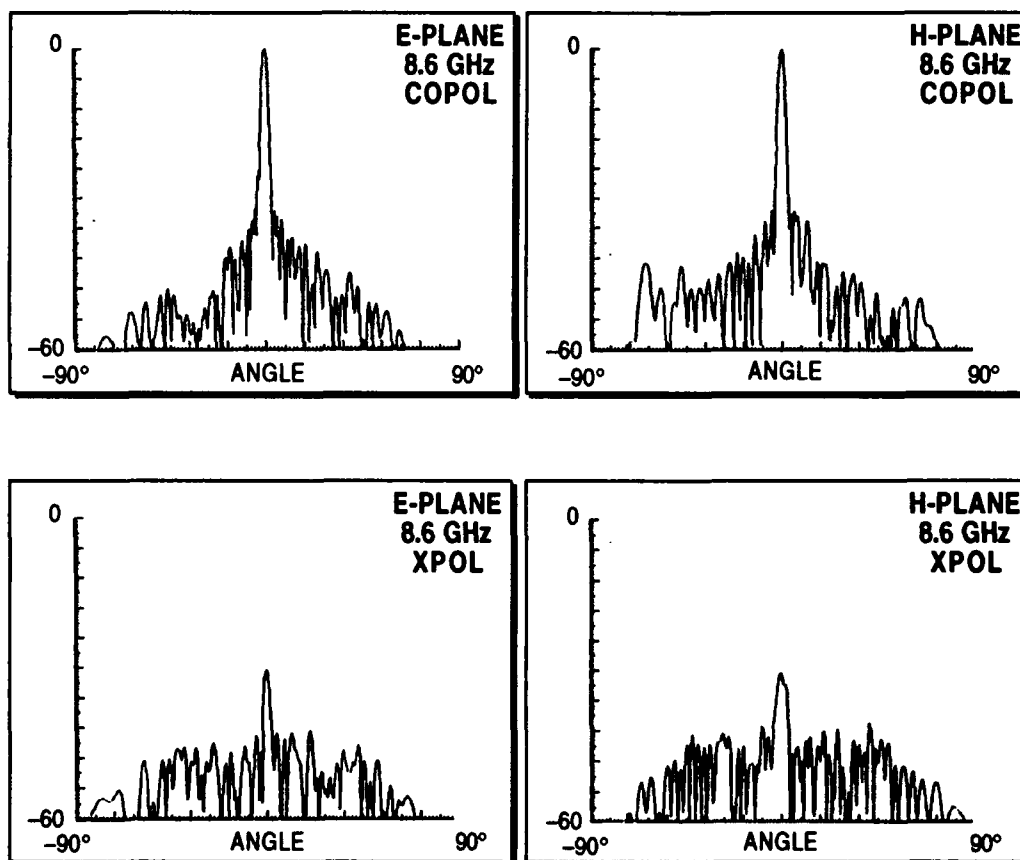
Table 4-1
Antenna Boresight Characteristics

			RHCP CLUSTER*		LHCP CLUSTER*	
Freq.	Beam Position	Parameter	Transmit Port	Receive Port	Transmit Port	Receive Port
8.3 GHz	Az:0/El:0	Gain	36.31	36.36	36.23	36.51
		AZ Beamwidth	1.97	1.98	1.96	1.96
		EL Beamwidth	1.98	1.98	1.98	1.97
		Crosspol	-30.44	-26.4	-28.86	-26.22
		AZ Sidelobe	-27.44	-23.87	-26.03	-26.17
		El Sidelobe	-30.06	-29.88	-29.28	-29.69
8.5 GHz	Az:0/El:0	Gain	36.08	35.94	35.83	36.26
		AZ Beamwidth	1.97	1.95	1.99	1.99
		EL Beamwidth	1.97	1.97	1.98	1.98
		Crosspol	-26.06	-26.4	-24.32	-26.25
		AZ Sidelobe	-31.59	-24.80	-27.86	-27.75
		El Sidelobe	-29.66	-29.19	-27.16	-26.34
8.9 GHz	Az:0/El:0	Gain	37.44	35.76	36.27	34.75
		AZ Beamwidth	1.87	1.95	1.86	1.87
		El Beamwidth	1.88	1.91	1.85	1.84
		Crosspol	-23.09	-24.07	-26.32	-26.19
		AZ Sidelobe	-19.94	-28.58	-21.72	-21.90
		EL Sidelobe	-25.43	-26.67	-21.33	-23.54
*REFERS TO SENSE OF CIRCULAR POLARIZATION TRANSMITTED BY ANTENNA						

When, for example, the phase shifters are set to collimate a RHCP beam on transmit, LHCP energy resulting from the feed cross polarization is not collimated. Therefore, this energy is radiated over a wide range of spatial angles thereby yielding lower secondary cross polarization levels.

Representative co and cross-polarized RCP transmit sum patterns are shown in Fig. 4-5 and 4-6 while the corresponding receive sum patterns (LCP) are shown

Figure 4-5
Transmit RCP Patterns
Boresight



in Fig. 4-7 and Fig. 4-8 for boresight and maximum scan condition. Figure 4-9 shows the co and cross-polarization LCP transmit patterns and Fig. 4-10 shows the receive sum patterns (RCP) for maximum scan.

