A TIME-HARMONIC ELECTROMAGNETIC ANALYSIS

OF

SHIELDED MICROSTRIP CIRCUITS

by

James Clinton Rautio
B.S., Cornell University, 1978
M.S., University of Pennsylvania, 1981

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of Syracuse University
August 1986

Approved R F Harrington

Date 8/18/86

ABSTRACT

The precise analysis of microstrip circuits at microwave and millimeter wave frequencies has become critically important in the design of monolithic microwave GaAs integrated circuits. This is because the usual post-fabrication 'tweaking' of circuit response is rarely feasible and the resulting lengthy repeated design modifications drive design costs to unacceptable levels.

This dissertation will describe a Galerkin method of moments analysis for microstrip circuits of arbitrary planar geometry enclosed in a rectangular conducting box. The technique entails a time-harmonic electromagnetic analysis evaluating all fields and surface currents. This analysis is suitable for the accurate verification of designs prior to fabrication.

A computer program has been written in Pascal on an IBM-PC implementing the analysis. Several microstrip circuits were built and measured. The resulting agreement with calculations verifies the accuracy of the analysis.

A TIME-HARMONIC ELECTROMAGNETIC ANALYSIS

OF

SHIELDED MICROSTRIP CIRCUITS

by

James Clinton Rautio
B.S., Cornell University, 1978
M.S., University of Pennsylvania, 1981

DISSERTATION

Submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of Syracuse University
August 1986

Approved R F Harrington

Date 5/18/86

ACKNOWLEDGEMENTS

The author expresses thanks to Dr. Roger Harrington whose guidance was critical to the success of this project and to General Electric, Electronics Laboratory, for funding the fellowship which made this work possible and for providing access to the HP8510 automated network analyzer used to make measurements.

CONTENTS

Chapter 1

INTRODUCTION

Microstrip circuits form an integral part of most current microwave systems. The attractiveness of implementing a microwave network in microstrip results from the ease with which an intricate circuit can be manufactured. If the increased losses (as compared with, for example, waveguide) can be tolerated, the microstrip implementation of a circuit is frequently preferred.

Manufacture of a microstrip circuit starts with a substrate, a rectangular slab of dielectric. Common dielectric materials include Teflon, sapphire, quartz, Alumina, silicon and gallium arsenide. Substrates with large dielectric constants are frequently selected so as to confine most of the fields within the body of the substrate.

Initially, the substrate is completely covered by a thin conductor commonly called metalization. The metalization may have several layers of different metals to enhance substrate adhesion.

The top side of the substrate is then covered by a photoresist which is exposed with the desired microstrip circuit pattern. The unexposed (or the exposed) portions of the metalization are then etched away leaving the desired circuit. The bottom side of the substrate remains metalized, forming a ground plane.

Frequently, additional components (e.g., transistors, capacitors, resistors) are then soldered or bonded to the surface of the microstrip circuit. Unless the microstrip circuit is to be used as an antenna, the final step of manufacture is to place the circuit in a shielding, conducting box. Coaxial connectors penetrating the shield

provide connections to the circuit. If this final step is not performed, unwanted interactions may develop when the circuit is placed in a system.

At this point the circuit response is measured, usually as a function of frequency. Since the initial circuit response is rarely satisfactory, the circuit must be 'tweaked'. Tweaking consists of removing metalization with a diamond scribe, or adding metalization by means of silver epoxy or by adding wire bonds connecting small patches of metalization included on the microstrip substrate for just that purpose.

In some cases, tweaking is not feasible. For example, with monolithic microwave integrated circuits, the substrate (usually gallium arsenide) and circuit are so fragile and so small (geometries on the order of microns), that any physical modification of the circuit is likely to result in the destruction of the circuit. In this case, the only alternative is to repeat the entire circuit design and manufacture. This can cost several tens of thousands of dollars and occupy a significant fraction of a year.

The precise analysis of microstrip circuits is the objective of the work described in this paper. Such an analysis will allow any circuit modification to be performed prior to manufacture resulting in the reduction or elimination of post-manufacture tweaking and redesign.

Current microstrip analyses start with a model based on circuit theory [1]. This usually includes TEM transmission lines, stubs, capacitors, inductors and resistors. More advanced analyses will

include transmission line loss and dispersion modeled by closed form approximations based on more extensive electromagnetic analyses. Microstrip discontinuities (e.g., step in line width, junction of three or more lines, etc.) may also be included in a similar manner. Restrictions on the closed form approximations are sometimes stringent.

The circuit model of a large microstrip network can become complicated. As such, interconnection errors in the model may be difficult to detect and remove. In addition, while the resulting accuracy is sufficient to approximately realize a desired circuit response, the circuit model accuracy falls short of that which would be desirable for many tasks.

Most other work to date has concentrated on development of numerically efficient circuit models of microstrip elements. Such models can be included in extensive existing microwave circuit analysis programs in hopes of improving the accuracy of the microstrip circuit representation. The concentration on numerical efficiency results from the requirement for automatic circuit optimization. Optimization algorithms require the repeated analysis of a circuit, changing parameters of the circuit each time. An analysis which can not be performed quickly, such as an electromagnetic analysis, is unacceptable in an optimization loop.

Electromagnetic analyses have been used to provide data points for fitting approximate closed form expressions. Also, due to the complexity of an electromagnetic analysis, the vast majority of work has been restricted to, or closely related to, two dimensional

microstrip structures (i.e., no variation of geometry along one axis).

The technique presented here will take a different tack. A precise three dimensional electromagnetic analysis of arbitrary microstrip geometries is the primary objective. While this essentially precludes using the analysis in an optimization loop (at the present time), the analysis will still be sufficiently fast that meaningful results can be obtained, even on a personal computer, for circuits of moderate complexity.

Chapter 2

PREVIOUS WORK

Most current applied microstrip circuit analysis is based on circuit theory approximations [1] to the actual microstrip structure. For example, a microstrip line can be approximated by a TEM transmission line [2]. More elaborate analyses will include closed form approximations to circuit theory models of various microstrip discontinuities [1]. The approximations are typically based upon an electromagnetic analysis. Some circuit theory analyses are more closely related to an underlying electromagnetic analysis [3], [4] in order to achieve greater accuracy. In all cases, the objective is to reduce a microstrip analysis problem to a circuit theory problem with emphasis on speed so that the analysis may be included in an optimization loop. Electromagnetic analysis of an entire microstrip circuit, without resort to circuit theory, has not yet been pursued due to excessive numerical requirements.

Here, we will explore just such an electromagnetic analysis of arbitrary microstrip structures. The numerical requirements will be large but manageable, even on current personal computers. The time required for analysis will preclude its inclusion in an optimization loop, however, the accuracy provided will justify the use of the analysis prior to circuit fabrication. We will also note that in many other fields (e.g., structures, thermodynamics, fluid flow, seismology, etc.) the value of an accurate, numerically intensive, general purpose computer analysis is well established [5].

The analysis follows from section 8-11 of [6] which describes a

technique for the analysis of planar waveguide probes. The technique is modified for a shielded microstrip box and is used to calculate the fields in the box due to any rectangular patch of current. These results are then applied in a Galerkin implementation of the method of moments [7].

The analysis described here is closely related to the spectral domain technique [9]. The spectral domain technique was first applied in the analysis of microstrip dispersion [10], [11].

A number of techniques, including the spectral domain technique, have been applied to the analysis of three dimensional microstrip resonators [12] - [38]. The vast majority of this work has been directed toward the evaluation of resonant frequencies and modes. Little work has been reported concerning the analysis of the forced response of microstrip circuits. Potential difficulty evaluating the forced response of a shielded microstrip circuit due to difficulty in modeling sources has been mentioned [9].

One exception is a time domain technique which models the microstrip circuit as a mesh of transmission lines with current on a line representing the magnetic field and line voltage representing electric field [31] - [33]. The volume of the microstrip box is divided into a mesh of these transmission lines. For each point in the mesh, six transmission lines must be used in order to represent the six possible vector components. While this can result in extremely large numerical requirements, it can provide detailed information on the resulting fields and currents.

Another forced response analysis [34] models a microstrip circuit

using planar waveguide (the microstrip line is replaced by a waveguide of wider width with magnetic walls along the sides and electric walls on top and bottom). In this case, the surface of the microstrip is divided into a mesh with inductors connecting nodes and capacitors connected in shunt to ground. This technique can be used when the substrate is thin with respect to the wavelength and when coupling between microstrip edges may be neglected. Since only the surface of the equivalent planar waveguide is divided into a mesh, the numerical requirements are reduced.