Figure 4-6
Transmit RCP Patterns
Maximum Scan Condition

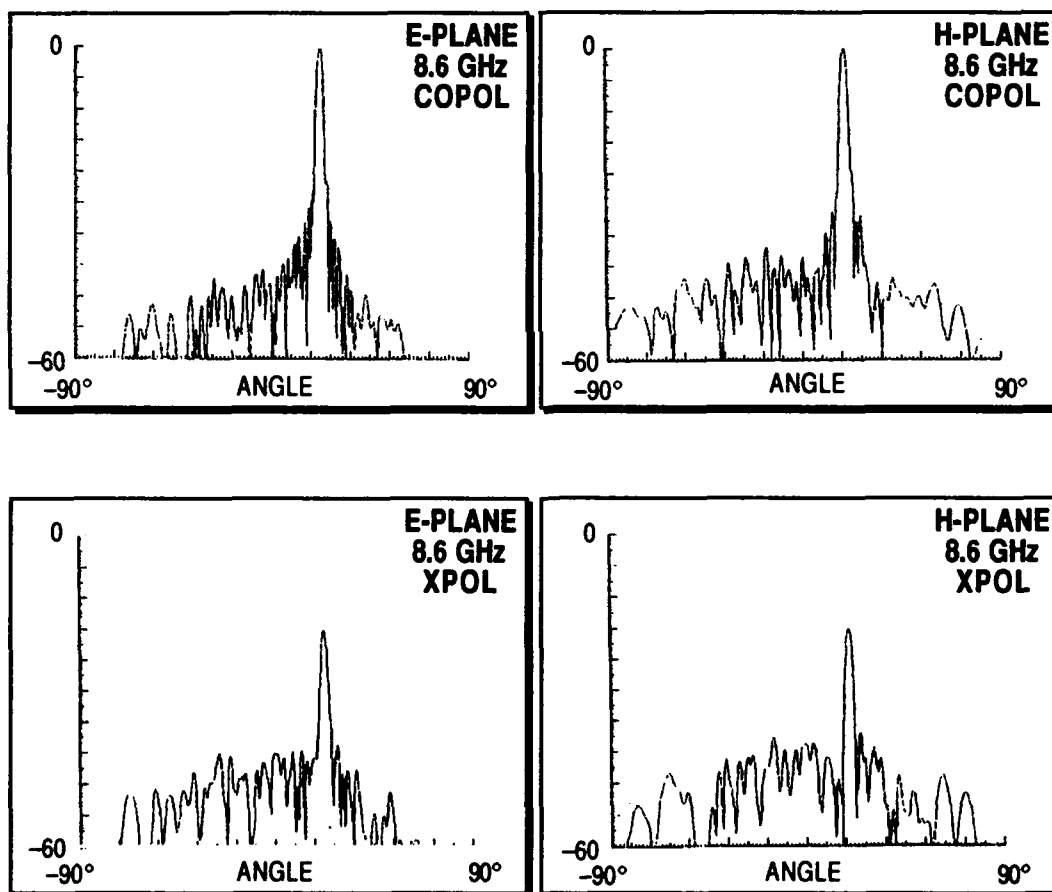


Figure 4-7
Receive LCP Patterns
Boresight

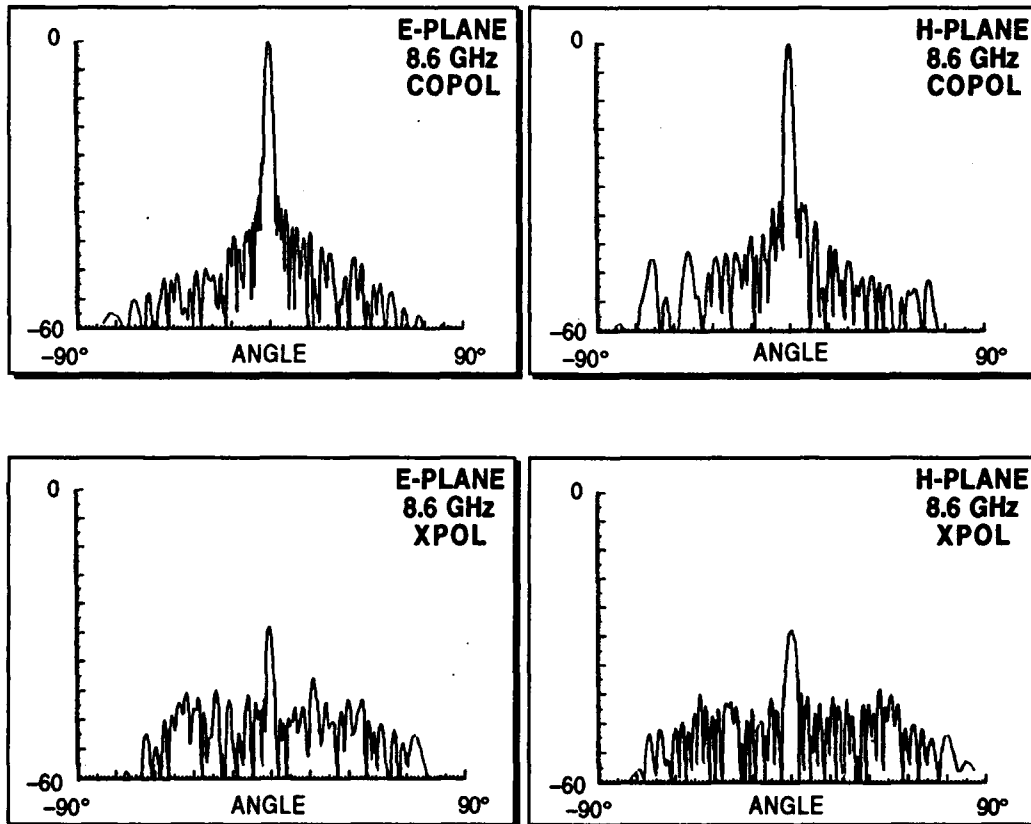


Figure 4-8
Receive LCP Patterns
Maximum Scan Condition

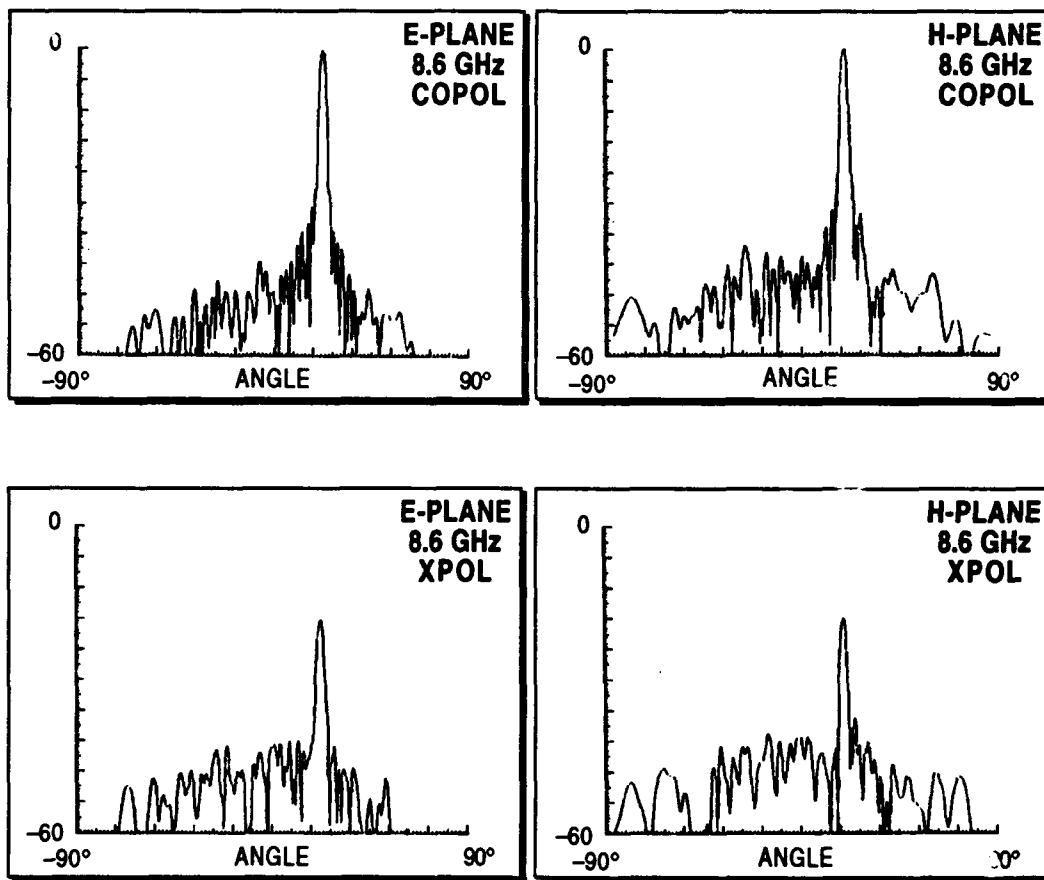


Figure 4-9
Transmit LCP Patterns
Maximum Scan Condition

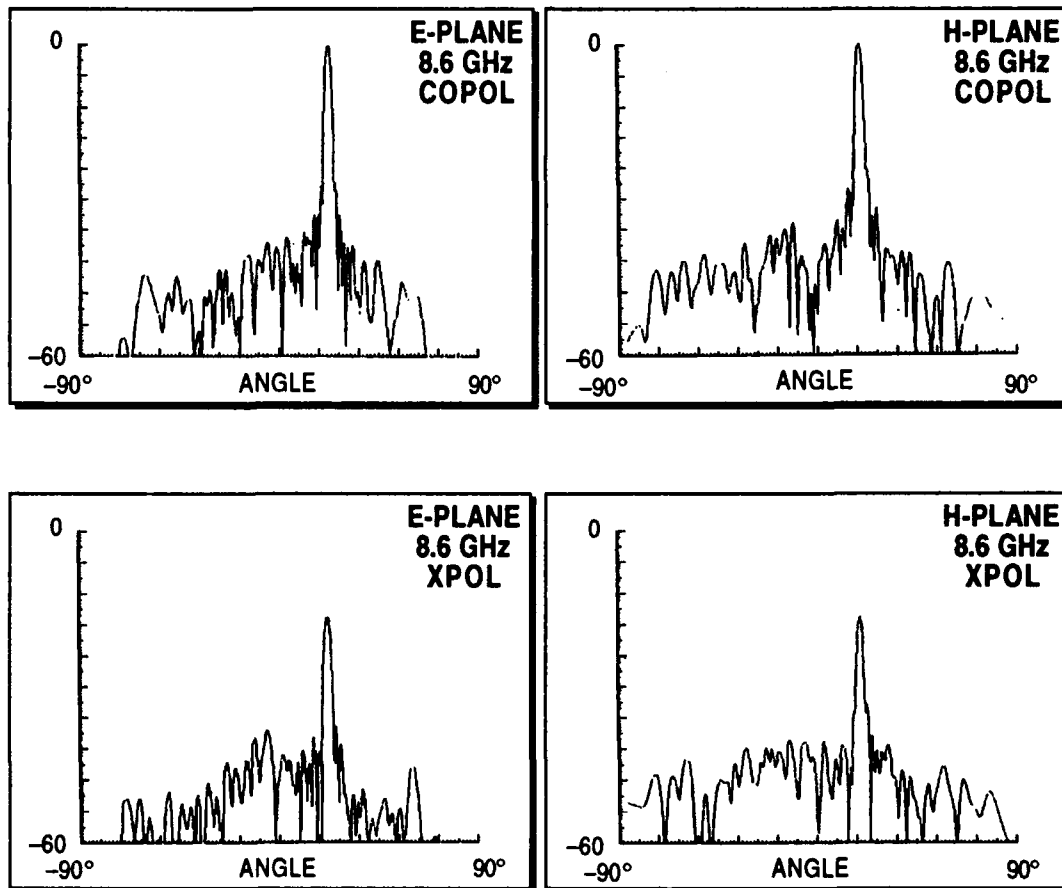
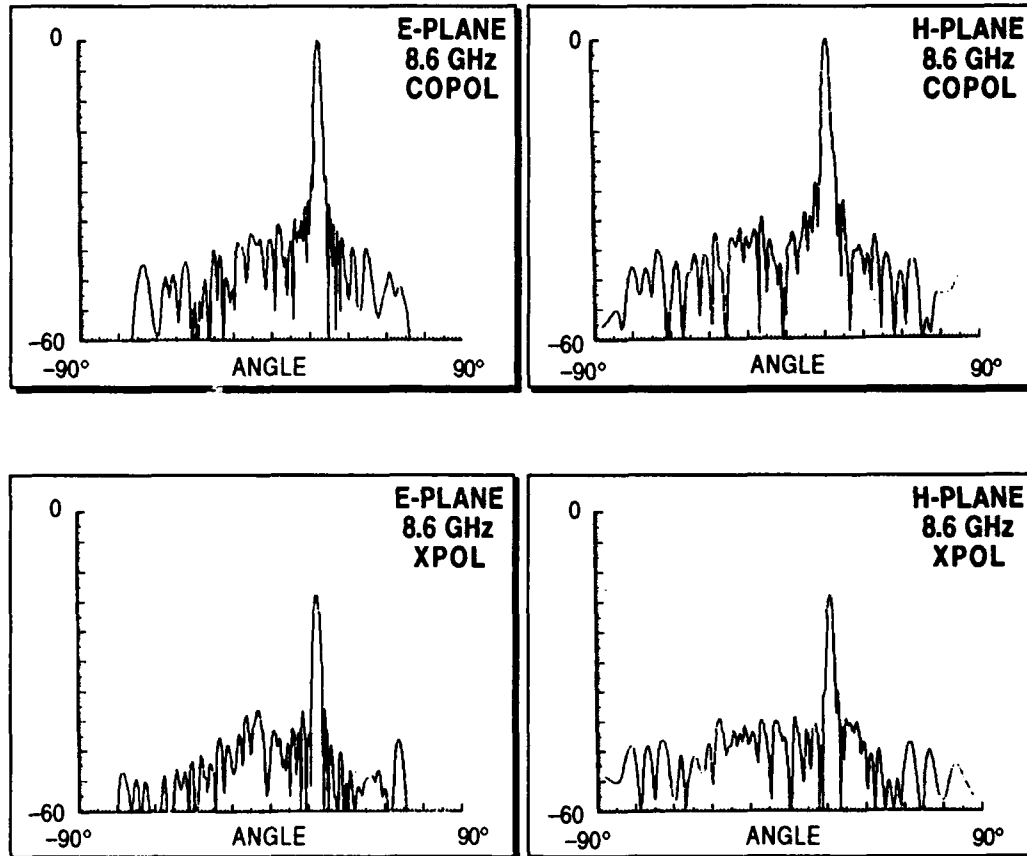
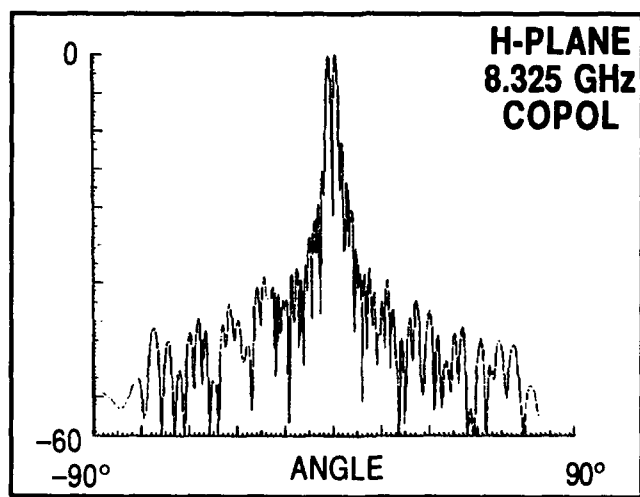
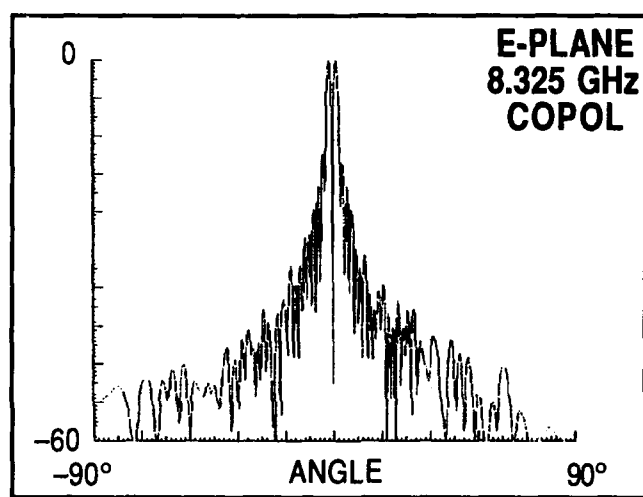


Figure 4-10
Receive RCP Patterns
Maximum Scan Condition



Typical difference patterns in the elevation and horizontal planes at boresight are shown in Fig. 4-11. The low sidelobes exhibited by these patterns are inherent in space fed arrays using a monopulse multimode feed.

Figure 4-11
Receive LCP Difference Patterns
Boresight



5.0 Conclusions

The results of this work indicate that an efficient circularly polarized feed capable of transmitting high power for a space fed phased array can be realized using four stepped septum polarizers in a cruciform arrangement. Use of a multimode multi-flare horn is shown to provide very good secondary pattern performance in both the transmit and monopulse modes of operation.

6.0 Acknowledgements

This work was performed for the U.S. Naval Air Warfare Center, China Lake, CA under a subcontract from Loral Space & Range Systems, Ridgecrest, CA. The authors wish to acknowledge the support of NAWC, Loral and Microwave Applications Group.

7.0 References

1. Goudey, K.R. and Sciambi, A.F. (1978) "High Power X-Band Monopulse Tracking Feed for the Lincoln Laboratory Long-Range Imaging Radar," IEEE Trans. on Microwave Theory and Techniques, Vol. MTT-26, No. 5, May 1978, pp. 326-332.
2. Chen, M.H. and Tsandoulas, G.N. (1973) "A Wide-Band Square Waveguide Array Polarizer," IEEE Trans. on Antennas and Propagation, Vol. AP-21, No. 3, pp. 389-391.
3. Cohn, S.B., (1970) "Flare-Angle Changes in a Horn as a Means of Pattern Control," Microwave Journal, Vol. 13, No. 10, Oct. 1970, pp. 41-ff.

**INVISIBLE U-SHAPED SLOT
COMMUNICATIONS ANTENNA**

X00170

H. Shnitkin

**Norden Systems, Inc.
P.O. Box 5300
Norwalk, CT 06856**

1992 Antenna Applications Symposium

September 1992

INVISIBLE U-SHAPED SLOT COMMUNICATIONS ANTENNA

**H. Shnitkin
Norden Systems, Inc.
P.O. Box 5300
Norwalk, CT. 06856**

Abstract

The increase of vandalism and pilferage of vehicle installed communication equipment and antennas has created a need for invisible radio antennas in the UHF band. Users of cellular telephone and law-enforcement under-cover teams desire to disguise their radio and antenna equipment to prevent theft, breakage, and exposure. This paper presents an inexpensive invisible, i.e. covert, antenna design for two-way communication, for both vehicle mounted and below-ground installations.

1. Background

Covert and disguised antenna types have been used by law-enforcement agencies since the advent of mobile radio. They have been disguised as AM broadcast antennas, mounted on fenders and side view mirrors. Some agencies have cut a large hole into the car's roof or trunk to form a half-wave slot antenna, which was then covered with fiberglass and refinished to blend in with the rest of the car body. There have even

been some military studies on the feasibility of insulating the engine hood or trunk lid from the rest of the body to form a resonant slot antenna.

None of the above mentioned solutions are completely satisfactory. AM broadcast and side view mirror antennas are still subject to vandalism, side view mirror antennas are quite inefficient, and the trunk lid, hood or roof slot antenna are impractical since they are far too expensive and can never be transferred from one vehicle to another.

In 1978 a UHF U-shaped slot antenna was invented (U.S. PAT. No. 4,124,851), ideally suited for an invisible communication antenna installation in an automobile (See Figure 1). It is essentially a probe-fed, single ridge waveguide, cavity-backed, U-shaped slot radiator. It is configured to fit precisely into the position originally designed for mounting an automobile rear-deck loudspeaker. The cavity is located in the trunk of the car and radiation takes place through the loud speaker grill and the car's rear window. (See Figure 2)

The U-shaped slot antenna produces a vertically polarized, omnidirectional radiation pattern. Because it exhibits a high radiation resistance, shunted by a resonant high-Q cavity, a microstrip transformer network is required to match its impedance to that of a 50-ohm cable.

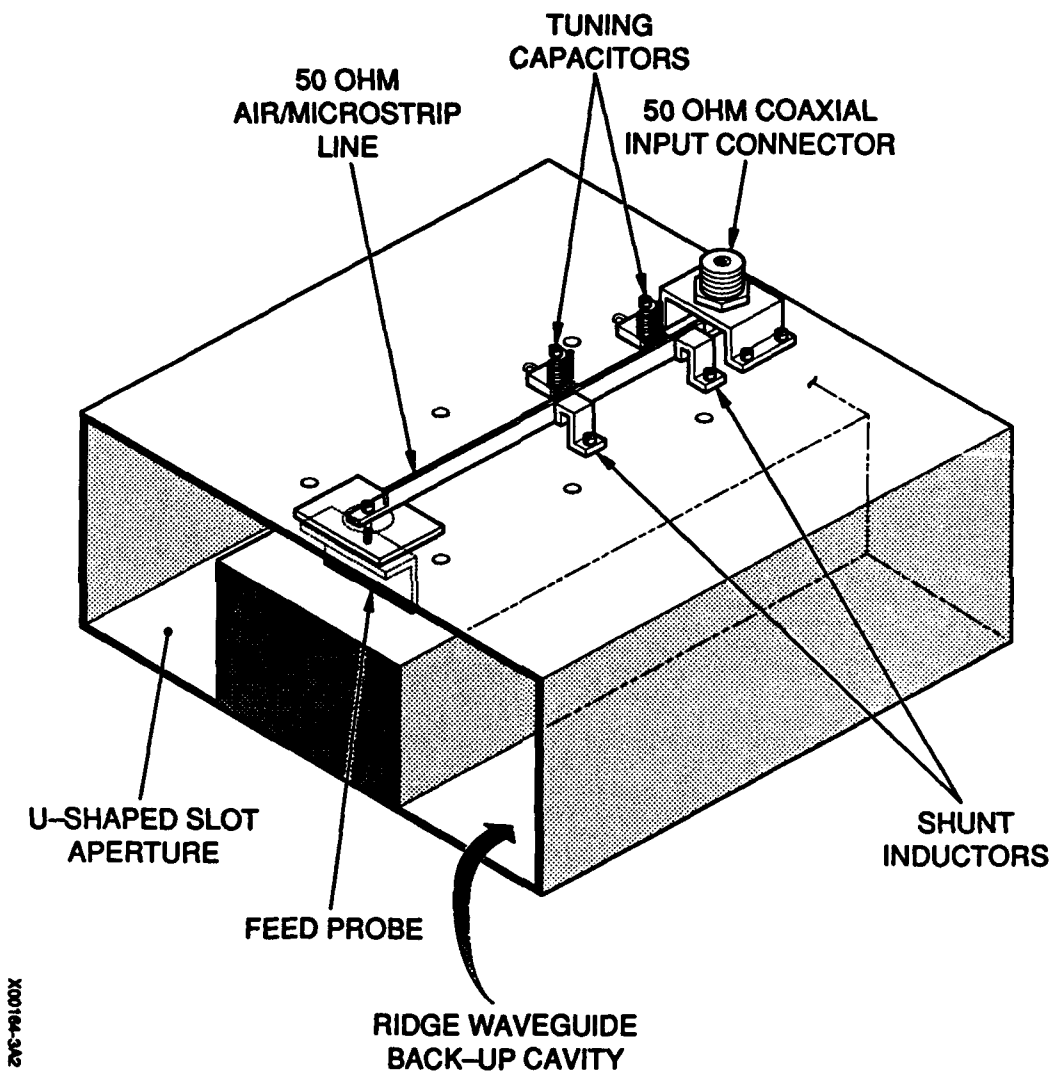


Figure 1. U-Shaped Slot Antenna Assembly

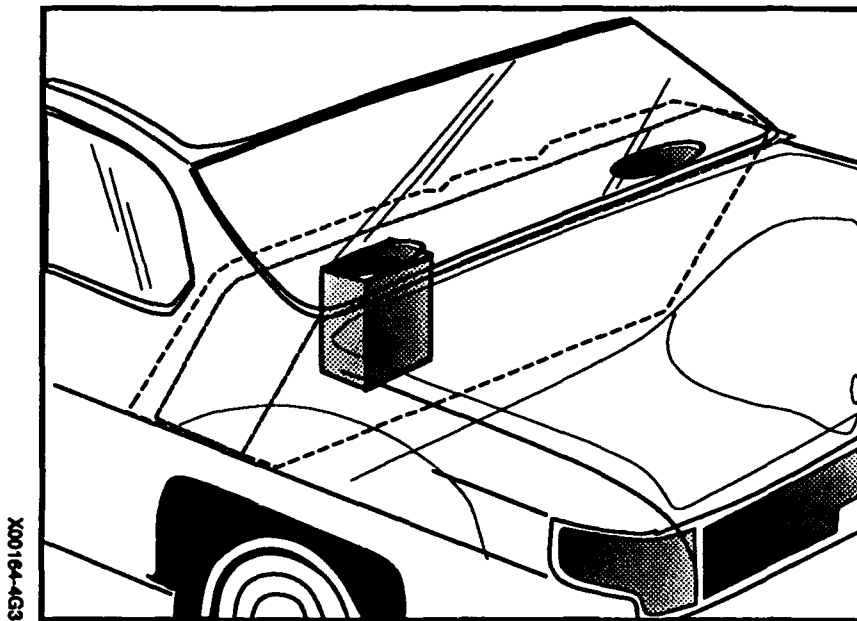


Figure 2. Automobile Installed Invisible Antenna

2. Antenna Gain and Radiation Pattern

To understand the U-shaped slot aperture, it is best to divide it into a mid-section with left/right oriented magnetic field and a two-element anti-phased array consisting of the two end-sections with front/rear oriented magnetic fields, as shown in Figure 3. The mid-section radiates vertically polarized maxima in the front and rear directions, and exhibits nulls in the left/right directions. The end-section two-element array is approximately 104 electrical degrees spaced and excited with 180-degree out-of-phase H-fields. Thus, its radiation pattern exhibits nulls in the forward, rearward, and upward directions as well as vertical polarization maxima in the left and right directions.

X00184-1G2

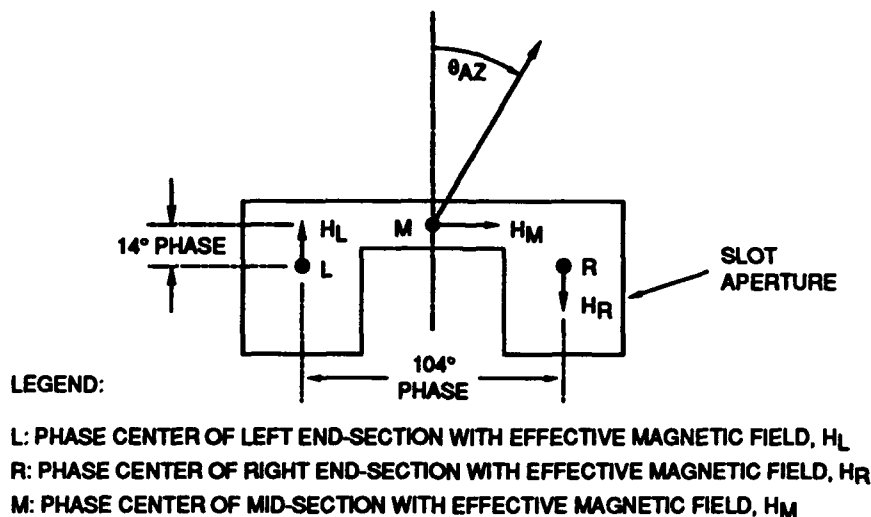


Figure 3. U-Shaped Aperture

The omnidirectional nature of the radiation pattern is illustrated in the vector diagram of Figure 4. This diagram is referenced to the radiated phase of the mid-section and shows how the phase of the end-

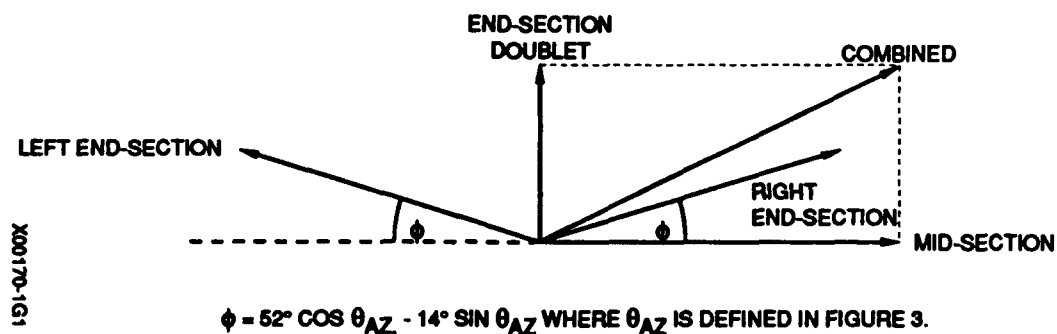
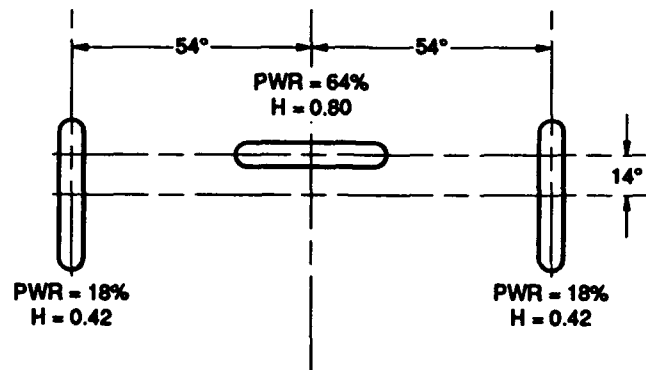


Figure 4. Vector Diagram of Mid and End-Section of U-Shaped Aperture

section anti-phased doublet is always orthogonal to that of the mid-section. It should be noted that vector magnitudes are proportional to the magnetic field excitation and the sine of the angle formed by that field and the azimuth direction. Since the phase center of the mid-section and that of the two-element array formed by the two end-sections are both located at the physical center of the antenna, and because their relative phase of radiation is 90 degrees, no sharp radiation nulls can occur in any direction. This behavior is analogous to that of the omnidirectional turnstile antenna.

To compute both gain and radiation pattern, the power division between the center and end-sections was first determined by integrating the electrical field over the respective areas. This resulted in 64 percent of the power in the mid-section and 18 percent in each end-section.

For computational purposes, the U-shaped aperture has been replaced by three separate coplanar slot radiators, polarized and located as shown in Figure 3 and 5. A computer program was written, which evaluates the radiated voltage pattern for any azimuth and elevation angle throughout a hemispherical volume for arbitrarily located radiators with arbitrary polarization angles. In addition, the program integrates the total radiated power and thus derives absolute directivity values. Applying this program to the horizontally oriented U-shaped slot aperture results in the antenna directivity values shown in Table 1. It can be seen that at an elevation angle of 5 degrees, the directivity of the invisible U-slot antenna is 2.7 ± 0.8 dB. This mean value directivity is 2.3 dB less than that of the standard visible quarter-wave monopole.



1. COMPUTE THREE SLOT PATTERNS OVER HEMISPHERE
2. AT EACH SPATIAL DIRECTION, ADD THREE VECTORS IN PROPER POLARIZATION AND PHASE (H_ϕ, θ)
3. INTEGRATE POWER OVER HEMISPHERE: $P = \oint |H_\phi, \theta|^2 \sin\theta d\theta d\phi$
4. CALCULATE GAIN IN ϕ, θ DIRECTION, BY: $10 \text{ LOG} |H_\phi, \theta|^2 \cdot 4\pi / P$

Figure 5. Pattern and Gain Computation

Table 1. U-shaped Slot Directivity Over Hemisphere

EL Angle (deg)	Azimuth Angle (deg)								
	(Rear) -90	-67.5	-45	-22.5	(Left & Right) 0	22.5	45	67.5	(Fwd) 90
85	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
75	3.5	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.5
65	3.5	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.5
55	3.5	3.4	3.3	3.2	3.2	3.2	3.3	3.4	3.5
45	3.5	3.4	3.2	3.1	3.0	3.1	3.2	3.4	3.5
35	3.5	3.3	3.0	2.8	2.7	2.8	3.0	3.3	3.5
25	3.5	3.2	2.8	2.5	2.4	2.5	2.8	3.2	3.5
15	3.5	3.2	2.6	2.2	2.1	2.2	2.6	3.2	3.5
5	3.5	3.1	2.5	2.1	1.9	2.1	2.5	3.1	3.5

Note: All values are in decibels

3. Impedance Matching Tuner

The cavity behind the U-shaped radiator, essentially a foreshortened quarter wavelength long, short circuited ridge waveguide, appears in parallel with a radiation resistance of approximately 900 ohms. An end-loaded, open-circuited voltage probe excites this cavity from a 50-ohm microstrip line, one inch from the aperture. The electrically short probe (0.035 wavelengths) and the low ridge-waveguide impedance cause the antenna to have a Q of about 50, resulting in a narrow instantaneous bandwidth.

A tuner employing two variable trimmer capacitors, as shown in Figure 1, is provided to allow the operator to adjust for a 2.0:1 VSWR impedance match to 50 ohms at the center frequency of the 1.3% instantaneous band, anywhere within a 20% operating frequency band. The equivalent circuit of the antenna with tuner is shown in Figure 6.

This tuner is constructed of a 50-ohm, air-dielectric microstrip line to accommodate the high RF voltages and currents associated with 125 watts of RF power without overheating and with less than 0.05 dB of insertion loss. The tuner consists of a fixed shunt inductor in parallel with an air-trimmer capacitor (1.5 to 14 pF) separated by about one-eighth wavelength from a second shunt inductor in parallel with a trimmer capacitor. Adjustment of each trimmer capacitor produces a net normalized shunt susceptance in the range of $-j1.0$ to $+j1.0$. The tuner's

operation is analogous to that of a double-stub microwave tuner, with the trimmer capacitors replacing the tunable stubs. The capacitor closest to the aperture acts mostly as a line stretcher to bring the input admittance at the second capacitor unto the $G = 1.0$ circle of the Smith Chart, between the $B = -j1.0$ and $B = +j1.0$ values. Typical computed antenna input impedances for a 400 MHz to 500 MHz antenna tuned for 450 MHz are shown in Table 2.

The shunt inductors are fabricated from heavy gauge copper sheet in order to carry high RF transmit currents with minimum power loss and heat dissipation. In addition they serve as mechanical supports for the 50-ohm air/microstrip line.

4. Below-ground Installation

The U-slot antenna, discussed above, also lends itself to an underground installation as shown in Figure 7. The antenna's back-up cavity is placed into a hole in the ground so that the radiating slot aperture is flush with the terrain. About eight radial counter-poise wires, about 24 inches long, should be attached to the periphery of the slot and placed flush with the ground surface. The aperture should then be covered with camouflage cloth in order to disguise it.

Table 2. Computed Antenna Input Impedance

Input Zo: 50.00 ohms

Load Impedance: 900.00 + 0.00 j ohms

Frequency: 0.444 →→→→ 0.456 GHz by 0.001

Number	Circuit Element	Param 1	Param 2
6	Shunt Cap. (pF)	0.50000	0.00000
9	Shunt S.C. Stub	50.00000 Ω	17.91000 in.
3	Series Cap. (pF)	4.00000	0.00000
11	Transmission Line Length	50.00000 Ω	2.00000 in.
6	Shunt Cap. (pF)	10.00000	0.00000
5	Shunt Inductor (nH)	14.00000	0.00000
11	Transmission Line Length	50.00000 Ω	5.50000 in.
5	Shunt Inductor (nH)	8.00000	0.00000
6	Shunt Cap. (pF)	4.80000	0.00000

Frequency (GHz)	Input R (ohms)	Input X (ohms)	Reflection Coefficient Magnitude	Reflection Phase (deg)	VSWR
0.444	23.565	47.894	0.623	85.8	4.31
0.445	31.029	48.646	0.552	80.3	3.47
0.446	40.565	46.764	0.468	74.1	2.76
0.447	50.903	40.080	0.369	67.0	2.17
0.448	58.316	27.487	0.257	58.9	1.69
0.449	58.478	11.964	0.134	48.4	1.31
0.450	51.431	-0.671	0.016	-24.8	1.03
0.451	41.462	-7.537	0.124	-133.9	1.28
0.452	32.188	-9.712	0.245	-144.7	1.65
0.453	24.854	-9.275	0.355	-152.7	2.10
0.454	19.394	-7.685	0.452	-159.6	2.65
0.455	15.386	-5.712	0.535	-165.6	3.30
0.456	12.426	-3.707	0.604	-171.0	4.05

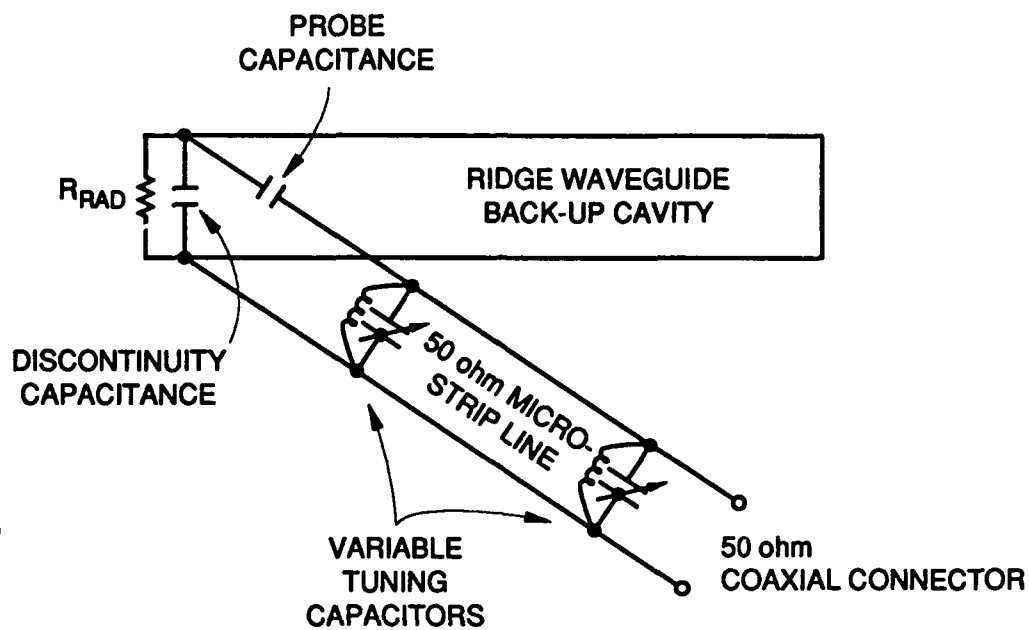


Figure 6. Equivalent Circuit

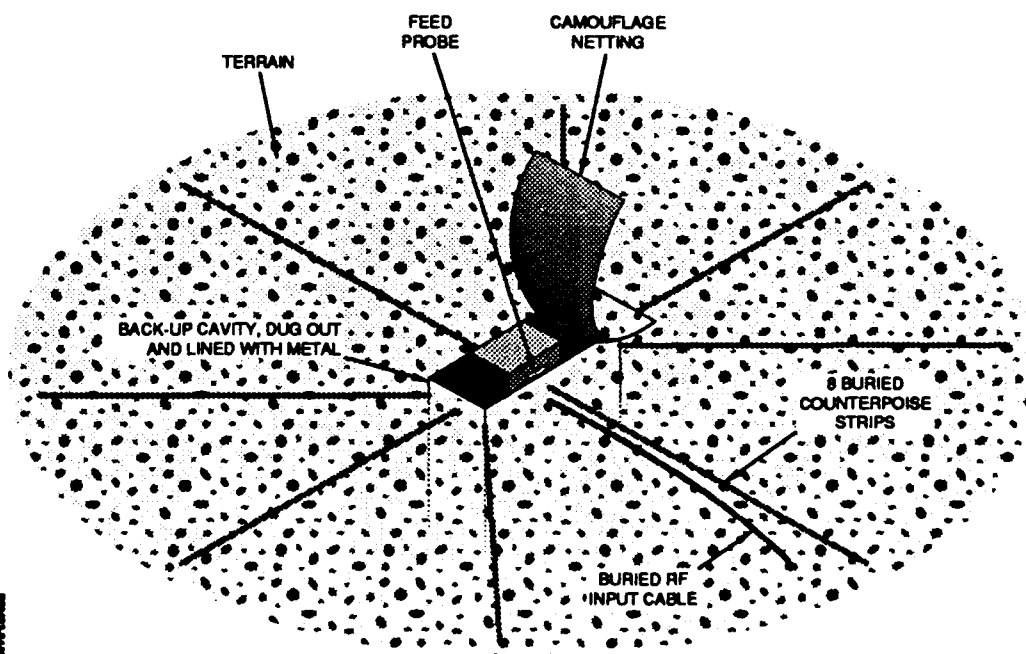


Figure 7. Underground Antenna Installation

5. Conclusion

An invisible mobile or below-ground communications slot antenna for use with cellular telephone or police radio equipment has been presented. This antenna possesses adequate gain, omnidirectionality, bandwidth, and transmit power handling capability for these applications, while offering a covert as well as vandal and burglar-proof alternative to the conventional monopole.

DESIGN CONSIDERATIONS AND ANALYSIS OF A CROSS INTERFEROMETER ANTENNA FOR AIRBORNE APPLICATIONS*

T. F. Carberry
The MITRE Corporation
Bedford, MA 01730

ABSTRACT

Most airborne interferometers are oriented in the horizontal plane to measure the azimuth angle of emitter signals. This arrangement imposes limitations for the associated direction finding (DF) system such as providing only one-dimensional emitter direction information and an azimuth error (known as coning error) that is dependent on the emitter's elevation angle. This coning error limits the vertical coverage of the horizontal interferometer. The "cross interferometer," employing both horizontal and vertical arrays, offers not only two-dimensional emitter position information but also improved azimuth angular accuracy since the vertical array provides the elevation angle needed to mitigate the effect of the coning error bias and allows the azimuth angle to be measured accurately. As a result, the cross interferometer expands the operational flexibility of the DF system since it significantly increases the vertical coverage of the DF system for a specified azimuth accuracy when compared with the horizontal interferometer.

The azimuth accuracy improvement is, of course, dependent on the angular accuracy of the vertical array. In addition, the interferometer is subject to angular ambiguities (since it employs elements with spacings greater than a half wavelength). Also, since the probability of ambiguity is additive, a vertical array can increase the DF antenna's probability of ambiguity.

This paper examines the effect of the vertical array on the performance of the DF system. Error statistics of the azimuth error as a function of channel-pair phase error, azimuth angle off broadside, elevation angle, and vertical interferometer height are developed. In addition, design factors required to achieve a low probability of ambiguity are presented.

1.0 INTRODUCTION

The interferometer is a relatively simple and accurate antenna that is effective for many direction-finding (DF) system applications. Most airborne interferometers, however, are horizontally-oriented, consisting of a single linear multiple-element receive array that mea-

*The views expressed in this paper are those of the author and do not reflect the policy or position of The MITRE Corporation.

sures the direction-of-arrival (DOA) of an incoming signal only in the azimuth or horizontal plane using relative channel-pair phase measurements. These constraints impose limitations on the DF system's performance. For example, if the emitter is not in the horizontal plane, the azimuth angle measured by a DF system is in error by an amount dependent on the elevation angle of the emitter. If high azimuth accuracy is required, this error, known as coning error, seriously restricts the horizontal interferometer to signals within a few degrees above or below the horizontal plane. This reduces the operational flexibility of the DF system. In addition, the horizontal interferometer provides only one-dimensional direction information of the signal's DOA. One-dimensional direction information limits the DF systems ability to properly sort emitter signals. Therefore, it is possible to misassociate emitters at different altitudes but at the same azimuth with the wrong aircraft. In addition, one-direction information limits the DF system's ability to discriminate against multipath.

This paper describes the generalized design and analysis of a "dual-linear" interferometer known as the "cross interferometer" that consists of horizontal and vertical interferometers. The objectives of the analysis are to demonstrate how well this dual linear antenna strategy overcomes the above described deficiencies of the horizontal interferometer. In order to compare the cross interferometer with the horizontal interferometer on a quantitative basis, it is necessary to consider specific design examples.

To conduct this analysis, some assumptions and design objectives must be made; they are as follows: (1) the DF system is broadband and operates over a 4:1 frequency range, (2) the angular accuracy design objective over the azimuth and elevation coverage sector is 1 degree or less, (3) the DF system is designed for maximum antenna gain to achieve maximum detection range, (4) the azimuth coverage sector is plus or minus 60 degrees, (5) the elevation coverage design objective is at least plus or minus 30 degrees, (6) the horizontal

and vertical interferometers are oriented on the aircraft to account for the mean aircraft pitch angle during flight, (7) the angular measurement accuracy is dominated by phase measurement errors, and (8) the maximum probability of ambiguity is 0.1 or less over the band.

These assumptions are used to frame the analysis and not tied to any specific program. One purpose of the analysis will be to determine if these design objectives can be achieved and, if not, what can be achieved.

2.0 HORIZONTAL INTERFEROMETER DESIGN ISSUES

The two major design issues associated with all interferometers are angular error and angular ambiguities. The angular error is based on the spacing of the outer-pair elements. However, the outer-pair elements, normally spaced much greater than a half-wavelength, generate angular ambiguities. These angular ambiguities are resolved using information from the inner-pair elements.