More typical are analyses for the resonant frequencies of microstrip resonators. These are usually accomplished in the spectral domain using techniques related to [12]. For example, [13] describes the analysis of triangular microstrip resonators. Full wave (as opposed to quasi-static) analysis of rectangular microstrip resonators is the topic of [15] - [17], [19] and [20]. Rectangular slot resonator analysis is presented in [24]. Some of the above analyses are for microstrip resonators in a rectangular waveguide tube, while others are for resonators completely shielded in a box. These analyses can be used in the evaluation of microstrip discontinuities [38].

Several other analyses [14], [18], [21] approach the resonator problem with a quasi-static assumption.

A number of papers have dealt with resonators of arbitrary shape [25] - [37]. Farrar and Adams [25] describe the quasi-static approach, while Jansen [26] - [29] details a full wave approach. Jansen uses a polynomial to represent the resonator surface current

and applies a spectral domain solution. A finite difference method, the method of lines, has also been applied to microstrip resonators [30] - [33]. This technique uses the difference in field between adjacent points on the substrate surface to represent derivatives of the field. The initial phase of the analysis requires points covering the entire substrate surface while the final phase need consider only those points on metalization.

The frequency domain analysis of the forced response of shielded microstrip circuits merited little attention in the literature reviewed here. It is not known why the above resonator analyses were not applied to a forced response analysis.

Chapter 3

METHOD OF ANALYSIS

The rectangular conducting box will be treated as two separate waveguides joined at z=h (Figure 3.1 and 3.2). We will work with waveguides of identical dimensions joined with no offset. The waveguide for z<h will be short circuited at z=0 and will be filled with a dielectric with dielectric constant ϵ_1. This region will be referred to as region 1. The waveguide for z>h will be short circuited at z=c and will be filled with a dielectric with dielectric constant ϵ_o. This will be referred to as region 0. While region 0 is usually air, ϵ_o may take on a value other than that of free space. Using regions 1 and 0, rather than regions 1 and 2, simplifies subsequent notation. The analysis is extended in a straightforward manner to multi-layered geometries.

The dielectric in region 1 will be called the substrate. Microstrip circuitry is realized in printed circuit fashion on the surface of the substrate. The analysis we will pursue divides the microstrip circuitry (the metalization on the surface of the substrate) into rectangular subsections. An appropriate form for the surface current distribution on the surface of each subsection is assumed. Then each subsection is taken individually and the electric fields tangent to the surface of the substrate due to a unit current on that subsection are calculated. After the tangential fields due to each subsection have been calculated, the magnitude of current on each subsection is selected such that an integral of the total tangential electric fields over each subsection goes to zero. A source is

modeled by a subsection on which the tangential electric field goes to a value other than zero.

All boundary conditions except tangential electric field on the metalization are met in the initial calculation of the electric field from each individual subsection. By selecting the appropriate currents for each subsection, the tangential electric field boundary condition on the metalization is also satisfied, providing us with the solution. Once all the currents are determined, the N-port circuit parameters follow immediately.

3.1 Expansion of the Green's Function in Terms of Orthonormal Waveguide Modes

The tangential (or transverse to z) fields in a given region due to current on a given subsection will be expressed as a sum of homogeneous waveguide modes. Modes transverse to z will be used with a z dependence such that the boundary conditions at z=0 for region 1 and at z=c for region 0 are met. Expressions for the tangential fields will be written as a weighted sum of these modes

$$
\begin{aligned}
E_t^1 &= \sum V_i \frac{\sin(K_{iz}^1 z)}{\sin(K_{iz}^1 h)} e_i \\
H_t^1 &= \sum V_i Y_i^1 \frac{\cos(K_{iz}^1 z)}{\sin(K_{iz}^1 h)} h_i \\
E_t^0 &= \sum V_i \frac{\sin[K_{iz}^0 (c-z)]}{\sin[K_{iz}^0 (c-h)]} e_i \\
H_t^0 &= \sum V_i Y_i^0 \frac{\cos[K_{iz}^0 (c-z)]}{\sin[K_{iz}^0 (c-h)]} h_i
\end{aligned}
\qquad (3.1)
$$

where i = Summation index over all modes TE_z and TM_z.
V_i = the modal coefficient (weight) of the i^{th} mode.
Y_i = the admittance of the i^{th} mode as follows:

$$Y_i^{1TE} = jK_{iz}^1/(\omega\mu_1)$$

$$Y_i^{1TM} = j\omega\epsilon_1/K_{iz}^1$$

$$Y_i^{0TE} = -jK_{iz}^0/(\omega\mu_0)$$

$$Y_i^{0TM} = -j\omega\epsilon_0/K_{iz}^0$$

$$K_{iz}^1 = -\sqrt{K_1^2 - K_x^2 - K_y^2}$$

$$K_{iz}^0 = +\sqrt{K_0^2 - K_x^2 - K_y^2}$$

$K_x = M\pi/a$, for rectangular waveguide.

$K_y = N\pi/b$, for rectangular waveguide.

$$K_1 = \omega\sqrt{\mu_1 \epsilon_1}$$

$$K_0 = \omega\sqrt{\mu_0 \epsilon_0}$$

Note that K_z for region 1 is taken as the negative square root and for region 0 is the positive square root. Also, the modal admittances are the admittances of the standing wave modes rather than those of the usual traveling wave modes (they differ by the constant

j). The constant in the denominators guarantees continuous tangential electric field at the surface of the substrate (z=h), i.e., we are assuming there is no magnetic current.

If other than a short circuit forms the top (z=c) or bottom (z=0) of the shielding box, then only the z dependence (and corresponding normalizing constant) of the above equations need be modified.

The $\mathbf{e}_i$ and $\mathbf{h}_i$ are the orthonormal mode vectors. The mode vectors are determined by the geometry of the shielding waveguide. We will consider only rectangular waveguide here. For rectangular waveguide, the mode vectors may be written as follows

$$\begin{aligned}
\mathbf{e}_i^{TE}(x,y) &= N_1 g_1 \mathbf{u}_x - N_2 g_2 \mathbf{u}_y \\
\mathbf{e}_i^{TM}(x,y) &= N_2 g_1 \mathbf{u}_x + N_1 g_2 \mathbf{u}_y \\
\mathbf{h}_i &= \mathbf{u}_z \times \mathbf{e}_i \qquad \mathbf{e}_i = -\mathbf{u}_z \times \mathbf{h}_i
\end{aligned} \tag{3.2}$$

where $g_1 = \cos(K_x x)\sin(K_y y)$

$g_2 = \sin(K_x x)\cos(K_y y)$

$N_1 = \sqrt{2/ab}$, M=0 and N≠0

$= 0$, M≠0 and N=0

$= 2(N/b)\sqrt{ab/(N^2a^2+M^2b^2)}$, M≠0 and N≠0

$N_2 = 0$, M=0 and N≠0

$= \sqrt{2/ab}$, M≠0 and N=0

$= 2(M/a)\sqrt{ab/(N^2a^2+M^2b^2)}$, M≠0 and N≠0

The functions g_1 and g_2 provide the x and y dependence of the mode vectors. The constants N_1 and N_2 normalize the modes such that the square (the mode vector dotted with itself) of any mode vector

integrated over the x-y cross section of the guide will be unity. Since the mode vectors are also orthogonal, we have an orthonormal basis for the expansion of any field in the waveguide due to current on the surface of the substrate.

The superscripts TE and TM mean that the index i is limited to the TE or TM modes respectively. The functions g_1, g_2, N_1 and N_2 are implicitly dependent on the index i through the conventional mode numbers M and N. Note that the M=0, N=0 mode need not be included as all current is transverse to the z direction [39], [40].

If a different geometry is selected for the waveguide shield, only the above mode vectors need be changed.