The rms angular error of an interferometer, σ_a , in degrees is a direct function of the channel-pair phase error of the outer-most channel-pair and an inverse function of the outer-most channel-pair distance and the cosine of the angle off broadside. This well known relationship is presented below:

$$\sigma_a = \frac{\sigma_{cp}}{2\pi D_{max} \cos(AZ) / \lambda} \quad (1)$$

where

σ_{cp} = rms channel-pair phase error, degrees

D_{max} = outer-most channel-pair spacing

AZ = angle off broadside

λ = wavelength

An important sub-issue concerning angular error is the operating frequency range of the DF system. According to equation 1, it is clear that the angular error increases as the frequency decreases. This property establishes a lower limit to the frequency for a given physical dimension available for the interferometer.

For an emitter in the interferometer's horizontal plane, figure 1 presents the interferometer's angular error at broadside as a function of interferometer length in terms of wavelengths ($D_{\max}/\lambda$). Also presented on the figure is a table of multiplying constants as a function of angle off broadside. As shown, for an angle off broadside by 60 degrees, the broadside error is multiplied by a factor of 2.0. At the high end of the band, let us assume that the interferometer is 14 wavelengths long. A review of this figure reveals that the channel-

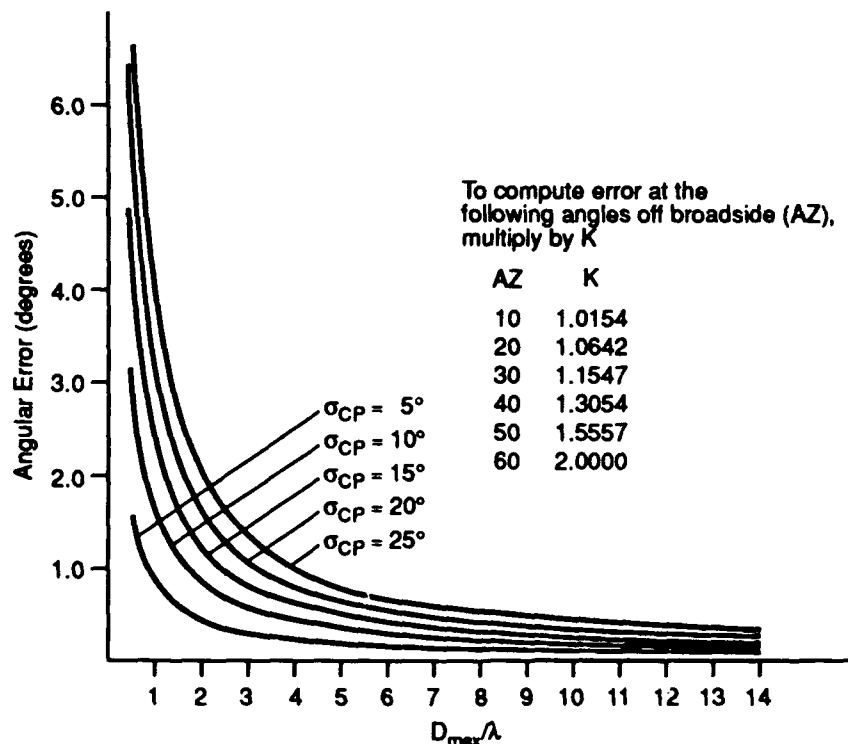


Figure 1. Interferometer Broadside Angular Error as a Function of RMS Channel-Pair Phase Error

pair phase error must be 10 degrees or less to meet the 1 degree angular error design objective. As previously mentioned, it is the low end of the band that limits the achievable angular error. At the low frequency, the interferometer is only 3.5 wavelengths long. If the channel pair phase error is 10 degrees, the angular error is 0.45 degrees at broadside and 0.91 degrees at 60 degrees off broadside. For a 5 degree channel pair phase error, the angular error is 0.23 degrees at broadside and 0.45 degrees at 60 degrees off broadside.

With reference to figure 2, if the interferometer is oriented in the horizontal plane and the emitter is not in the horizontal plane, a coning bias error results. This occurs because the interferometer actually measures the angle (DOA_y) between the unit direction vector, $\hat{r}$, and the unit horizontal interferometer axis vector, $\hat{y}$, or

$$DOA_y = \cos^{-1}(\hat{r} \cdot \hat{y}) \quad (2)$$

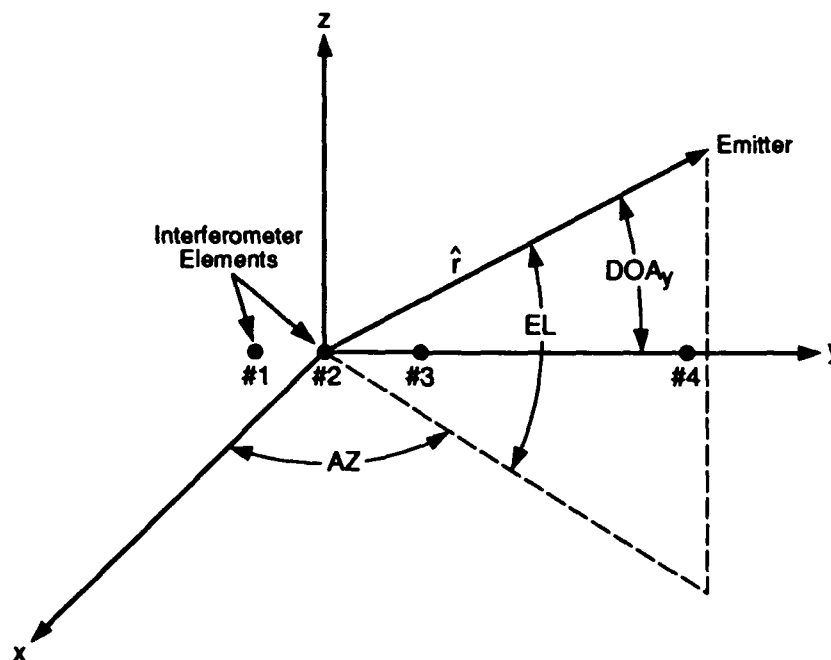


Figure 2. Four-Element Horizontal Interferometer and Coordinate System

where

- $\hat{r}$ = unit radial vector from the interferometer origin to the emitter
 $= \hat{x} \cos(EL) \cos(AZ) + \hat{y} \cos(EL) \sin(AZ) + \hat{z} \sin(EL)$
- $\hat{x}$ = unit vector normal to interferometer panel
- $\hat{y}$ = unit vector along axis of interferometer
- $\hat{z}$ = unit vector perpendicular to interferometer axis but in plane of
interferometer panel

Coning error may be computed as the magnitude of the difference between the complement of DOA_y , which is denoted herein as AZ_m , and the emitter azimuth angle (AZ), assuming that the interferometer has no errors. Thus,

$$\text{Coning error} = |AZ - AZ_m| \quad (3)$$

where

$$AZ_m = \sin^{-1} [\sin (AZ) \cos (EL)]$$

and EL = elevation angle of the signal source with respect to the horizontal plane.

Figure 3 presents the coning error introduced in the horizontal interferometer's azimuth measurement as a function of elevation angle and angle off broadside. This data assumes that the horizontal interferometer installation angle is adjusted to compensate for the mean pitch angle of the aircraft during flight. As shown, if the emitter is above or below the horizontal plane by 10 degrees and at an angle off broadside of 60 degrees, there is a coning error of 1.48 degrees. Thus, if the emitter has an elevation angle of 10 degrees and an azimuth angle of plus or minus 60 degrees, it could be misassociated with an aircraft in or near the horizontal plane that is separated from it by 1.48 degrees. This "bias error" is based on the assumption that the complement of the DOA_y (AZ_m) is equal to the AZ , which only occurs when the

emitter is in the horizontal plane. Thus, this error is a geometric derived bias that is independent of interferometer design parameters such as channel-pair phase error and frequency. Further, it cannot be calibrated out unless the DF system has elevation angle information. As shown in figure 3, there is a trade-off between the azimuth and elevation coverage sectors. For example, if the maximum allowable coning error is 1 degree, the maximum elevation coverage sector for a plus or minus 60 degree azimuth coverage sector is plus or minus 8.0 degrees. This is substantially below the elevation coverage design objective of plus or minus 30 degrees (the coning error for an emitter at 30 degrees is 11.41 degrees which is outside the field of view of figure 3.)

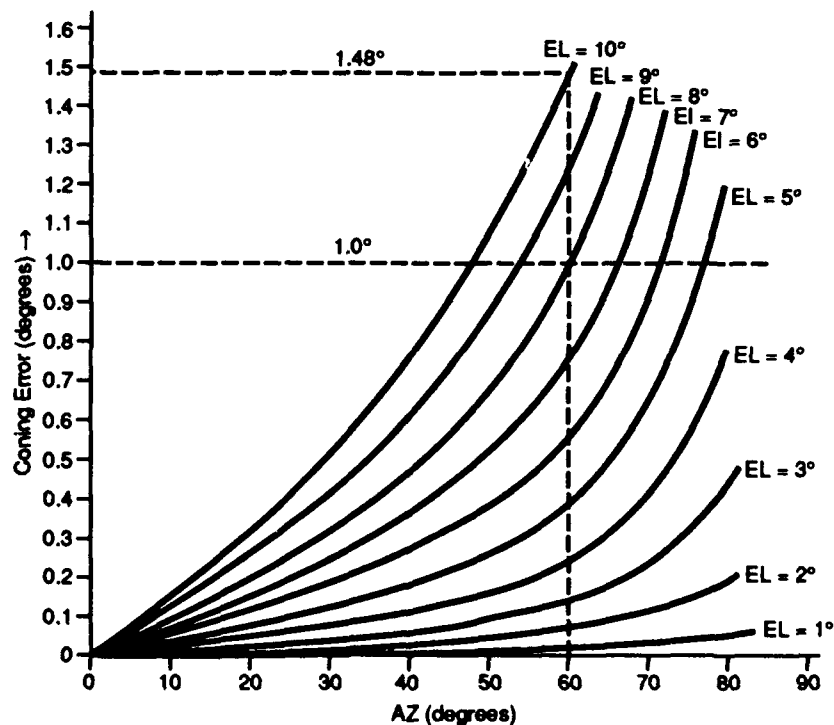


Figure 3. Coning Error as a Function of Azimuth and Elevation Angle

Clearly, coning error can limit the achievable angular accuracy and operational effectiveness of the horizontal interferometer since the horizontal interferometer is only accurate in a narrow vertical region about the horizontal plane.

The other major design issue is concerned with angular ambiguities associated with the interferometer. These result because the interferometer employs element spacings greater than a half wavelength and the phase comparator is capable of measuring channel-pair relative phase to only modulo 360 degrees. Since the number of possible DOA solutions is equal to the distance between channel-pair elements expressed in half wavelengths, the number of ambiguities and the probability of ambiguity increases with frequency. Using the previous example, at the low end of the 4:1 band, there are 7 possible DOA solutions for the outer-pair elements within visible space (plus or minus 90 degrees), only one of which is correct. However, at the high end of the band, there are 28 possible DOA solutions within visible space. Since there are multiple DOA solutions at each frequency, it is necessary to use the inner channel-pairs to provide ambiguity resolution.

An important sub-issue to angular ambiguities is the minimum element spacing which must be consistent with other requirements, the most important being the operating frequency range of the DF system. It is desirable to have some of the spacings of the inner channel-pairs be as small as possible to minimize ambiguities. If the minimum spacing were a half-wavelength or less, there would be only one DOA solution (and no ambiguities) for that channel-pair. Even if there were phase errors, there would still be one DOA solution for that channel-pair. For interferometers with more than 2 elements, it is possible for an ambiguous solution to be generated. The reason is because the DOA solution for the half-wavelength channel pair could have a large error. Since this channel-pair level result is then compared with other

channel-pair level results, it is possible to arrive at an overall interferometer level DOA result that is seriously in error because the half-wavelength spacing channel-pair DOA solution can misassociate with ambiguous solutions of the other channel-pairs.

If the DF system operates over a 4:1 frequency range, a one-half wavelength spacing at the low end of the band corresponds to 2 wavelengths at the high end of the band. Further, a 4:1 frequency range requires a broad-band flush-mounted element such as a cavity-backed spiral. If element efficiency is required to maximize antenna gain, the element diameter will be slightly under 0.5 wavelengths at the low end of the band (including the termination region of the spiral element). Thus, the 4:1 frequency range and maximum element efficiency requirements limit the minimum element spacing to approximately 2 wavelengths at the high end of the band. However, as described later, this may significantly increase the probability of ambiguity at the high end of the band.

The procedure to design the interferometer includes two steps which are: (1) an initial design based on zero channel-pair phase error using relatively prime number set element spacings, and (2) a finalized design that accounts for the effects of channel-pair phase error.

In the first step, relatively prime number set element spacings are used to eliminate unresolvable angular ambiguities. This is theoretically possible if the channel pair phase error are zero.

Figure 4 presents a 4-element horizontal interferometer. As shown, it has a 4:7:17 spacing sequence. Each spacing value refers to the distance between adjacent elements

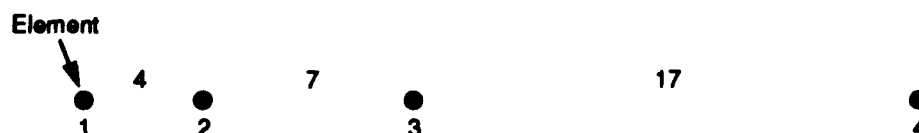


Figure 4. Candidate Interferometer Spacing Sequence

usually expressed in half wavelengths at the maximum frequency (the maximum number of DOA solutions is equal to twice the maximum dimension of the interferometer in wavelengths). The total distance between the outer elements (1,4) is 28 half-wavelengths. It should be noted that the minimum element spacing (elements (1,2)) is 4 half-wavelengths. These spacings are defined in terms of a relatively prime number set (2 out of 3 spacings should be prime numbers in a 4-element interferometer to avoid unresolvable ambiguities.)

Figure 5 illustrates the channel-pair combinations used in the interferometer. The outer pair (1,4) associated with sub-array #2 is used to measure the azimuth of the emitter while the inner pairs (1,3) and (1,2) associated with sub-array #1 are used to resolve ambiguities associated with sub-array #2.

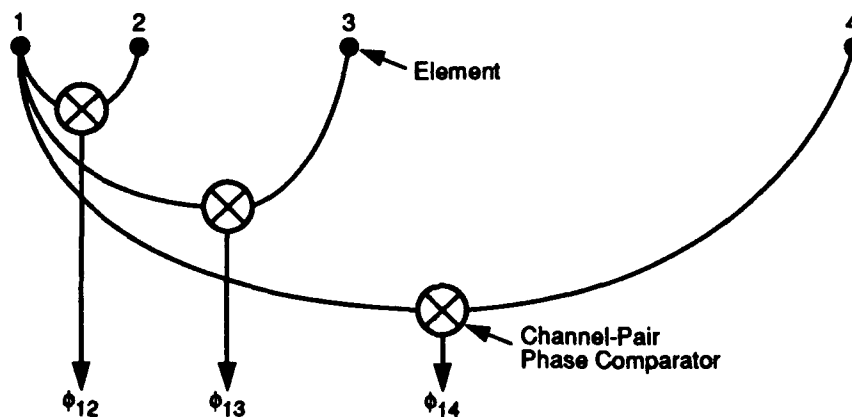


Figure 5. Assumed Channel-Pair Combinations (End-phase left)

To examine the angular ambiguities of the interferometer with zero phase error, the measured azimuth values were computed for an emitter over a range of elevation angles for specified azimuth angles. Figures 6(a), 6(b), and 6(c) present these measured azimuth values (including ambiguities over a range of elevation angles for an incoming signal at an azimuth of 0 degrees for channel-pairs (1,4), (1,3), and (1,2) respectively. As shown, the curves are

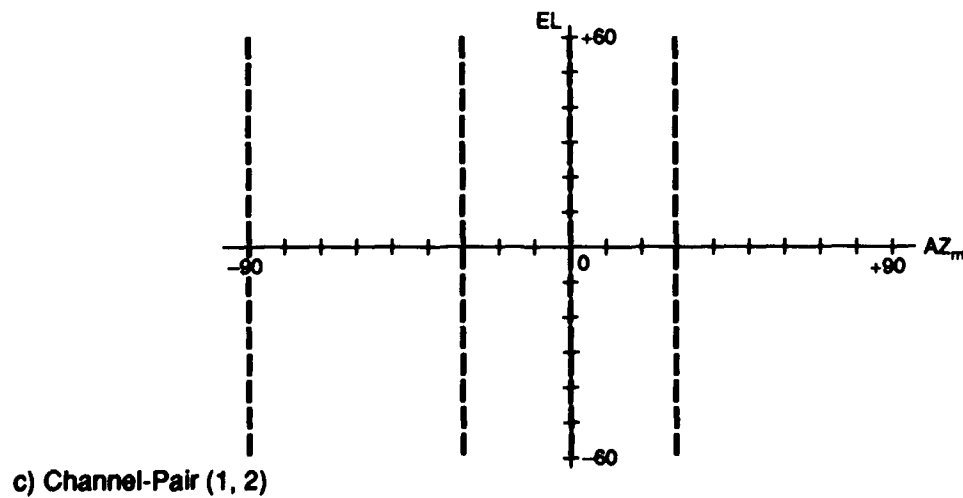
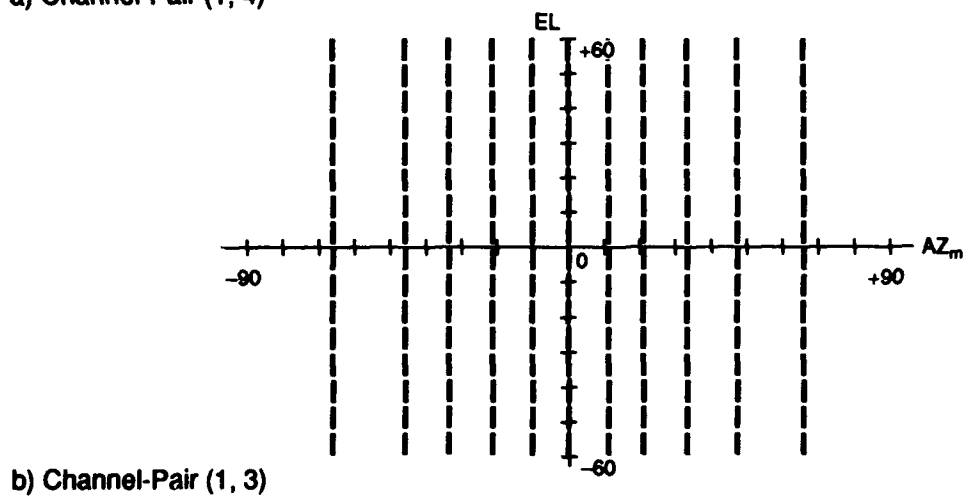
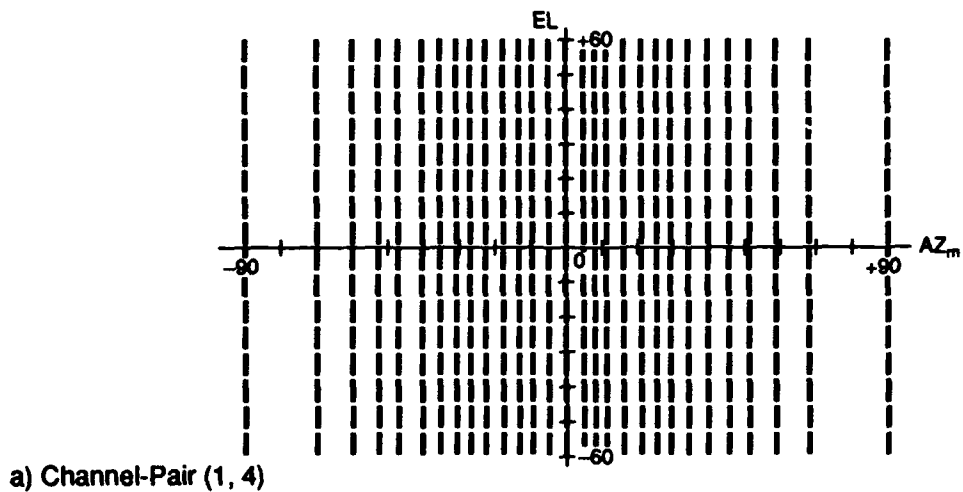


Figure 6. Measured Azimuth Values for an Emitter at $AZ = 0$ Degrees for Channel Pairs (1, 4), (1, 3) and (1, 2)

loci for a set of constant AZ values. The loci are presented over a plus or minus 90 degree azimuth range and over a plus or minus 60 degree elevation range. As seen, the loci are all straight lines. Each AZ solution has only one measured azimuth value, AZ_m , independent of elevation. Further, the only common AZ_m value for all three channel-pairs is 0 degrees. Figures 7(a), 7(b), and 7(c) present the measured azimuth values for an emitter at an azimuth of +30 degrees for channel-pairs (1,4), (1,3), and (1,2), respectively. The only common AZ locus is the one for +30 degrees. However, the coning error effect is obvious. The AZ loci are curved lines; thus, each locus encompasses a range of possible AZ_m values. Further, each locus overlaps with other AZ loci measured at different elevation angles. For example, an emitter that is at an azimuth of +34 degrees and at an elevation of plus or minus 26 degrees could be misassociated with an aircraft at an azimuth of +30 degrees and an elevation of 0 degrees. This will restrict the usefulness of the horizontal interferometer to signals near the horizontal plane if angular accuracies of 1 degree or less are required.

This does not account for the effects of channel-pair phase errors. Assuming the rms phase error is equal for all three channel-pairs, the angular accuracy of each channel-pair is an inverse function of the channel-pair spacing (apply equation 1 to the other channel-pairs). Thus, a signal at +30 degrees may have an rms (1-sigma) angular error of 0.525 degrees for channel-pair (1,4), 1.34 degrees for channel-pair (1,3), and 3.68 degrees for channel-pair (1,2), assuming the rms channel-pair phase error is 10 degrees. Figure 8 illustrates these angular error distributions at +30 degrees. As shown, channel-pair (1,4) has ambiguous loci nearby at 25.38 and 34.85 degrees. Clearly, an ambiguous azimuth measurement is possible.

In the second step of the design procedure, the final design examines the effect of phase errors on the probability of ambiguity. This requires more detailed discussion of the synthesis procedure employed in designing the 4-element interferometer described earlier.

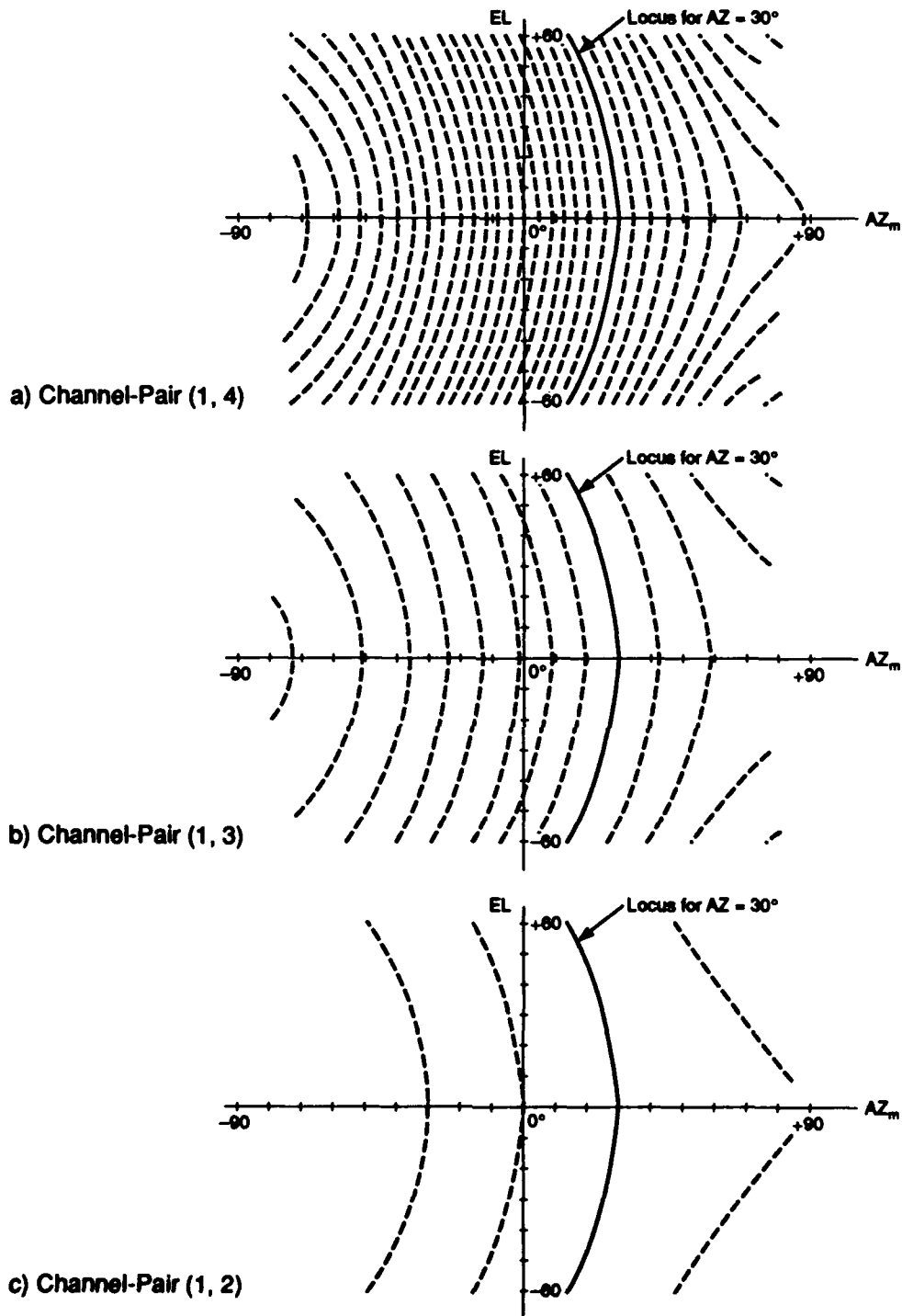


Figure 7. Measured Azimuth Values for an Emitter at $AZ = 30$ Degrees for Channel Pairs (1, 4), (1, 3), and (1, 2)

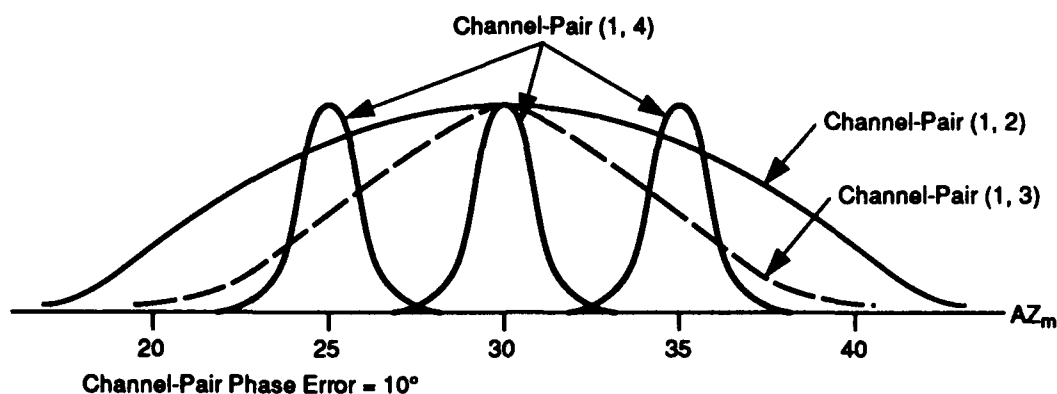


Figure 8. Channel-Pair Angular Error Distributions Near $AZ_m = 30^\circ$ for an Incident Wave at Azimuth = 30°

A 4-element interferometer is actually synthesized by combining two 3-element interferometers "sub-arrays". Three-element interferometers are used as building blocks for higher order interferometers because they are well understood. Sub-array #1 consists of channel-pairs (1,2) and (1,3), while sub-array #2 consists of channel-pairs (1,3) and (1,4). Sub-array #1 has a spacing sequence 4:7, while sub-array #2 has a spacing sequence 11:17. Associated with each sub-array is the sub-array design ratio, R ; that is

$$R_1 = (d_{12} + d_{23})/d_{12} = m_1/n_1 \quad (4a)$$

and

$$R_2 = (d_{12} + d_{23} + d_{34})/d_{13} = m_2/n_2 \quad (4b)$$

where d_{12} , d_{23} , and d_{34} are the element spacings expressed usually in half-wavelengths and m_1 , n_1 , m_2 , and n_2 are the design ratio integer numerators and denominators after any common factors are removed. For sub-array #1, $R_1 = (4 + 7)/4 = 11/4$, while for sub-array #2, $R_2 = (4 + 7 + 17)/11 = 28/11$. Hence, $m_1 = 11$, $n_1 = 4$, $m_2 = 28$, and $n_2 = 11$. Usually, these

numerator and denominator values are computed only at the maximum frequency in the synthesis procedure. Thus, they are normally integer values. However, in the analysis of probability of ambiguity (described later), they assume non-integer values since they are computed as a function of frequency.

The probability of ambiguity of the 4-element interferometer can now be calculated by first calculating the probability of ambiguity of each 3-element sub-array used to synthesize the 4-element interferometer. The probability that an ambiguity will occur for a 3-element interferometer can be computed by the following expression:

$$P_{3a} = \operatorname{erfc} \left[\frac{\pi}{\sqrt{2} K_a \sigma_{cp}} \right] \quad (5)$$

where

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-\xi^2) d\xi$$

$$K_a = \sqrt{m^2 - mn + n^2}$$

and

m = sub-array design ratio numerator value

n = sub-array design ratio denominator value

Thus, the probability of ambiguity for sub-array #1, ($P_{3a}(\#1)$), and sub-array #2, ($P_{3a}(\#2)$), can then be calculated from equation 5. However, to minimize the overall 4-element interferometer probability of ambiguity, sub-array #2 should incorporate the resolving action of sub-array #1. Since sub-array #1 increases the ambiguity boundaries of sub-array #2 by m_1 when sub-array #1 is used to resolve sub-array #2's ambiguities, the argument of equation 5 for $P_{3a}(\#2)$ is multiplied by m_1 .

The overall 4-element interferometer probability of ambiguity is

$$P_{4a} = 1 - [1 - P_{3a}(\#1)] [1 - P_{3a}(\#2)] \quad (6)$$

where

$P_{3a}(\#1)$ = probability of ambiguity of sub-array #1

$P_{3a}(\#2)$ = probability of ambiguity of sub-array #2

(resolved by sub-array #1)

It should be noted that the probability of ambiguity of $P_{3a}(\#2) \ll P_{3a}(\#1)$ since the ambiguities of sub-array #2 have been effectively resolved by sub-array #1. Thus, the probability of ambiguity of this four element interferometer is virtually identical to the probability of ambiguity of sub-array #1 or $P_{4a} \approx P_{3a}(\#1)$.

The 4-element interferometer probability of ambiguity as a function of frequency and channel-pair phase error is presented in figure 9. As shown, the probability of ambiguity increases with frequency and channel-pair phase error. The probability of ambiguity at the high end of the 4:1 frequency range is over 0.4 for a 25 degree phase error and approximately 0.0625 for a 10 degree rms phase error. Not shown is the 5 degree rms phase error result since it is too small for the scale. It is approximately 0.000175 at the high end of the band. The relatively large probability of ambiguity values at the high end of the band are due largely to the 4:1 frequency range and the requirement for the minimum spacing associated with elements-pair (1,2). Assuming the previously stated assumptions are not violated, to achieve a probability of ambiguity of 0.1 or less requires tight channel-pair rms phase errors (10 degrees or less).

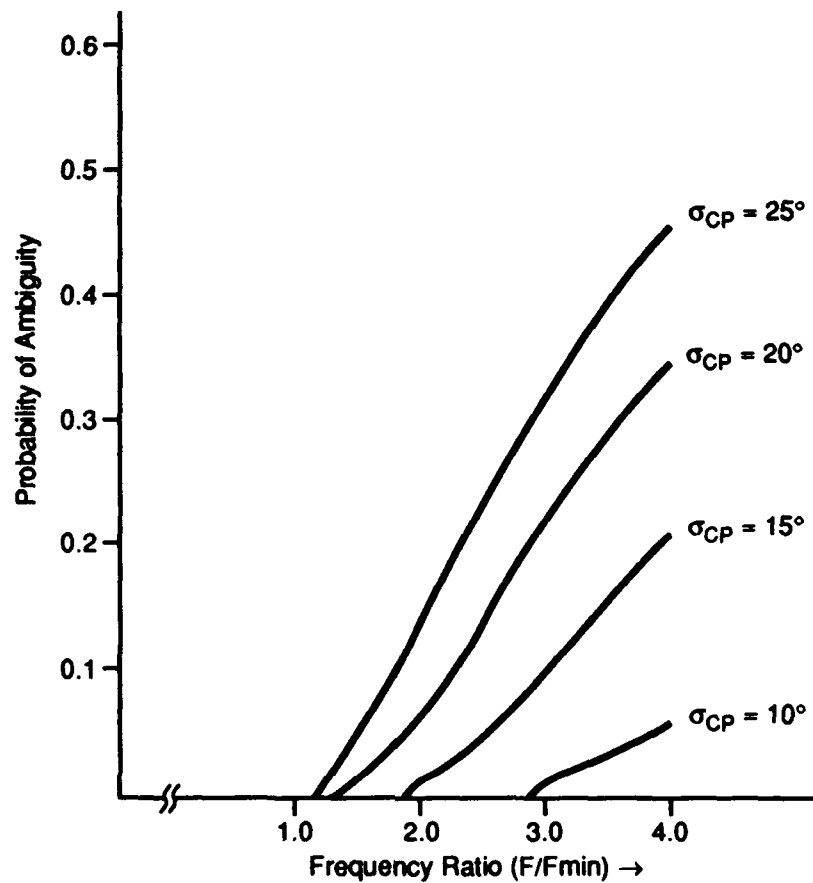


Figure 9. Probability of Ambiguity of the Four-Element Interferometer (4:7:17) as a Function of Frequency Ratio and RMS Channel-Pair Phase Error

3.0 THE CROSS INTERFEROMETER

Figure 10 presents the basic "cross interferometer" concept. As shown, it consists of a horizontal array of elements along the y-axis and a vertical array of elements along the z-axis. The total of seven elements (four horizontal, four vertical, with one common) is strictly illustrative. The actual number depends on DF system performance requirements and available aircraft real estate.