Given a specific current distribution, $\mathbf{J}(x,y) = J_x\mathbf{u}_x + J_y\mathbf{u}_y$, on the surface of the substrate, we must determine the modal coefficients, the V_i, of the field generated by that surface current. This is accomplished by setting the discontinuity in tangential $\mathbf{H}$ equal to the assumed surface current distribution. Then, using the orthogonality of the modal vectors, we may determine the V_i of the field generated by the current.

$$\mathbf{J}_s = \mathbf{u}_z \times \left[\mathbf{H}_t^o \big|_{z=h} - \mathbf{H}_t^i \big|_{z=h} \right]$$

Substituting the modal expansion for $\mathbf{H}_t$ into the above expression yields

$$\mathbf{J}_s = \sum V_i \hat{Y}_i (\mathbf{u}_z \times \mathbf{h}_i)$$

$$\text{where } \hat{Y}_i = Y_i^o \operatorname{ctn}[K_{iz}^o (c-h)] - Y_i^i \operatorname{ctn}(K_{iz}^i h)$$

The admittance $\hat{Y}_i$ is the parallel connection of the admittances of the two shorting planes at z=0 and z=c transformed back to the substrate surface, z=h. This expression can be rewritten as

$$J_s = -\sum V_i \hat{Y}_i \mathbf{e}_i$$

Dotting both sides with $\mathbf{e}_j$, integrating over the waveguide cross section, and noting the orthonormality of the modal vectors, we find

$$V_i = -\hat{Z}_i \iint J_s \cdot \mathbf{e}_i \, ds \qquad (3.3)$$

where $\hat{Z}_i$ is the inverse of the $\hat{Y}_i$ described above.

Substitution of (3.2) and (3.3) into (3.1) will yield the tangential fields everywhere in the waveguide. Specialization of (3.3) to a delta function for J_s will provide the Green's function in the 'spatial' domain for current on the surface of the substrate. The Green's function is a cosine and sine series in two dimensions with the coefficients of the series representing the Green's function in the 'spectral' domain.

Since we wish to have quick access to the actual fields, rather than the transform of the fields, we will remain in the spatial domain. Remaining in the spatial domain also appears to involve little additional computational expense and will ease the task of maintaining conceptual clarity.

In what follows, we will determine the V_i for various current

distributions defined over rectangular subsections. The fields due to these currents will be applied in a method of moments solution.

The derivation of the modal expansions presented above follows a similar derivation in Harrington [6], section 8-11, pages 425-428, for waveguide probes.

3.2 Evaluation of the V_i For Various Current Distributions

Evaluation of the V_i will require the evaluation of surface integrals of the current distribution dotted with a mode vector. We will consider only current distributions over rectangular subsections. The current distributions will be symmetrical with respect to the center of the subsection. Further, only current distributions which are separable with respect to x and y will be evaluated. One component of current, either x or y, will be evaluated at a time.

Since we will be working with separable current distributions, the integrals of (3.3) reduce to the product of two one dimensional integrals. We will evaluate the one dimensional integrals for a number of cases.

While the evaluation of the integrals is tedious, the assumption of current distribution symmetry will provide a simple result. Specifically, the result of each integral will be a constant, dependent only on the dimensions of the subsection, multiplied by the appropriate component of the mode vector evaluated at the center of the subsection. The simplicity of this result will provide considerable flexibility in the choice of expansion functions.

3.2.1 The Rectangular Pulse Function

The rectangular pulse function is the simplest pulse function that we shall consider. The rectangular pulse is defined as (Figure 3.3)

$$f(x) = 1, \quad x_o - \Delta x/2 \leq x < x_o + \Delta x/2$$
$$= 0, \quad \text{otherwise}$$

When the rectangular pulse is used as part of a current distribution, we will require the evaluation of

$$F_C = \int f(x)\cos(Kx)dx \qquad \text{and} \qquad F_S = \int f(x)\sin(Kx)dx$$

The constant K will be the wavenumber corresponding to the variable of integration, for example, $M\pi/a$. Evaluation of the above integrals is straightforward yielding

$$F_C = \frac{2}{K}\sin(K\Delta x/2)\cos(Kx_o), \quad K\neq 0$$
$$F_S = \frac{2}{K}\sin(K\Delta x/2)\sin(Kx_o), \quad K\neq 0$$
$$F_C = \Delta x \quad \text{and} \quad F_S = 0, \quad K=0$$

Note that the constant which depends on the subsection dimensions is the same for both cases, as one would expect from the nature of the functions being integrated. In subsequent sections, we will refer to

that constant as $G(\Delta x)$. In some cases, the integrand will be a function of y instead of x. In such a case we will refer to $G(\Delta y)$. For the rectangular pulse we have

$$G(\Delta x) = \frac{2}{K}\sin(K\Delta x/2), \quad K \neq 0$$

$$= \Delta x, \quad K=0$$

We will use F_c, F_s, $G(\Delta x)$ and $f(x)$ as a generic notation. The specific functions which they represent will depend on the type of pulse being considered.

3.2.2 The Triangular Pulse Function

For reasons to be detailed later, it is desirable to investigate higher order pulse functions. We will next consider the triangle pulse function. The triangle pulse may be viewed as the convolution of two of the rectangular pulses considered above. It is defined as

$$f(x) = \frac{x-x_o}{\Delta x} + 1, \quad x_o - \Delta x \leq x < x_o$$

$$= \frac{x_o - x}{\Delta x} + 1, \quad x_o \leq x < x_o + \Delta x$$

$$= 0, \quad \text{otherwise}$$

The integrals of interest are

$$F_C = \int f(x)\cos(Kx)dx \qquad \text{and} \qquad F_S = \int f(x)\sin(Kx)dx$$

The indicated integrations (over the entire domain of f(x)) are tedious but, again, the result is simple:

$$F_C = G(\Delta x)\cos(Kx_o) \qquad \text{and} \qquad F_S = G(\Delta x)\sin(Kx_o)$$

$$\text{with } G(\Delta x) = \frac{2}{\Delta x K^2}\left(1-\cos(K\Delta x)\right), \quad K\neq 0$$

$$= \Delta x, \quad K=0$$

3.2.3 The Parabolic Pulse Function

A further convolution of the triangular pulse with a rectangular pulse yields a piecewise parabolic pulse. When the x-axis is scaled so that the area of the pulse is Δx (the same as the other pulse types), we have

$$f(x) = 2\left(1+\frac{x-x_o}{\Delta x}\right)^2, \quad x_o-\Delta x \le x < x_o-\Delta x/2$$

$$= 1-2\left(\frac{x-x_o}{\Delta x}\right)^2, \quad x_o-\Delta x/2 \le x < x_o+\Delta x/2$$

$$= 2\left(1-\frac{x-x_o}{\Delta x}\right)^2, \quad x_o+\Delta x/2 \le x < x_o+\Delta x$$

$$= 0, \quad \text{otherwise}$$

As before, the integrals of interest are

$$F_C = \int f(x)\cos(Kx)dx \qquad \text{and} \qquad F_S = \int f(x)\sin(Kx)dx$$

Evaluation of these integrals must be performed over three separate ranges. Since f(x) is symmetric about x_o, the result simplifies to

$$F_C = G(\Delta x)\cos(Kx_o) \qquad \text{and} \qquad F_S = G(\Delta x)\sin(Kx_o)$$

$$\text{with } G(\Delta x) = \frac{8}{\Delta x^2 K^3}\left(2\sin(K\Delta x/2) - \sin(K\Delta x)\right), \quad K \neq 0$$

$$= \Delta x, \quad K=0$$

Note that even the piecewise parabolic pulse, has a result very nearly as simple as the rectangular pulse result.

3.3 General Expressions for V_i

We will use products of the above functions; the rectangular, triangular and parabolic pulses; to form the desired current distributions. For example, two rectangular pulses, one a function of x and the other a function of y, would be multiplied to represent a uniform current distribution over a rectangular subsection. A more useful example is a rectangular pulse as a function of x multiplied by a triangular pulse which is a function of y. This would provide a 'roof top' current distribution as shown in Figure 3.4. If the triangular pulse is replaced with a parabolic pulse, the flat portions of the roof would become curved.

The roof-top distribution in Figure 3.4 depicts the current in a

rectangular subsection. The base of the three dimensional drawing represents the rectangular subsection. The height of the figure above the base is proportional to the magnitude of the current density at that point. The current flows in the direction of the triangle. Figure 3.5 shows two roof-top functions which have been placed on adjacent, overlapping rectangles. Note that the sum of the currents provides a piecewise linear approximation to the actual surface current in the direction of current flow and a step approximation to the current in the lateral direction.