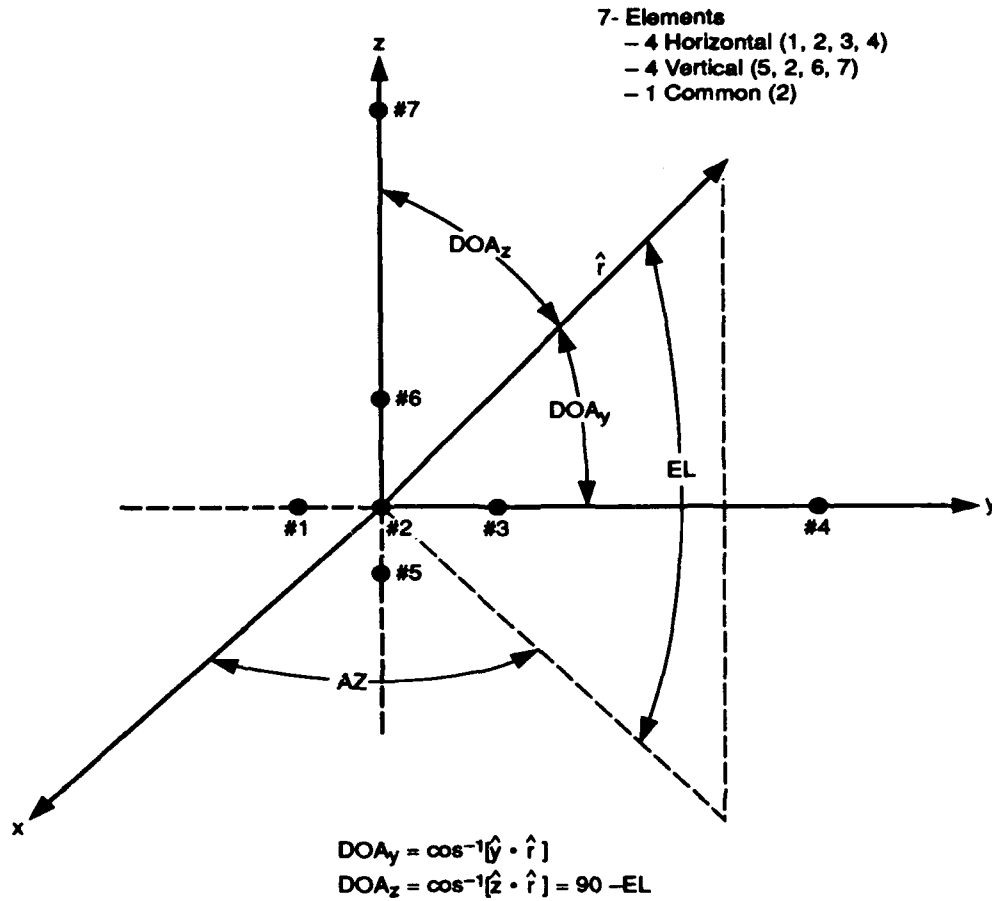


Figure 10. Illustrative Cross-Interferometer Concept

For a signal at an azimuth angle, AZ, and elevation angle, EL, the measured phase differences across the horizontal and vertical interferometer outer pair elements (Ψ_h and Ψ_v , respectively, assuming no phase error) are:

$$\Psi_h = 2\pi D_{hmax} \cos(EL) \sin(AZ)/\lambda \text{ (modulo } 2\pi) \quad (7a)$$

and

$$\Psi_v = 2\pi D_{vmax} \sin(EL)/\lambda \text{ (modulo } 2\pi) \quad (7b)$$

where λ is the wavelength and D_{hmax} and D_{vmax} are the distances between the outermost channel-pair elements for the horizontal and vertical interferometers, respectively.

As shown, these phase measurements are modulo 2π since the phase comparator is limited to this range.

Like the horizontal interferometer, the vertical interferometer can have ambiguities. If we assume that the vertical interferometer is identical to the horizontal interferometer, figures 11a, 11b, and 11c present the loci of EL values for the vertical interferometer for a signal incident on the interferometer at an elevation angle of -10 degrees over a range of azimuth angles. As shown, the vertical interferometer directly measures the complement of the elevation angle (as shown in figure 10). As shown, the loci are all straight lines.

Two vertical interferometers will be examined in this study. These are: (1) a 4 element interferometer identical to the horizontal interferometer and (2) a 3 element interferometer identical to sub-array #1 of the horizontal interferometer. As noted in the previous section, the probability of ambiguity of sub-array #1 (4:7) is virtually identical to the four-element interferometer (4:7:17) since it controls the overall probability of ambiguity of the four-element interferometer.

Both configurations employ the horizontal interferometer defined in the previous section but one configuration (denoted configuration A) uses the 4-element vertical interferometer described above (see elements #5, #2, #6, and #7 in figure 10) while the second configuration (denoted B) uses a 3-element vertical interferometer also described above (see elements #5, #2, and #6 in figure 10).

The overall probability of ambiguity for the combined horizontal and vertical interferometer configuration is as follows:

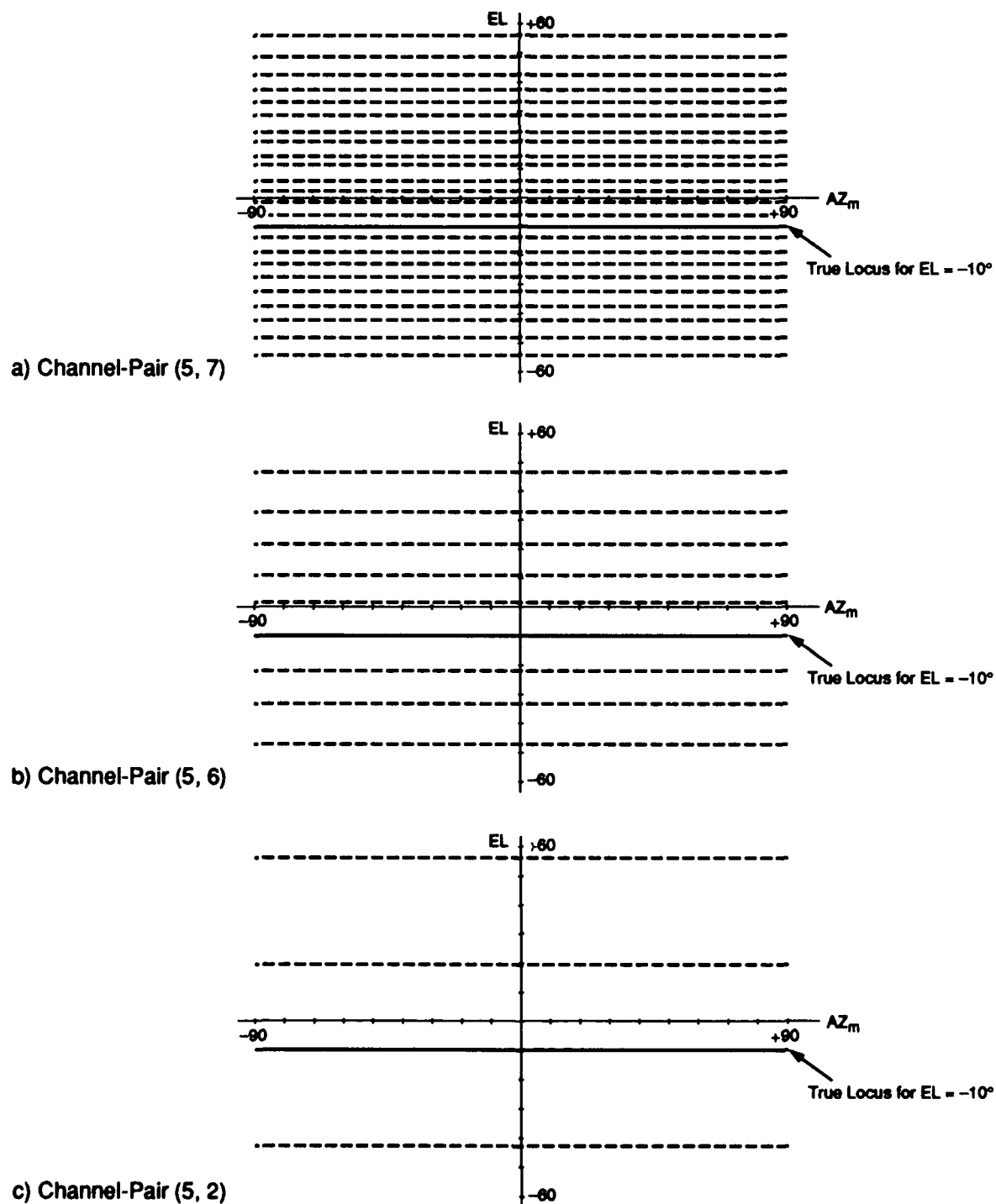


Figure 11. Loci of EL Values for a Vertical Interferometer for an Incident Wave at Elevation = -10 Degrees (Channel Pairs (5, 7), (5, 6), and (5, 2))

$$P_c = 1 - [1 - P_{4h}] [1 - P_{4v}] \quad (8)$$

or

$$P_c = P_{4h} + P_{4v} - P_{4h} P_{4v}$$

As shown, the combined probability of ambiguity of the cross interferometer configurations is approximately the sum of the two probabilities of the horizontal and vertical interferometers. Assuming the rms channel-pair phase error is 10 degrees, the probability of ambiguity for the horizontal interferometer at the maximum frequency is 0.0625, and the probability of ambiguity of the cross interferometer is slightly more than 0.12. However, reducing the channel-pair phase error to 5 degrees, the probability of ambiguity is less than 0.00035. It should be noted that this is the worse case probability of ambiguity at the high end of the band.

Assuming the inner-pair ambiguity resolving strategy is effective, the two-dimensional position of the emitter is

$$EL = \sin^{-1}[\Psi_v \lambda / (2\pi D_{vmax})] \quad (9a)$$

$$AZ = \sin^{-1}[\Psi_h \lambda / (2\pi D_{hmax} \cos(EL))] \quad (9b)$$

Equation 9b is the corrected AZ measurement accounting for the elevation angle.

Figure 12 presents the corrected locus (minus ambiguities) for a signal incident on the cross interferometer at an azimuth angle of +30 degrees and an elevation angle of -10 degrees (with no phase error). As shown, with no phase error, the corrected azimuth measurement locus over a range of elevation angles is now a straight line.

However, it is clear from equation 9b that an elevation angular error will influence the azimuth error, since the elevation angle, EL, is contained in this expression. Using standard

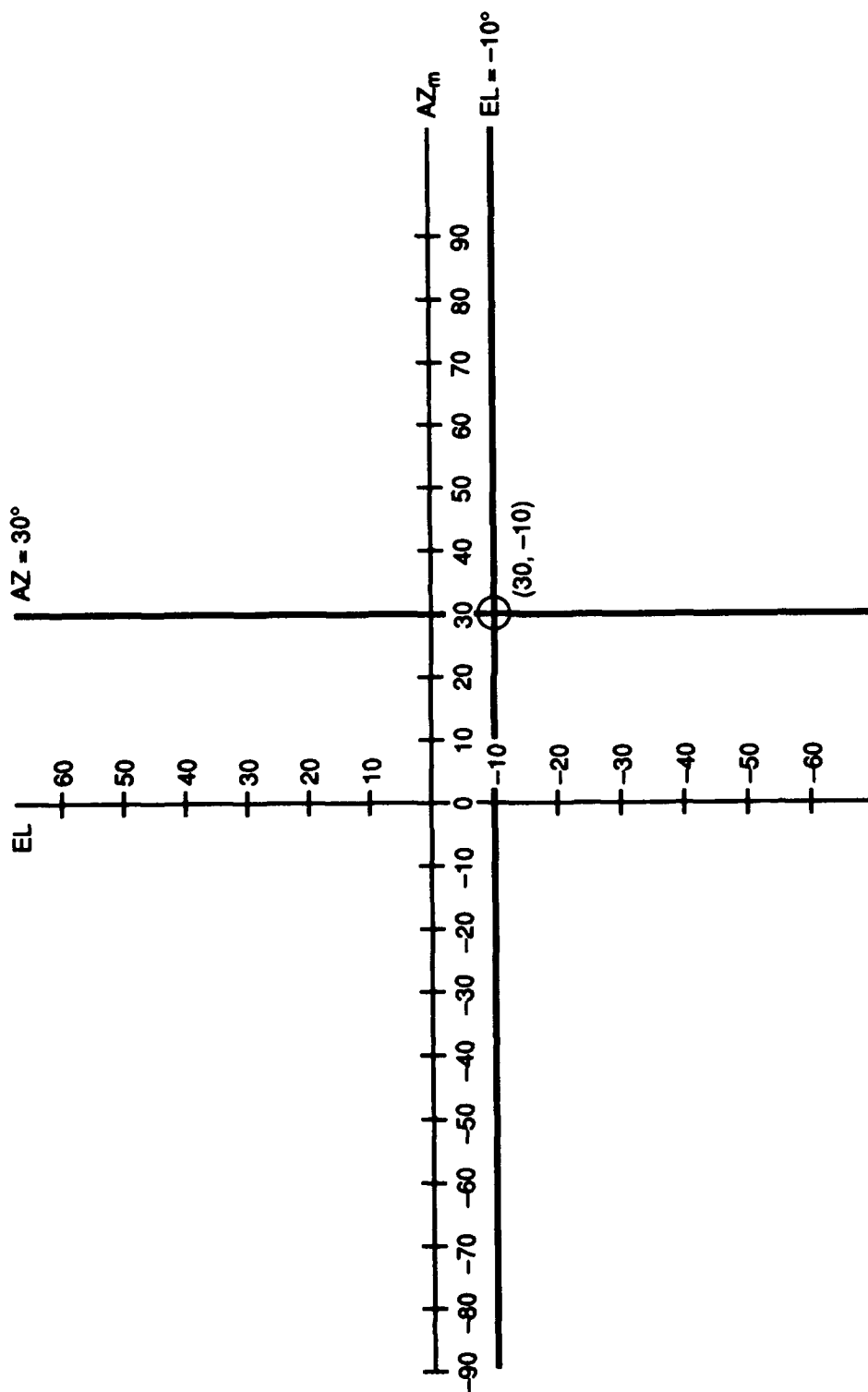


Figure 12. Measured AZ and EL Values for an Emitter at $AZ = +30^\circ$ and $EL = -10^\circ$
(No Phase Error)

small error analysis procedures. The error is approximated by linearization of equations 9a and 9b. Further, any residual bias error associated with the coning bias error is negligible. The remaining errors are random. To determine the magnitude of these random errors, the elements of the covariance matrix for (AZ,EL) were derived and are presented below:

$$\text{Var}(AZ) = (\partial AZ/\partial \Psi_h)^2 \text{Var}(\Psi_h) + (\partial AZ/\partial \Psi_v)^2 \text{Var}(\Psi_v) \quad (10a)$$

$$\text{Var}(EL) = (\partial EL/\partial \Psi_h)^2 \text{Var}(\Psi_h) + (\partial EL/\partial \Psi_v)^2 \text{Var}(\Psi_v) \quad (10b)$$

$$\begin{aligned} \text{Covar}(AZ,EL) &= (\partial EL/\partial \Psi_h) (\partial AZ/\partial \Psi_h) \text{Var}(\Psi_h) \\ &+ (\partial EL/\partial \Psi_v) (\partial AZ/\partial \Psi_v) \text{Var}(\Psi_v) \end{aligned} \quad (10c)$$

These partial derivatives were derived employing equations 9a and 9b used to calculate the channel-pair phases for the horizontal and vertical interferometers. Thus,

$$(\partial EL/\partial \Psi_h)_{\Psi_v} = 0 \quad (11a)$$

$$(\partial EL/\partial \Psi_v)_{\Psi_h} = \lambda/[2\pi D_{v\max} \cos(EL)] \quad (11b)$$

$$(\partial AZ/\partial \Psi_h)_{\Psi_v} = \lambda/[2\pi D_{h\max} \cos(EL) \cos(AZ)] \quad (11c)$$

$$(\partial AZ/\partial \lambda_v)_{\Psi_h} = F_1(\lambda, AZ, EL)/F_2(\lambda, AZ, EL) \quad (11d)$$

where

$$F_1(\lambda, AZ, EL) = \Psi_h (\lambda/(2\pi))^2 \tan(EL)$$

$$F_2(\lambda, AZ, EL) = D_{v\max} D_{h\max} \cos(AZ) \cos^2(EL)$$

Of particular interest is the covariance matrix element $\text{Var}(AZ)$, which is computed as a function of channel-pair phase error, elevation angle (EL), and azimuth angle (AZ). It is used to determine the influence of the vertical interferometer height ($D_{v\max}$) on the azimuth measurement. Since the angular error increases as the frequency decreases, the calculations were performed at the lowest frequency within the band. Thus, the assumed horizontal

interferometer length (D_{hmax}) was 3.5 wavelengths. The two vertical interferometer heights were 3.5λ (configuration A) and 1.375λ (configuration B). Figure 13 presents the azimuth error as a function of elevation angle for three azimuth angles: 0, 30, and 60 degrees. Figures 13a, and 13b present this data for 5 and 10 degree rms channel-pair phase errors, respectively. This azimuth error is the standard deviation (or square root of the variance, $\text{Var}(AZ)$). The azimuth error increases as the azimuth angle or elevation angle increases. Further, it increases as the vertical interferometer height, D_{vmax} , decreases. As shown, for off-principal-plane positions ($AZ \neq 0$ degrees or $EL \neq 0$ degrees), for a given elevation angle, increasing vertical interferometer height reduces azimuth error. However, the azimuth error increases with the angle off the principal plane. In addition, it increases with channel-pair phase error. (Note: The scales on figures 13a, and 13b change to better present the data because the azimuth error increases in proportion to the channel-pair phase error.)