We will not reference specific pulse functions in the derivations that follow in this section. Rather, we will use the generic notation introduced above. Specifically, $f(x)$ will denote a pulse function which is a function of x. Similarly, $f(y)$ will denote a pulse function which is a function of y. $G(\Delta x)$ is the constant, derived above, which is obtained when $f(x)$ is multiplied by a sine or cosine and integrated over the domain of $f(x)$. Similarly for $G(\Delta y)$. We will indicate distinct $f(x)$ and $G(\Delta x)$ functions by means of subscripts.

In subsequent derivations it will be necessary to differentiate between source subsections and field subsections. For this analysis, we will consider a surface current on only one subsection at a time. That subsection will be termed the source subsection. With current on one subsection, we must find the electric field generated by the current on any other subsection. If we are considering the field on a subsection, it will be called a field subsection. When we are considering the field on the source subsection itself, the field subsection and the source subsection are the same. We will indicate

quantities relating to source subsections with a prime. Quantities relating to a field subsection will remain unprimed. Thus $G'(\Delta x)$ will refer to a source subsection and $G(\Delta x)$ will refer to a field subsection.

A primed modal vector will be used to indicate that the modal vector is to be evaluated at the center of the source subsection, e.g., $g_i' = g_i(x_o, y_o)$.

To evaluate the V_i, the modal coefficients, we will use (3.3)

$$V_i = -\hat{Z}_i \iint \mathbf{J}_s \cdot \mathbf{e}_i \, ds \tag{3.3}$$

The modal vectors, $\mathbf{e}_i$, are repeated here for convenience

$$\begin{aligned} \mathbf{e}_i^{TE}(x,y) &= N_1 g_1 \mathbf{u}_x - N_2 g_2 \mathbf{u}_y \\ \mathbf{e}_i^{TM}(x,y) &= N_2 g_1 \mathbf{u}_x + N_1 g_2 \mathbf{u}_y \end{aligned} \tag{3.2}$$

We will first consider an x directed surface current $\mathbf{J}_s = J_x f_1(x) f_2(y) \mathbf{u}_x$. After evaluating the V_i for both the TE and TM modes, we will then consider a y directed current, $\mathbf{J}_s = J_y f_3(x) f_4(y) \mathbf{u}_y$. All integrations are over the entire guide cross section.

$$\begin{aligned} V_{ix}^{TE} &= -\hat{Z}_{mn}^{TE} \iint J_x f_1'(x) f_2'(y) N_1 g_1 \, dxdy \\ V_{ix}^{TM} &= -\hat{Z}_{mn}^{TM} \iint J_x f_1'(x) f_2'(y) N_2 g_1 \, dxdy \\ V_{iy}^{TE} &= +\hat{Z}_{mn}^{TE} \iint J_y f_3'(x) f_4'(y) N_2 g_2 \, dxdy \\ V_{iy}^{TM} &= -\hat{Z}_{mn}^{TM} \iint J_y f_3'(x) f_4'(y) N_1 g_2 \, dxdy \end{aligned}$$

Evaluating the indicated integrals, we obtain

$$V_{ix}^{TE} = -\hat{Z}_{mn}^{TE} G_1'(\Delta x) G_2'(\Delta y) N_1 g_1' \; J_x$$

$$V_{ix}^{TM} = -\hat{Z}_{mn}^{TM} G_1'(\Delta x) G_2'(\Delta y) N_2 g_1' \; J_x$$

$$V_{iy}^{TE} = +\hat{Z}_{mn}^{TE} G_3'(\Delta x) G_4'(\Delta y) N_2 g_2' \; J_y$$

$$V_{iy}^{TM} = -\hat{Z}_{mn}^{TM} G_3'(\Delta x) G_4'(\Delta y) N_1 g_2' \; J_y$$

These V_i may be used directly in (3.1). Evaluating the transverse electric fields at z=h due to the surface current, J_s, we find

$$E_t\big|_{z=h} = \sum V_i e_i$$

Substituting in the V_i and evaluating E_t due to the TE field first

$$E_x^{TE} = \sum \left(-\hat{Z}_{mn}^{TE} N_1 G_1'(\Delta x) G_2'(\Delta y) g_1' J_x + \hat{Z}_{mn}^{TE} N_2 G_3'(\Delta x) G_4'(\Delta y) g_2' J_y \right) N_1 g_1$$

$$E_y^{TE} = \sum \left(+\hat{Z}_{mn}^{TE} N_1 G_1'(\Delta x) G_2'(\Delta y) g_1' J_x + \hat{Z}_{mn}^{TE} N_2 G_3'(\Delta x) G_4'(\Delta y) g_2' J_y \right) N_2 g_2$$

$$E_x^{TM} = \sum \left(-\hat{Z}_{mn}^{TM} N_2 G_1'(\Delta x) G_2'(\Delta y) g_1' J_x + \hat{Z}_{mn}^{TM} N_1 G_3'(\Delta x) G_4'(\Delta y) g_2' J_y \right) N_2 g_1$$

$$E_y^{TM} = \sum \left(-\hat{Z}_{mn}^{TM} N_2 G_1'(\Delta x) G_2'(\Delta y) g_1' J_x + \hat{Z}_{mn}^{TM} N_1 G_3'(\Delta x) G_4'(\Delta y) g_2' J_y \right) N_1 g_2$$

Summing the TE and TM modes, we have

$$E_x = \sum_{m,n} \left[-G_1'(\Delta x) G_2'(\Delta y) g_1' g_1 \left(N_1^2 \hat{Z}_{mn}^{TE} + N_2^2 \hat{Z}_{mn}^{TM} \right) \right] J_x + \\ + \left[G_3'(\Delta x) G_4'(\Delta y) N_1 N_2 g_2' g_1 \left(\hat{Z}_{mn}^{TE} - \hat{Z}_{mn}^{TM} \right) \right] J_y \qquad (3.4)$$

$$E_y = \sum_{m,n} \left[G_1'(\Delta x) G_2'(\Delta y) N_1 N_2 g_1' g_2 \left(\hat{Z}_{mn}^{TE} - \hat{Z}_{mn}^{TM} \right) \right] J_x + \\ + \left[-G_3'(\Delta x) G_4'(\Delta y) g_2' g_2 \left(N_2^2 \hat{Z}_{mn}^{TE} + N_1^2 \hat{Z}_{mn}^{TM} \right) \right] J_y$$

Equation (3.4) represents the electric field tangent to the surface of the substrate. This equation is similar to eq. 18 in [9], illustrating the close resemblance of this technique to the spectral domain approach.

A quantity which will be of interest is a weighted integral of a tangential component of the electric field on a subsection. We will choose the weighting function for the x component of the electric field to be $f_1(x)f_2(y)$, where f_1 and f_2 correspond to a field subsection. Multiplying E_x by this function and integrating term by term, we find that simply multipling each term in the summation by $G_1(\Delta x)G_2(\Delta y)$ effects the integration.

Similarly, we will choose a weighting function for the y component of the electric field to be $f_3(x)f_4(y)$, where f_3 and f_4 correspond to a field subsection. In a like manner, the integration is effected by multiplying each term in the summation for E_y by $G_3(\Delta x)G_4(\Delta y)$.

These integrals (known as reactions) of the electric field will be useful for implementing a Galerkin analysis and for calculating circuit input admittances.

3.4 Implementation of the Method of Moments Solution

A method of moments solution will be implemented by expanding the current on the microstrip metalization into a sum of currents on rectangles, the rectangles subdividing the entire metalization. The x and y currents may (and will be) expanded onto distinct sets of rectangles. The total current will be written as a sum of the current on all rectangles.

$$J_s = \sum J_{xk} f_1(x,x_k) f_2(y,y_k) u_x + \sum J_{y\ell} f_3(x,x_\ell) f_4(y,y_\ell) u_y$$

The k^{th} subsection is for x directed current and is centered at x_k, y_k. The ℓ^{th} subsection is for y directed current and is centered at x_ℓ, y_ℓ. The centers for the x and y current subsections will, in general, be different. For complete generality, the pulse functions, f, can be different on different subsections and the subsections may be different sizes. For the present task, we will assume that the four pulse functions are not a function of the indices k and ℓ and that all subsections have the same Δx and Δy.

The simplest choice for the pulse functions would be rectangular pulses for all four, f_1 through f_4. Then dividing the microstrip metalization into subsections will simply consist of selecting adjacent, non-overlapping rectangles which cover the entire metalization. A problem with the rectangular current distribution is that the discontinuity in the current at the edges of the rectangular subsection results in infinite line charges. The difficulties encountered in this situation are discussed later.

A second choice for pulse functions is to use a triangular pulse in the direction of the current and a rectangular pulse in the other direction. This would correspond to making f_1 and f_4 triangles while leaving f_2 and f_3 as rectangles. Now the divergence of the current is merely discontinuous rather than undefined. The electric field at the edge of the pulse is still singular. However, when weighted by the same triangle-rectangle pulse functions, the singularity is integrable. This is one reason why we are interested in the weighted integral of the electric field over each subsection.

The triangle basis function has been used [43] for two dimensional open microstrip problems. The triangle-rectangle pulse function was introduced by Glisson and Wilton [41] as the 'roof-top' function. In order to use the roof-top function as a basis function, the subdivision of the metalization into rectangles must be performed carefully.

First, the triangle portions of the current distributions will overlap as in Figure 3.5. This provides a piecewise linear approximation to the current in the direction of current flow. The rectangle portions of the current distributions must be adjacent, this results in a step approximation to the current transverse to the direction of current flow. This must be true of both the x directed and y directed current. If all subsections possess the same $\Delta x, \Delta y$ parameters, such an arrangement can be realized by spacing the rectangle centers on a $\Delta x, \Delta y$ grid. As will be pointed out next, the same $\Delta x, \Delta y$ grid cannot be used for both the x and y directed currents.

The rectangle function has a width of Δx (or Δy) while the

triangle function has a total width of $2\Delta y$ (or $2\Delta x$). The x directed current has a triangle in the x direction, while y directed current has a triangle in the y direction. We could place the centers of both x and y directed current subsections on the same grid. However, any edges in the metalization current expansion will not be clearly defined. For example, take an edge which is parallel to the x-axis. Is that edge located where the x directed current (rectangle function) ends or is it $\Delta y/2$ further away, where the y directed current (triangle function) ends? While this is a problem, in the limit, as Δx and Δy go to zero, it will make little difference. We must, however, turn our attention to a more important problem.

A rectangle of current on a substrate surface, much like a current element in free space, will generate only an electric field parallel to the direction of the current at the center. There is no central electric field generated perpendicular to the direction of the current. The same is true of the integral of the electric field over the center of a source subsection. If we use the same grid for the centers of the x and y directed currents, then x directed current can not generate any y directed electric field on its own subsection. The reverse is true for y directed currents. Thus, (given an arbitrarily large conductivity) a current on a subsection cannot generate any perpendicular current on a colocated subsection. As will be discussed later, this kind of subsection arrangement generates incorrect results.

This problem is resolved in a straight forward manner as suggested by Glisson and Wilton [41]. Simply offset the grid for one

current component by $\Delta x/2, \Delta y/2$ with respect to the other grid. This will allow every subsection to generate perpendicularly directed current on all other subsections. As an added benefit, it will also cause the edges of the x directed and y directed current expansions to line up perfectly.

Since there is little additional numerical effort, we will also investigate a parabolic-rectangle distribution function. Since this function is smoother than the triangle pulse function, we will be able to accurately represent the parabolic pulse with fewer waveguide modes than for the triangular pulse. However, it will be constrained to zero derivative at the subsection centers, in contrast to the discontinuous derivative of the triangle function.

Implementation of a Galerkin method of moments solution is now straightforward. Divide the microstrip into an x and a y directed grid of subsections as discussed above. The current densities on the subsections will form a set of dependent variables in a system of equations. Using (3.4), evaluate the weighted integral of the electric field on all subsections due to each current. The weighted integrals of the electric field on the subsections will form the set of independent variables. The independent variables will now be related to the dependent variables by an impedance matrix. Pick one (or more) subsections as a source, set the integral of the electric field on that subsection equal to one and all the others to zero (zero tangential electric field on a conductor). Invert the matrix and you have a solution, the currents on the subsections which give the desired tangential electric field.

The selection and characterization of the source will be discussed next.

3.5 The Source Model

Microstrip circuit inputs and outputs are usually taken at the edge of the substrate by means of a coaxial cable penetrating the shielding sidewall at z=h. The coax shield is connected to the microstrip shield and the coax center conductor is attached to a microstrip conductor.

The coax aperture can be modeled by a conductor backed circulating magnetic current. We will assume that the aperture is small and that the aperture current has negligible effect. When we compare measured data with calculated data in a later section, we will find that the contribution from the circulating magnetic current is important and can be modeled, for practical situations, as a small fringing capacitance in shunt with the connector.

The primary effect will be the electric current injected by the center conductor. The center conductor is circular and the microstrip is a two dimensional surface. Assuming the center conductor is small, we can model the current injected by the coax center conductor as a rectangular subsection of surface current directed perpendicularly to the shielding side wall. This subsection will be called a port subsection. It is convenient to use the same pulse function as with the other subsections of the microstrip circuit. This facilitates the transition from the port subsection to the microstrip subsections.

The center of the port subsection will be on the shielding sidewall with only half of the pulse protruding from the sidewall.

Evaluation of the V_i for a port subsection is similar to that of the microstrip subsections. Images provide us with exactly the same current distribution on the port subsection as the other subsections, only now half the current is inside the shield and half is outside. Thus the integral of (3.3) extends over half of its former domain (the inside half). Since the integrand is even about the sidewall, the $G(\Delta x)$ or $G(\Delta y)$ (and, in turn, the V_i) are simply half that of a full pulse.

In summary, a port subsection is centered on the shielding sidewall at the location of the coaxial input. Then, since only half the pulse is inside the shield, each row and column of the impedance matrix associated with the port subsection is multiplied by one half. Otherwise, port subsections are the same as all the other subsections.

3.6 Evaluation of the Input Admittance

For the initial portion of this section we will discuss the input admittance of a one port microstrip circuit. Quantities associated with that port will be designated by subscript 1. Elements in the admittance matrix of the entire microstrip system will have double numerical subscripts.

The input admittance will be evaluated by means of the usual variational expression (pp. 348-349 of [6])

$$Y_1 = - \frac{I_1^2}{\iint \mathbf{E} \cdot \mathbf{J} \, ds}$$

Since E is zero everywhere except at the port subsection, we need only consider the current on the port subsection. In the Galerkin solution outlined above, we represent the source as the weighted integral of the electric field over the area of the port subsection set equal to one. The current on the port subsection is proportional to that same weighting function, the constant of proportionality being $Y_{11} = J_1$. Thus, the denominator of the above expression is just Y_{11}. The input current is just the input current density multiplied by the width of the input, Δw, usually either Δx or Δy. Thus the input admittance is

$$Y_1 = - (Y_{11} \Delta w)^2 / Y_{11} = - Y_{11} (\Delta w)^2$$

In a like manner, the transfer admittance between any two ports, say port a and port b, of an N port circuit may be determined by

$$Y_{tran} = - Y_{ab} \Delta w_a \Delta w_b$$

Thus, the N port admittance matrix of circuit theory is formed from the elements of the Galerkin admittance matrix. Simply select all elements which lie at the intersection of any port subsection row and column. Then multiply each element by both associated port subsection widths. Tne array of these elements will form the Y

parameter matrix of the N-port. This may then be converted to Z parameters or S parameters by the usual transformations.

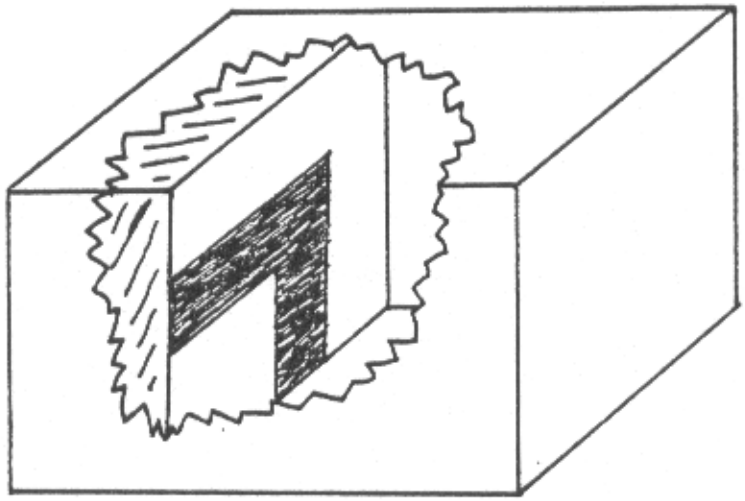


Figure 3.1. The microstrip geometry to be analyzed is realized on a dielectric substrate in a printed circuit fashion and is completely contained in a shielding, conducting rectangular box.