4.0 OVERALL COMPARISON

Table 1 presents an overall comparison of three DF antenna configurations, the horizontal interferometer configuration and both cross interferometer configurations (A and B). It should be noted that most of the data is referenced to the low end of the band where the angular error is maximum. The only exception is the probability of ambiguity. These values are at the high end of the band where the probability of ambiguity is maximum. As shown in columns 2 and 3, the azimuth angular error for an emitter at an elevation angle of 0 degree is the same for both horizontal and cross interferometer configurations.

Columns 4 and 5 provide a comparison of the azimuth error for an emitter at $AZ = 60$ degrees and $EL = 30$ degrees. The azimuth error is divided into two components; (1) bias (coning error), and (2) a random error. As shown, for the horizontal interferometer configura-

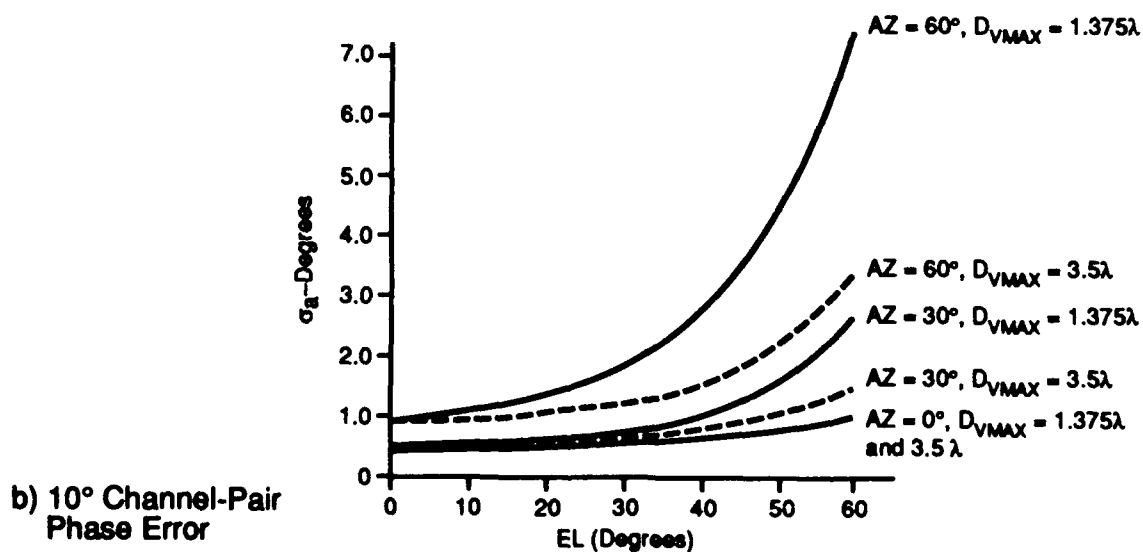
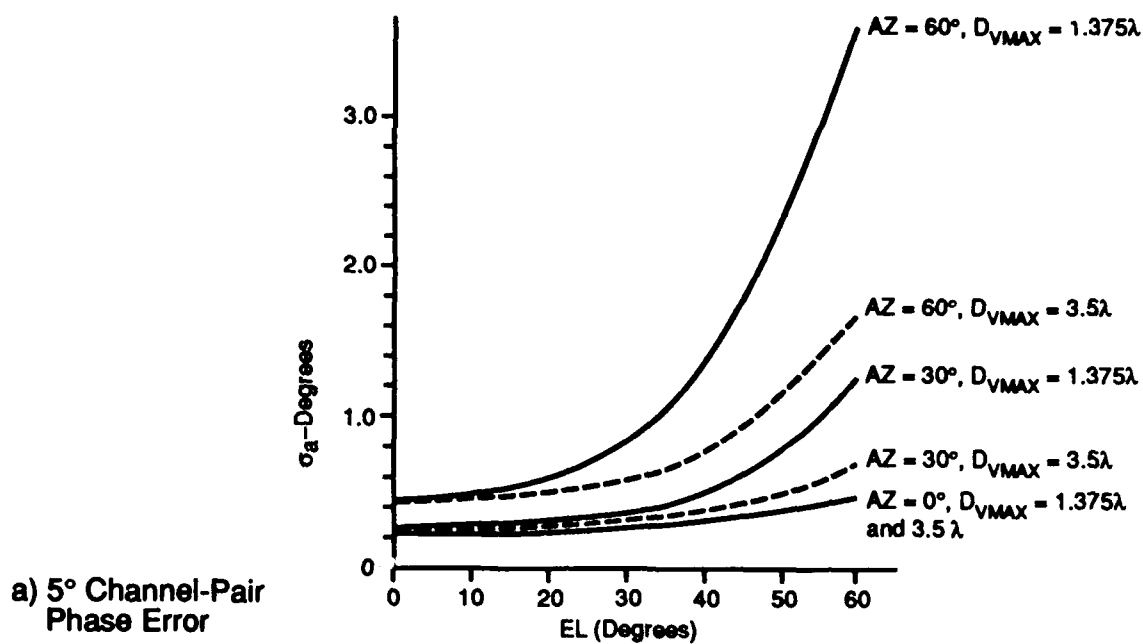


Figure 13. RMS Azimuth Error as a Function of RMS Channel Pair Phase Error, Elevation Angle, Azimuth Angle, and Vertical Interferometer Height

Table 1. Comparison of DF Antenna Configuration Performance Measures

DF Antenna Configuration	RMS Azimuth Error for Emitter at EL = 0° ⁽¹⁾		Azimuth Error for Emitter at AZ = 60°, EL = 30° ⁽¹⁾		Probability of Ambiguity ⁽²⁾	Achievable Elevation Coverage for an Azimuth Coverage of ±60°
	AZ = 0°	AZ = 60°	Mean Bias (Coning) Error	Random Error RMS, Degrees		
Horizontal Interferometer Only						
5° Channel-Pair Phase Error	0.23°	0.45°	11.41°	0.52°	1.75×10^{-4}	±6.2° ⁽³⁾
10° Channel-Pair Phase Error	0.45°	0.91°	11.41°	1.1°	6.25×10^{-2}	±2.2° ⁽³⁾
Cross Interferometer Configuration A						
5° Channel-Pair Phase Error	0.23°	0.45°	0	0.59°	3.5×10^{-4}	±47° ⁽⁴⁾
10° Channel-Pair Phase Error	0.45°	0.91°	0	1.17°	0.121	±20° ⁽⁴⁾
Configuration B						
5° Channel-Pair Phase Error	0.23°	0.45°	0	0.85°	3.5×10^{-4}	±34° ⁽⁴⁾
10° Channel-Pair Phase Error	0.45°	0.91°	0	1.70°	0.121	±10° ⁽⁴⁾

(1) Angular RMS Error Calculated at Low End of Band, F/Fmin = 1.0

(2) Probability Calculated at High End of Band, F/Fmin = 4.0

(3) Based on a Coning Error Plus RMS Error of 1°

(4) Based on an RMS Azimuth Error of 1°

tion, there is a coning bias error of 11.41 degrees that is independent of frequency or channel-pair phase error. Comparing these values with those of the cross interferometer, it is seen that the azimuth bias error is eliminated. However, the random error is not eliminated. As shown, each cross interferometer configuration (A and B) random error value, for a given channel-pair error, is larger than the horizontal interferometer random error value. For example, for a 5 degree channel-pair phase error, the horizontal interferometer introduces a 0.52 degree random error, while configuration A introduces a 0.59 degree error and configuration B a 0.85 degree error. For a 10 degree channel-pair error, these errors double in magnitude. The

horizontal interferometer introduces a 1.05 degree error, configuration A introduces a 1.17 degree error, and finally configuration B introduces a 1.70 degree error.

In column 6, the probability of ambiguity of each DF antenna configuration is presented for channel-pair phase errors of 5 degrees and 10 degrees. As seen, the probability of ambiguity of the cross interferometer is twice that of the horizontal interferometer. Also, there is also no difference between the probabilities of ambiguity for both cross interferometer configurations (A or B). For the purposes of discussion, if a maximum probability of ambiguity requirement of 0.1 is defined, it is seen that the horizontal interferometer configuration satisfies this requirement for a channel-pair phase error of 10 degrees. However, the cross interferometer requires a channel-pair phase error below 10 degrees. It turns out that if the channel-pair phase error were reduced to 9.7 degrees a 0.1 requirement could be met.

Finally, in the seventh column, the achievable elevation coverage, assuming the azimuth coverage is plus or minus 60 degrees, is presented. This is a measure of the operational flexibility of the DF antenna configuration. The elevation coverage is based on the maximum allowable azimuth error. The maximum azimuth error requirement is assumed to be 1 degree and is taken to be the sum of the bias coning error and the RMS azimuth error. As seen, the horizontal interferometer provides an elevation coverage of about plus or minus 6.2 degrees for a 5 degrees channel-pair phase and plus or minus 2.2 degrees for 10 degree channel-pair phase error. In comparison, the cross interferometer significantly increases the elevation coverage. Examining configuration B, the elevation coverage is plus or minus 10 degrees for a 10 degree channel-pair phase-error and plus or minus 34 degrees for a 5 degree channel-pair phase error. The elevation coverage can be further increased by increasing the vertical interferometer height as the data for configuration A demonstrates.

Not shown in table 1 are the elevation angular error characteristics of these DF antenna configurations. The horizontal interferometer provides no DF capability in the elevation plane; hence, this measure is not meaningful for this antenna. However, the cross interferometer does provide DF capability in the elevation plane. The vertical interferometer of configuration A provides angular error characteristics identical to those of the horizontal interferometer, while the configuration B vertical interferometer has increased elevation angular error due to the reduced height of the vertical interferometer. For a channel-pair phase error of 10 degrees at the low end of the band, configuration B provides an angular error of 1.16 degrees at an elevation angle of 0 degrees and 2.31 degrees at an elevation angle of plus or minus 60 degrees.

5.0 TECHNIQUES TO FURTHER REDUCE THE PROBABILITY OF AMBIGUITY

If the probability of ambiguity must be further reduced (for example, $P_a < 0.01$), it may be necessary to consider other techniques than relatively prime number set element spacings and low channel-pair phase error since there are practical limits to what can be achieved in terms of channel-pair phase error. These techniques were not considered in this study; but they should be mentioned. They are as follows: (1) a reduction of the frequency range, and (2) use of compact elements to reduce the minimum element spacing. In the first technique, the original 4:1 frequency range could be divided into two bands both covering a 2:1 frequency range. Essentially, there would now be two DF systems, each covering a 2:1 frequency range. This technique will dramatically reduce the probability of ambiguity of both the DF antenna configurations well below a probability of 0.01, at a penalty of increased equipment cost. In the second technique, the element diameter would be reduced using special loading materials within the spiral cavity such that the minimum spacing for channel

pair (1,2) would be 0.5λ at the high end of the band. This will reduce the probability of ambiguity well below 0.01. However, the penalty is a greatly reduced element efficiency at the low end of the band — possibly a 20 dB reduction or more.

6.0 CONCLUSIONS

The cross interferometer reduces coning error effects, improving DF system azimuth accuracy. The vertical array does contribute a random error to the azimuth accuracy and may increase the DF antennas' probability of ambiguity. However, increasing the vertical array height not only improves elevation accuracy but reduces this azimuth error. Further, the probability of ambiguity can be reduced by minimizing the channel-pair phase error.

Finally, an additional benefit of the cross interferometers is that it provides 2-dimensional DF information on the emitter's location, reducing the possibility of misassociating emitters with the wrong aircraft. In addition, it increases the operational flexibility of the DF system by increasing vertical elevation coverage capability of the DF system.

Concerning the specific cross interferometer design example examined, it is seen that it is necessary to have a channel-pair phase error below 10 degrees (actually 9.7 degrees or less to provide a probability of 0.1 or less. If these channel-pair phase error conditions are satisfied, the DF system can achieve a 4:1 bandwidth. Also, the azimuth angular error design objective of 1 degree or less for the cross interferometer is achievable, as shown in table 1, for both configuration A and B for a channel-pair phase error of 5 degrees. If the channel-pair phase error is 10 degrees, the angular error exceeds 1 degree only slightly for configuration A. Finally, the twin coverage design objectives of plus or minus 60 degrees for azimuth and at least plus or minus 30 degrees are mutually achievable for a channel-pair phase error of 5 degrees.

ACKNOWLEDGMENTS

The author would like to acknowledge the technical discussions and contributions of R. W. Bond, A. H. Corneliussen, C. G. Greenbaum, L. N. Honda, D. J. Kramer, D. Lambropoulos, W. R. Neal and G. A. Robertshaw of The MITRE Corporation and R. L. Goodwin of Naval Research Laboratory. In addition, the author would like to thank Joanne F. Richardson for her typing support.

REFERENCES

1. E. Jacobs and E. W. Ralston (Nov. 1981) "Ambiguity Resolution in Interferometry", IEEE Trans. Vol. AES-17, No. 6, pp 766-779.
2. G. V. Borgiotti (March 1977) "Maximum Theoretical Angular Accuracy of Planar and Linear Arrays of Sensors", IEEE Trans. Vol. AES-13, No. 2, pp 208-216.
3. W. B. Kendall (June 1965) "Unambiguous Accuracy of an Interferometer Angle-Measuring System", IEEE Trans. SET-11, pp 62-70.

Degree of Freedom Requirements for Angular Sector Nulling

Peter R. Franchi
USAF Rome Laboratory
Electromagnetics & Reliability Directorate
Hanscom AFB, MA 01731

Abstract

Analytic expressions are derived relating the degrees of freedom required to null the quiescent antenna pattern to an arbitrary desired level over a given angular sector. The results apply only to deterministic patterns and large arrays. In addition, the nulled region should not be near the main beam region. Earlier analytic results did not include the modifying effects of nearby nulls and are therefore not adequate for the general case. This inadequacy is overcome in this paper by including the effects of these nulls in an approach that is basically the original one.

Introduction

In a paper titled 'Constraints on Nulling Bandwidth' given several years ago at this conference, two closed form solutions were derived for the degree of freedom requirements needed to reduce the sidelobe level a specified amount over an arbitrary angular sector. A closer examination of the 'equal amplitude' solution indicates that it is indeed the equal amplitude answer but only for the case of large null depths. For sector null depths that are moderate or less, the amplitudes of the lobes within the sector are not all equal. The reason for this variation is that the lobe amplitudes are based on Tchebyscheff null spacings derived on the basis that all nulls outside the nulled sector are very far from that region. That condition is not met for moderate (or less) angular sector null depths. Figure 1 gives a graphic indication of this problem.

In order to compensate for the effects of the nulls that are close to the nulled sector, a new null placement generating function is introduced. This function determines the location of the nulls within the nulled sector, but, because the form of the correct generating function is unknown, an estimation for it is used. A single variable, 'a', controls the null spacing. As a varies from values of one to about two, the spacings change from Tchebycheff spacings to nearly equal spacings. The value of a is determined by setting the center lobe amplitude equal to the edge lobe amplitude inside the nulled sector.

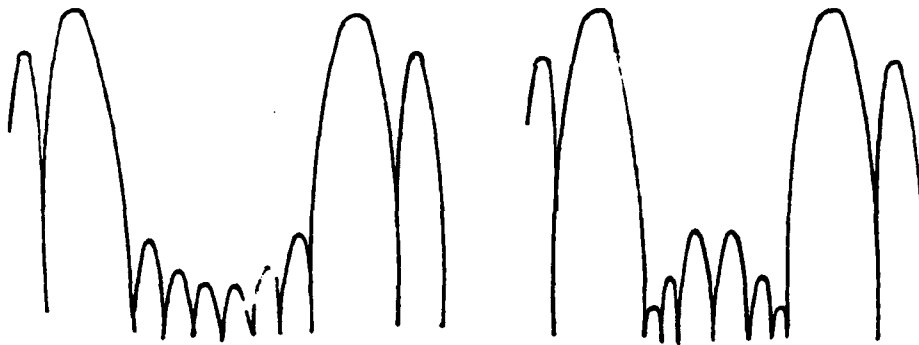


Figure 1: Uniform Spacing With Valley Tchebyshev Spacing With Peak

Review of Earlier Results

A brief review of the earlier work is given to aid understanding of these new modifications. The earlier work was based on the Schelkunoff polynomial with the pattern nulls spaced on a unit circle, (Figure 2).

The far field pattern can be written as

$$f(\theta) = \prod_{n=1}^{K-1} (s - s_n) \quad (1)$$

for an array of K elements and half wavelength spacings.

Simply stated, the field at any point, s , is proportional to the product of the distances from that point to all the nulls in the pattern. For our purposes, only the changes in the pattern over a small local region are needed. If the ratio of the perturbed pattern to the quiescent pattern (the pattern change) is taken, all the terms away from the perturbed region will cancel and so, the pattern change calculation is quite simple. If $s_0 = \lambda/D$, the beamwidth of a uniformly illuminated array and Δu is the angular sector null region, $\Delta u/s_0 = M$ or M is the number of beams in Δu . From Figure 3 and for the quiescent pattern,

$$SLL_0 \propto \prod_{n=1}^{K_0/2} \frac{s^2(2n-1)}{4} \quad (2)$$

where K_0 is a large number representing a null distance well removed from the region of interest. In the modified pattern, the new sidelobe level is SLL_1 .

$$SLL_1 \propto \prod_{n=1}^{N/2} \frac{s^2(2n-1)^2}{4} \prod_{\frac{N}{2}+1}^{K_0/2} s^2(2n-1)^2 \quad (3)$$

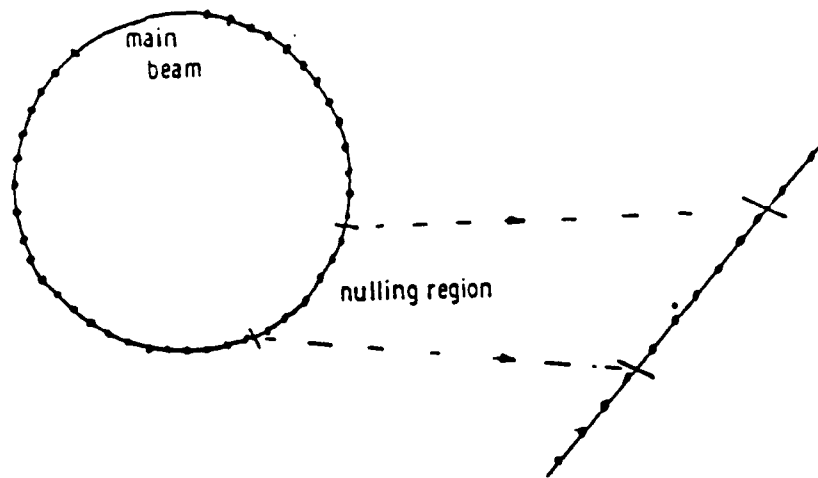


Figure 2: Schelkunoff Polynomial Unit Circle

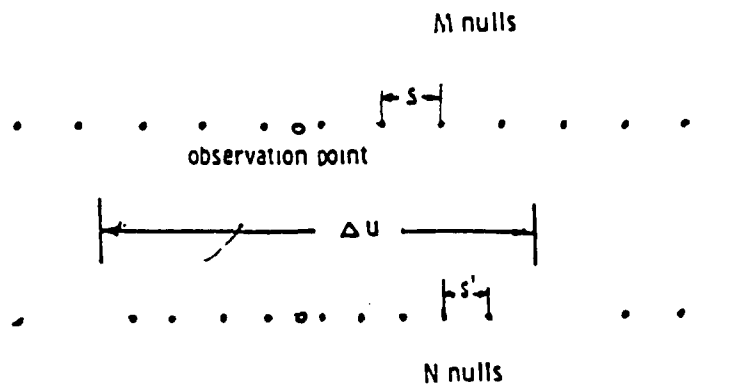


Figure 3: Null Location

where s' is the new spacing between nulls in Δu for N nulls or

$$\begin{aligned}
 \frac{Ns'}{SLL_1} &= \frac{Ms}{s'^N s^{K_0-N}} \\
 \frac{SLL_1}{SLL_0} = \Delta SLL &= \frac{s^{K_0}}{s'^N} \\
 &= (s'/s)^N \\
 &= (M/N)^N \quad \text{in field} \\
 &= (M/N)^{2N} \quad \text{in power} \\
 &= -20N \log(M/N) \quad \text{in dB.}
 \end{aligned} \tag{4}$$

In a similar manner, but for nulls set for Tchybshev spacings, N nulls are located at

$$\begin{aligned}
 \cos\left(\frac{\pi(2n+1)}{2N}\right) \quad 1 \leq n \leq N \quad \text{and } N \text{ even} \\
 \cos\left(\frac{\pi n}{N}\right) \quad N \text{ odd}
 \end{aligned} \tag{5}$$

Then

$$\Delta SLL = \left(\frac{Me}{N2}\right)^{2N} \tag{6}$$

Note the angular sector here is defined from just beyond the edge null on one side to the same point on the other side. In the earlier paper the angular width was defined by the distance between the two outer lobe peaks.

The Approximate Result

The function that gives the null locations (approximately) is

$$\Delta u \frac{\cos\left(\frac{\pi(2n-1)}{2Na} + \frac{\pi(a-1)}{2a}\right)}{\cos\left(\pi \frac{2n-1}{2a}\right)} \tag{7}$$

with a , the control variable. For $a = 1$ this reduces to

$$\Delta u \cos\left(\frac{\pi(2n-1)}{2N}\right) \tag{8}$$

and for $a = 2$

$$\Delta u \frac{\cos\left(\frac{\pi n}{2N}\right)}{\cos(\pi/4)} \tag{9}$$

and this expression gives null spacings that are nearly equal. With a similar procedure as the previous section, the center lobe expression is

$$SLL_c = \prod_{n=1}^{N/2} \left(\frac{\Delta u}{2d}\right)^2 \cos^2\left(\frac{\pi(2n-1)}{2aN} + \frac{\pi(a-1)}{2a}\right) \prod_{n=1+N/2}^{K_0} (s/2)^2 (2n-1)^2 \tag{10}$$

The edge lobe is

$$SLL_e = \prod_{n=1}^N \frac{\Delta u}{2d} \left[\cos\left(\frac{\pi}{Na} + b\right) - \cos\left(\frac{\pi(2n-1)}{2Na} + b\right) \right] \prod_{n=1+N/2}^{K_0} (s/2)^2 ((2n-1) - M^2) \quad (11)$$

where $b = \frac{\pi(a-1)}{2a}$.

Setting $SLL_c = SLL_e$ (the center lobe magnitude = the edge lobe magnitude)

$$\frac{\prod_{n=1}^N [\cos(\frac{\pi}{Na} + b) - \cos(\frac{\pi(2n-1)}{2Na} + b)]}{\prod_{n=1}^N \cos(\frac{\pi(2n-1)}{2Na} + b)} \approx \left[\frac{N + M/2 + 2}{N - M/2} \right]^{M/2} \quad (12)$$

The numerator can be approximated by $(.5)^N$ and the denominator by

$$\frac{2}{2^{aN} \cos^{(a-1)N}(\frac{\pi(a-1)}{4a})} \quad (13)$$

So

$$2^{a-1} \approx \left[\frac{1 + \frac{M}{2N}}{1 - \frac{M}{2N}} \right]^{M/2} \quad (14)$$

for large N and

$$a = 1 + 2.885 \left(\left(\frac{M}{2N} \right)^2 + 1/3 \left(\frac{M}{2N} \right)^4 + 1/5 \left(\frac{M}{2N} \right)^6 + \dots \right) \quad (15)$$

The ratio of the perturbed pattern center lobe to the quiescent pattern center lobe gives

$$\Delta SLL \cong \frac{\prod_{n=1}^{N/2} \frac{\Delta u}{2d} \cos^2(\frac{\pi(2n-1)}{2Na}) + \frac{\pi(a-1)}{2a} \prod_{n=1+N/2}^{K_0/2} \frac{s}{2} (2n-1)^2}{\prod_{n=1}^{K_0/2} \frac{s}{2} (2n-1)^2} \quad (16)$$

After several minor approximations and after the expression has been squared to put it in power terms

$$\Delta SLL \cong \left[\frac{e}{2^a \cos(\frac{\pi(a-1)}{2a})} \frac{M}{N} \right]^2 N \quad (17)$$

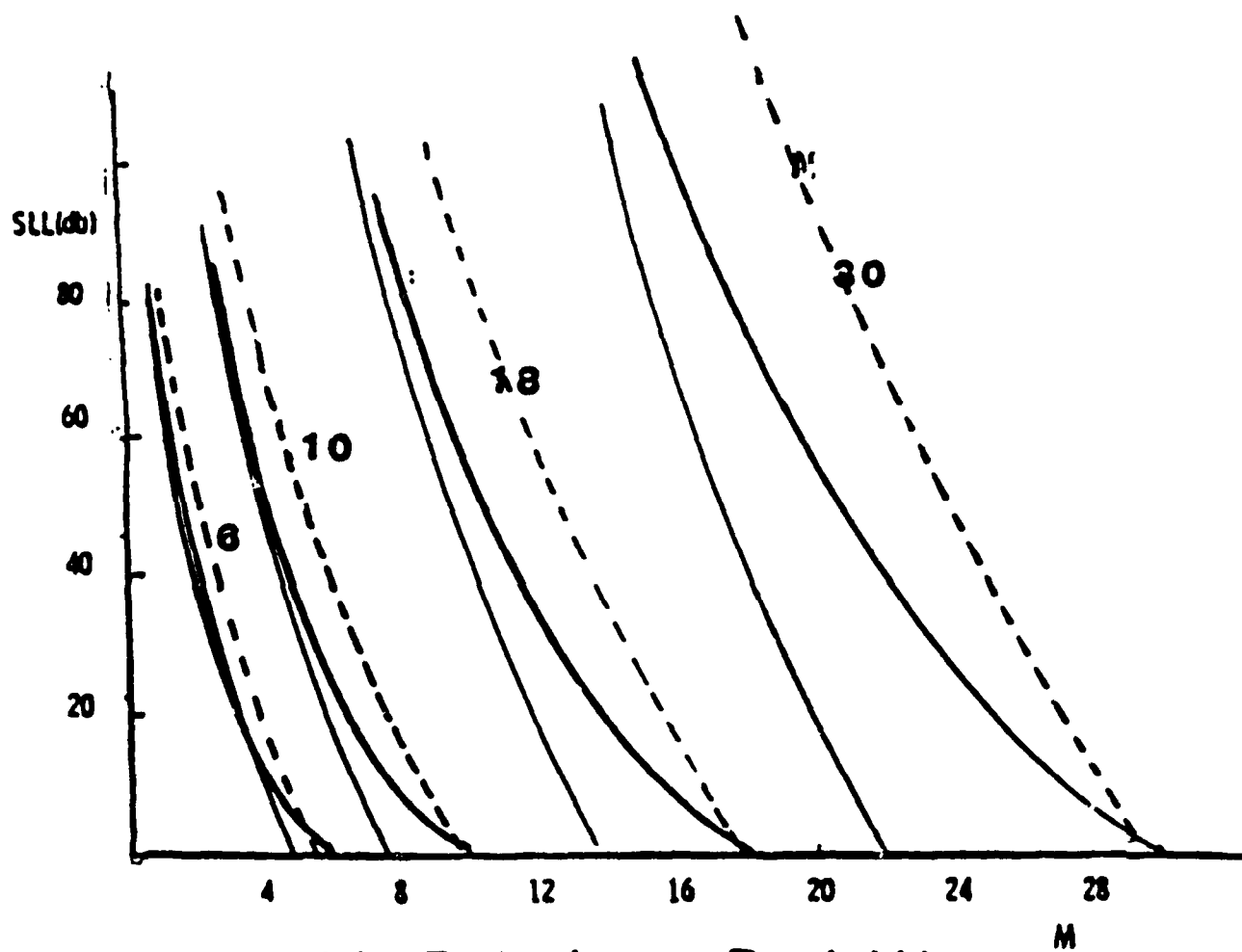
A polynomial match for the term, $\cos[\frac{(a-1)\pi}{2a}]$ gives

$$\cos\left(\frac{(a-1)\pi}{2a}\right) \cong 1 - .243(a-1) - .0502(a-1)^2 \quad (18)$$

So

$$\Delta SLL \cong \left[\left(\frac{e}{2} \right)^{1-(\frac{M}{N})^2} \frac{M}{N} \right]^{2N} \quad (19)$$

This is the approximate answer. The derivation outlined above is greatly condensed, with many approximations not mentioned, but the full derivation is too long for this paper.



Sidelobe Reduction vs. Bandwidth

Figure 4: Comparison of the Three Solutions

Comparisons

In Figure 4 a comparison is shown between this approximate solution and the equally spaced nulls and the Tchebyshev spaced nulls solutions. The dark line shows the approximate solution, the light line shows the Tchebyshev solution, and the dotted line the equal spacing solution. It is clear that the approximate solution becomes the equally spaced answer as N becomes equal to M , as might be expected. On the other hand, when N is much greater than M , the approximate solution approaches the Tchebyshev answer.

Another important point is apparent from these plots. For large values of N or M , the differences between these answers is quite large so using the equally spaced null answer, for example, to obtain a rough estimate of the degree of freedom requirements could lead to a serious underestimation. A comparison with the earlier work of H. Steyskel is shown in Figure 5. The agreement is quite good.

Summary

A closed form solution is given relating the width and depth of an angular null sector to the degrees of freedom required. This solution is only an approximate one that applies to regions that are at least a sector null width away from the mainbeam. While this answer only applies to theoretical array patterns, it will serve as a conservative estimate for errored array patterns.

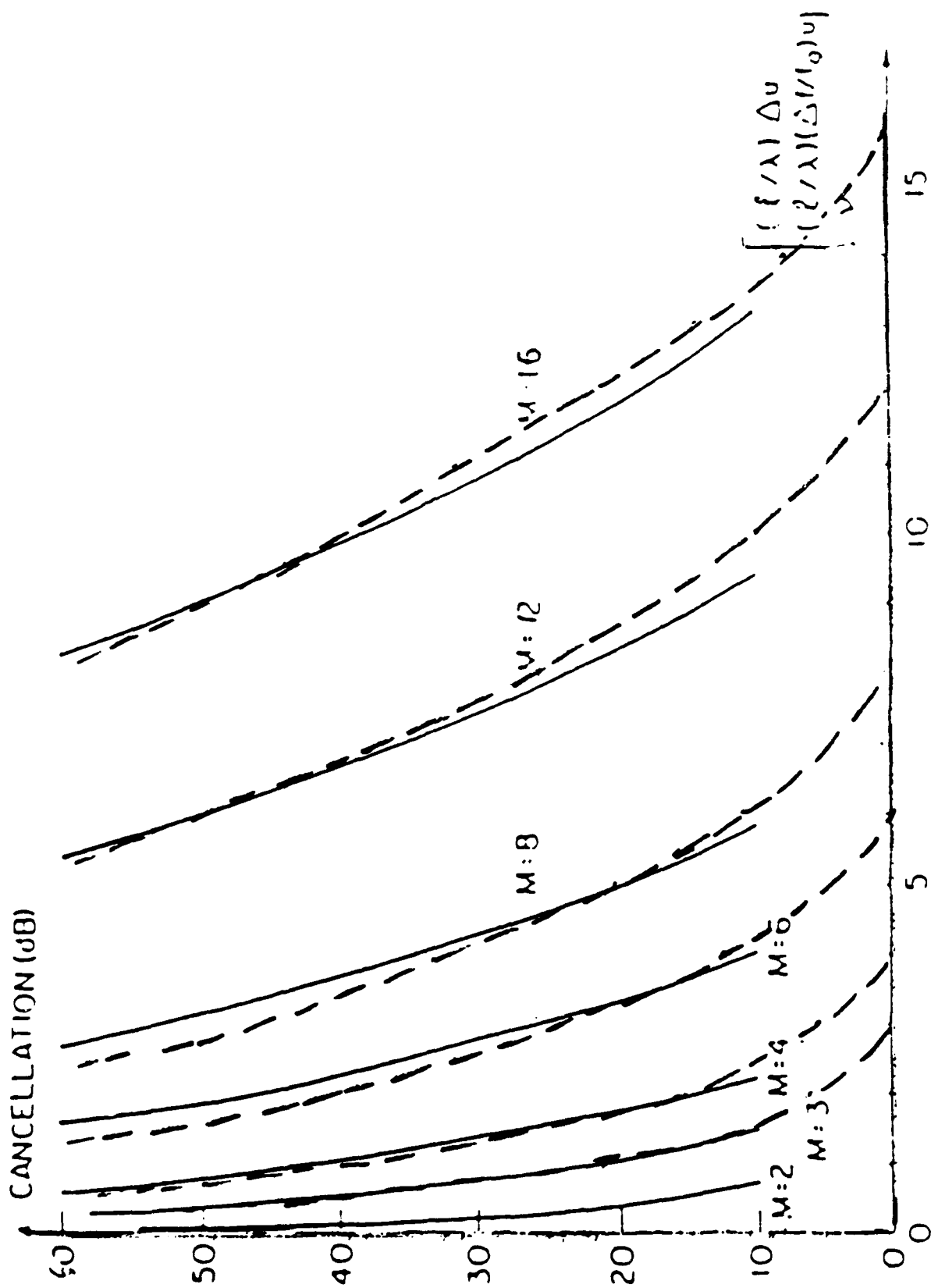


Figure 5: Comparison of Approximate Solution With H. Steyskel Results