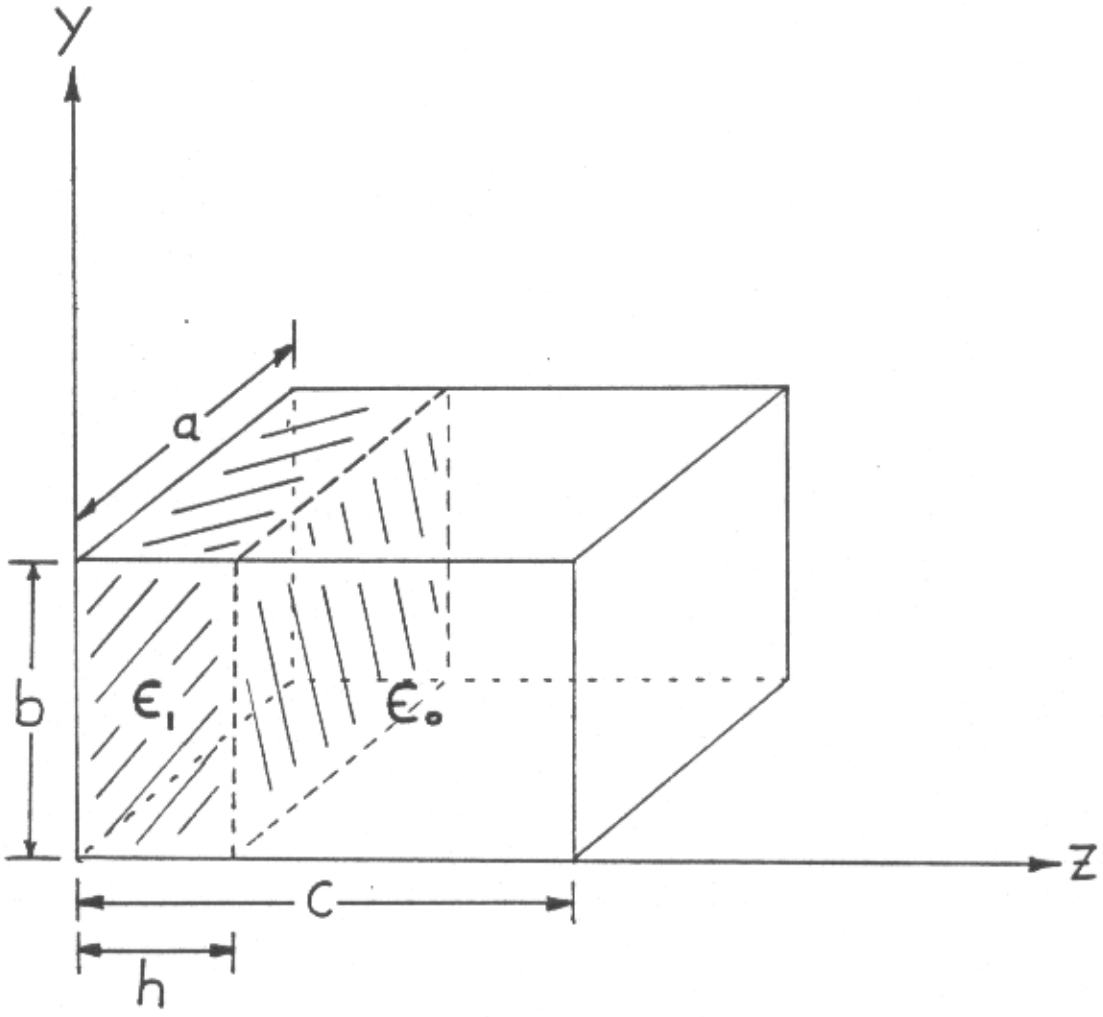


Figure 3.2. The coordinate system used is oriented so as to emphasize the fact that the fields are represented as a sum of homogeneous rectangular waveguide modes with the waveguide propagating in the z direction.

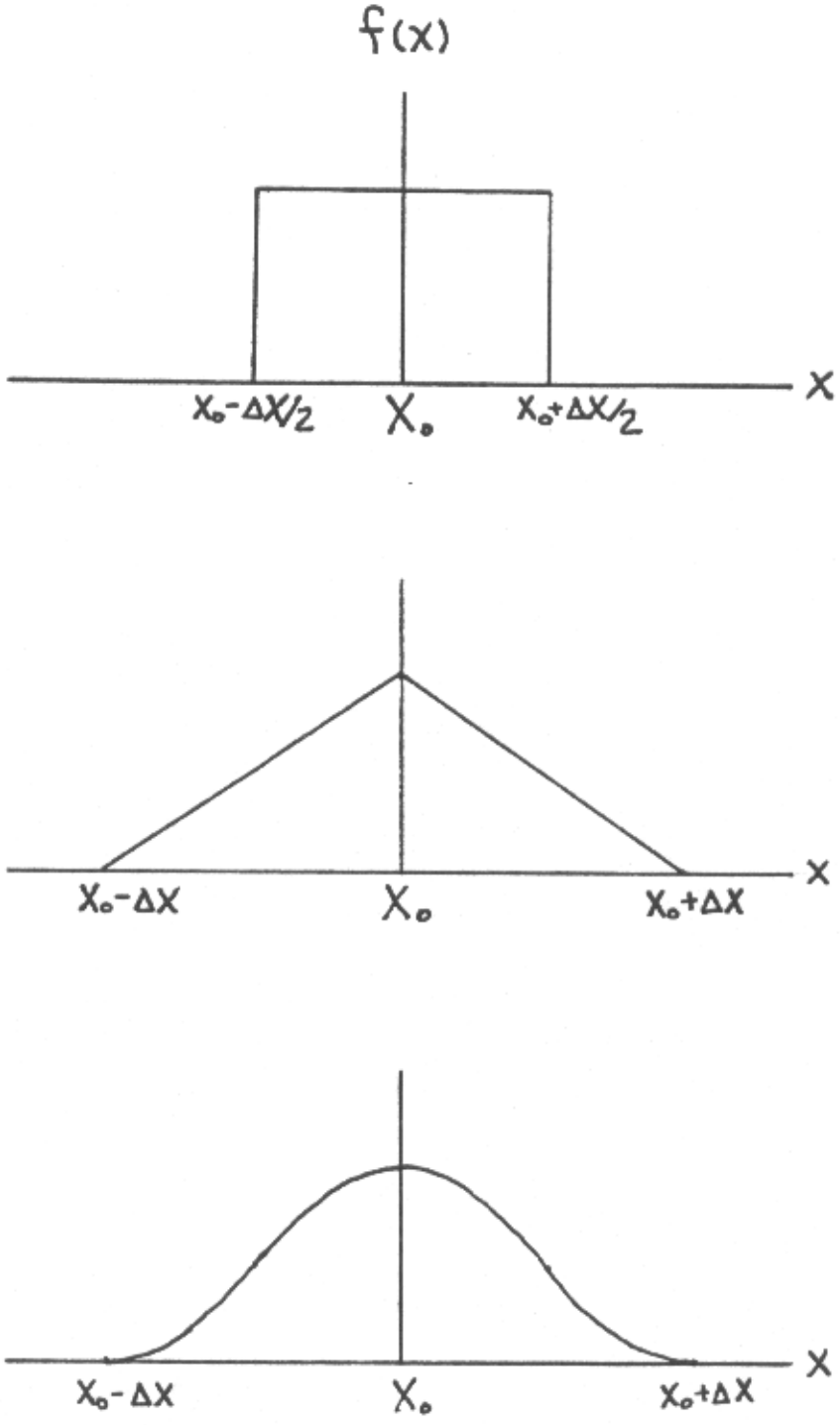


Figure 3.3. The rectangular (top), triangular (middle) and piecewise parabolic (bottom) functions used to build expansion functions for the representation of surface current on the microstrip metalization.

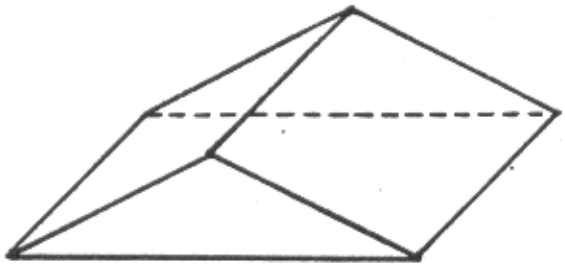


Figure 3.4. The product of a triangle function in one direction by a rectangle function in the lateral direction gives a roof-top function which will be used as an expansion function.

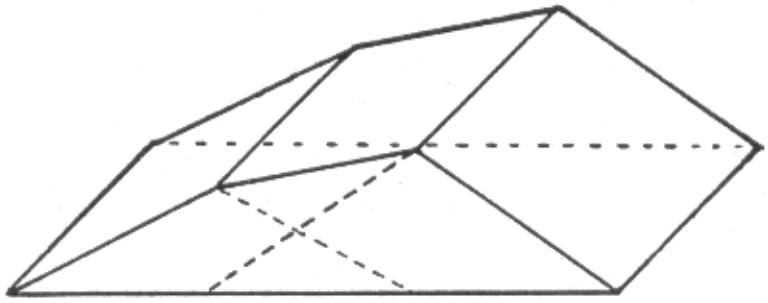


Figure 3.5. Two roof-top functions placed on overlapping rectangles give a piecewise linear approximation to the current in the direction of current flow. Additional roof-top functions placed side by side will provide a step approximation to the surface current in the direction lateral to current flow.

Chapter 4

EFFICIENT CALCULATION OF THE IMPEDANCE MATRIX ELEMENTS

The double summations for the impedance matrix elements require thousands of terms before accurate results can be obtained. This section describes a means of reorganizing the terms in the summation which results in a significant reduction of the time required to evaluate the matrix elements.

In addition, the summation is split into two stages. The first stage depends only on the shielding and substrate geometry. As long as the shield and the substrate remain unchanged, the first stage of the summation need not be recalculated. The number of floating point operations required for the first stage is proportional to the number of included modes and inversely proportional to the product of the x and y axis resolution (Δx and Δy).

The second stage of the summation is dependent only on the specific microstrip circuit geometry to be analyzed. The number of floating point operations required for each element of the system matrix is inversely proportional to the product of the x and y axis resolution (Δx and Δy) and is independent of the number of included modes. The microstrip circuit can be changed and only the second stage summation need be repeated.

Typical analyses have demonstrated a twenty fold improvement in speed for the complete analysis. If we ignore the circuit independent first stage and since the second stage of the summation is independent of the number of modes, we will realize an arbitrarily large relative improvement when including an arbitrarily large number of modes.

The summation is reorganized by restricting the centers of all subsections to a Δx, Δy grid and taking advantage of the periodicity of the sine and cosine functions. This allows the summation, which goes to infinity, to be reorganized into two summations. The final summation, which is circuit dependent and contains the sine and cosine functions, is taken over a specific finite range. This summation is the second stage referred to above. The terms of this summation themselves include summations which are taken over an infinite range. These infinite summations, which are terminated at some convenient point, form the circuit independent first stage summation.

4.1 Initial Identities Used in the Summation Reorganization

We wish to rewrite a double summation which includes sine and cosine functions so as to take advantage of the periodicities of the sine and cosine. Before approaching that problem directly, we will investigate rewriting some single summations. The following identities will be useful.

$$\sin(m\pi m_o/M) = (-1)^{im_o}\sin\big((iM+m)\pi m_o/M\big) = -(-1)^{im_o}\sin\big((iM-m)\pi m_o/M\big)$$

$$\cos(m\pi m_o/M) = (-1)^{im_o}\cos\big((iM+m)\pi m_o/M\big) = (-1)^{im_o}\cos\big((iM-m)\pi m_o/M\big)$$

$$\sin(i\pi m_o) = 0 \qquad \cos(i\pi m_o) = (-1)^{im_o}$$

$$\sin\big((i+1/2)\pi m_o\big) = 0, \qquad m_o \text{ even}$$

$$= (-1)^{(m_o-1)/2}, \quad m_o \text{ odd}$$

$$\cos\big((i+1/2)\pi m_o\big) = (-1)^{m_o/2}, \qquad m_o \text{ even}$$

$$= 0, \qquad m_o \text{ odd}$$

The above identities can be verified by inspection or by application of the trigonometric identities for the sum of angles.

The above identities applied to the three summations with which we must deal yield the following results.

$$\sum_{m=0}^{\infty}\cos(m\pi m_o/M)\cos(m\pi m_1/M)f(m) = \sum_{i=0}^{\infty}(-1)^{i(m_o+m_1)}f(iM) +$$

$$+ (m_o \text{ and } m_1 \text{ even})\sum_{i=0}^{\infty}(-1)^{(m_o+m_1)/2}\,f\big((i+1/2)M\big) +$$

$$+ \sum_{m=1}^{M/2-1}\cos(m\pi m_o/M)\cos(m\pi m_1/M)\sum_{i=0}^{\infty}\Big[(-1)^{i(m_o+m_1)}f(iM+m) +$$

$$+ (-1)^{i_1(m_o+m_1)}f(i_1M-m)\Big]$$

$$\sum_{m=0}^{\infty} \sin(m\pi m_0/M)\sin(m\pi m_1/M)f(m) =$$

$$= (m_0 \text{ and } m_1 \text{ odd}) \sum_{i=0}^{\infty} (-1)^{(m_0+m_1)/2} f\Big((i+1/2)M\Big) +$$

$$+ \sum_{m=1}^{M/2-1} \sin(m\pi m_0/M)\sin(m\pi m_1/M) \sum_{i=0}^{\infty} \Big[(-1)^{i(m_0+m_1)} f(iM+m) +$$

$$+ (-1)^{i_1(m_0+m_1)} f(i_1 M-m)\Big]$$

$$\sum_{m=0}^{\infty} \sin(m\pi m_0/M)\cos(m\pi m_1/M)f(m) =$$

$$= (m_0 \text{ odd and } m_1 \text{ even}) \sum_{i=0}^{\infty} (-1)^{(m_0+m_1-1)/2} f\Big((i+1/2)M\Big) +$$

$$+ \sum_{m=1}^{M/2-1} \sin(m\pi m_0/M)\cos(m\pi m_1/M) \sum_{i=0}^{\infty} \Big[(-1)^{i(m_0+m_1)} f(iM+m) -$$

$$- (-1)^{i_1(m_0+m_1)} f(i_1 M-m)\Big]$$

where m, i = summation indices.

i_1 = $i + 1$.

m_0, m_1, M = constant integers.

$f(m)$ = a function of the summation index, m.

In the above equations, any summation preceded by a statement in parentheses is performed only when the statement is true.

4.2 Rewriting the Summation

The expressions for the single summations of the previous section

may be applied twice to evaluate desired double summations. We have three distinct cases to evaluate:

1) The x directed electric field due to x directed current.
2) The x directed electric field due to y directed current.
3) The y directed electric field due to y directed current.

The fourth case, y directed electric field due to x directed current is the same as case 2 above.

We will restrict the centers of all subsections to a Δx, Δy grid with M Δx divisions from x=0 to x=a along the x axis and N Δy divisions from y=0 to y=b along the y axis. Placing the center of the source subsection at (x_o, y_o) and the center of the field subsection at (x_1, y_1) we will define m_o, m_1, n_o and n_1 with the following relations.

$$x_o/a = m_o/M, \quad x_1/a = m_1/M, \quad y_o/b = n_o/N, \quad y_1/b = n_1/N$$

We will evaluate each of the three cases in turn. In each case, the original summation will be written followed by the reorganized summation. The symbols used in the f(m,n) functions have been defined previously.

CASE 1) $J_x \Rightarrow E_x$

$$E_x = \sum_{m=0}^{\infty} \cos(m\pi m_o/M)\cos(m\pi m_1/M) \sum_{n=1}^{\infty} \sin(n\pi n_o/N)\sin(n\pi n_1/N) f_{xx}(m,n)$$

where $f_{xx}(m,n) = G_1(\Delta x)G_2(\Delta y)\left[N_1^2 \hat{Z}^{TE}_{mn} + N_2^2 \hat{Z}^{TM}_{mn}\right]$
and m_o, m_1, n_o, n_1 are all even.

$$E_x = \sum_{n=1}^{N/2-1} \sin(n\pi n_o/N)\sin(n\pi n_1/N) \sum_{j=0}^{\infty} \left(F_{xx}(jN+n)+F_{xx}(j_1 N-n)\right)$$

$$F_{xx}(n) = \sum_{i=0}^{\infty} \left[f_{xx}(iM,n)+(-1)^{(m_o+m_1)/2} f_{xx}\left((i+1/2)M,n\right)\right] +$$

$$+ \sum_{m=1}^{M/2-1} \cos(m\pi m_o/M)\cos(m\pi m_1/M) \sum_{i=0}^{\infty} \left(f_{xx}(iM+m,n)+f_{xx}(i_1 M-m,n)\right)$$

CASE 2) $J_y \Rightarrow E_x$

$$E_x = \sum_{m=1}^{\infty} \sin(m\pi m_o/M)\cos(m\pi m_1/M) \sum_{n=1}^{\infty} \cos(n\pi n_o/N)\sin(n\pi n_1/N) f_{xy}(m,n)$$

where $f_{xy}(m,n) = G_3(\Delta x)G_4(\Delta y)N_1 N_2\left[\hat{Z}^{TE}_{mn} - \hat{Z}^{TM}_{mn}\right]$
and m_o, n_o are odd and m_1, n_1 are even.

$$E_x = \sum_{n=1}^{N/2-1} \cos(n\pi n_o/N)\sin(n\pi n_1/N) \sum_{j=0}^{\infty} (-1)^j \left(F_{xy}(jN+n)+F_{xy}(j_1 N-n)\right)$$

$$F_{xy}(n) = (-1)^{(m_o+m_1-1)/2} \sum_{i=0}^{\infty} f_{xy}\left((i+1/2)M,n\right) +$$

$$+ \sum_{m=1}^{M/2-1} \sin(m\pi m_o/M)\cos(m\pi m_1/M) \sum_{i=0}^{\infty} (-1)^i \left(f_{xy}(iM+m,n)+f_{xy}(i_1 M-m,n)\right)$$

CASE 3) $J_y \Rightarrow E_y$

$$E_y = \sum_{m=1}^{\infty} \sin(m\pi m_o/M)\sin(m\pi m_1/M) \sum_{n=0}^{\infty} \cos(n\pi n_o/N)\cos(n\pi n_1/N) f_{yy}(m,n)$$

where $f_{yy}(m,n) = G_3(\Delta x)G_4(\Delta y)\left[N_2^2 \hat{Z}^{TE}_{mn} + N_1^2 \hat{Z}^{TM}_{mn}\right]$
and m_o, m_1, n_o, n_1 are all odd.

$$E_y = \sum_{j=0}^{\infty} F_{yy}(jN) + \sum_{n=1}^{N/2-1} \cos(n\pi n_o/N)\cos(n\pi n_1/N) \sum_{j=0}^{\infty} \left(F_{yy}(jN+n)+F_{yy}(j_1N-n)\right)$$

$$F_{yy}(n) = \sum_{i=0}^{\infty} (-1)^{(m_o+m_1)/2} f_{yy}\left((i+1/2)M,n\right) +$$

$$+ \sum_{m=1}^{M/2-1} \sin(m\pi m_o/M)\sin(m\pi m_1/M) \sum_{i=0}^{\infty} \left(f_{yy}(iM+m,n)+f_{yy}(i_1M-m,n)\right)$$

In the above equations, the summations over i and j represent the first stage and are truncated at some convenient point. The integer i_1 equals i+1 and j_1 equals j+1. The first stage summations need be done only once for a given substrate and shielding geometry with the results stored in memory. The summations over m and n are the second stage summations. The second stage needs to be executed each time metalization is added to the circuit.

The above equations are for the tangential E field. For a Galerkin technique, we need the weighted integral of the E field. This is realized by squaring each G(Δx) and G(Δy) function for cases 1 and 3. Case 2 is modified by multipling f_{xy} by $G_1(\Delta x)$ and $G_2(\Delta y)$.

4.3 Application of the Discrete Fourier Transform

A two-dimensional discrete Fourier transform implemented using an

efficient algorithm could realize the second stage summation. To do this, select a single source point. (One could also proceed by selecting a single field point.) This single point will be viewed as a fixed constant. The sines and cosines which are a function of the fixed point, combined with the first stage summations now form the Fourier coefficients of a sine or cosine series in two dimensions. The series extends from m=0 to M/2-1 and from n=0 to N/2-1.

To simply apply a discrete Fourier transform, the series should be extended in an odd fashion for a sine series and in an even fashion for a cosine series with the resulting summation going from m=0 to M-1 and n=0 to N-1. The second stage of the summation can now be realized by the application of a discrete Fourier transform. The transform is with respect to the spectral variables, m and n, to yield a sequence dependent on the spatial variables, m_1 and n_1 for a given source subsection (m_o, n_o). A pair of two dimensional transforms (one for x directed fields and one for y directed fields) must be performed for each subsection. This number can be reduced by invoking reciprocity.

Multiple application (twice for each possible source subsection) of the discrete Fourier transform provides us with the system matrix elements for all possible field subsections. Since a typical circuit will use only a small portion of all possible field subsections, the transform provides a great deal of unneeded information. Even for required field subsections, redundant information is calculated by the transform due to the symmetry of the system matrix. Thus a transform approach to the second stage may actually be less efficient than a direct summation. However, a distinct opportunity lies in the

possibility of implementing a discrete Fourier transform algorithm which does not calculate unneeded elements and takes advantage of the symmetry (even or odd) of the Fourier coefficients. Such an algorithm would have the potential of significantly improving the efficiency of the second stage calculation. This will be the subject of future research.

It should be noted that the application of a discrete Fourier transform does not imply any approximation. In fact, in the limit, as the first stage summation is carried to an infinte number of terms, the transform implementation of the second stage summation provides an exact result.

The FFT algorithm has been applied previously in the calculation of static microstrip capacitance [42].

Chapter 5

SOFTWARE DESIGN

To date, most numerical electromagnetic software (often referred to as 'code') has been written in FORTRAN, with a few programs written in BASIC. All software used in this work has been written in Pascal, in part, as an experiment to determine the strengths and weaknesses of Pascal in a numerical electromagnetics environment.

The initial, and most glaring, shortcoming of Pascal is the lack of an intrinsic complex data type. For most general programming tasks, this is of no consequence. However, extensive use of a complex data type is usually required in electromagnetics. BASIC, which also lacks a complex data type, has been used in complex arithmetic. While the programs often work very well, the source listing can be difficult to follow. Thus, the ease with which the software can be modified, maintained or ported to another system is compromised.

With Pascal, new data types can be defined. For example one could define the COMPLEX data type as a RECORD whose first component is a floating point number which we can call RE, and whose second component is also a floating point number which we could call IM. Then, if we declare a variable, EREL, as COMPLEX, it will have both a real and an imaginary part. If we wish to refer to the complex value, we would use EREL. To refer to the real part, we use EREL.RE, while EREL.IM refers to the imaginary part.

To perform complex manipulations, one now must write COMPLEX procedures (essentially the same thing as subroutines in FORTRAN, only the word CALL is not needed to invoke the procedure). For example,

one could write a complex multiply procedure which would be invoked by writing CMULT(A,B,C), where A, B and C are COMPLEX data types. A disadvantage here is that complicated equations must now be written as a series of procedure references, making it difficult to compare the source with the original equation. This, then is a real difficulty with Pascal. In fact, in applications requiring complex arithmetic which can not be vectorized (as described later) and in which the subsequently described advantages of Pascal are not overwhelming, FORTRAN is likely to be the better language.

This difficulty with Pascal could be eased somewhat if one could write COMPLEX functions. While this is allowed in some versions of Pascal, it is not standard and should be avoided if it is desired to make the software portable from system to system. This brings us to an advantage of Pascal. Standard Pascal is precisely and clearly defined. If one stays with the standard structures, plus, possibly a few fairly standard extensions, a highly portable program can result. If most of the software is numerical in nature, this portability does not represent a strong advantage over FORTRAN as FORTRAN can be highly portable as well. However, it is the author's experience that a numerically intensive program, once reduced to practice as a quality engineering computer aided design program, will have two to five times as much source allocated to user interface as it does to the original number crunching. In such a situation, FORTRAN has very poor portablity compared to Pascal. The reason is that there are about a dozen different 'standard' FORTRANs, each standard having many implementations, with each implementation having a wealth of

extensions. Frequently, the extensions are too tempting to pass up, the result being loss of portablity even if all systems were to use the same FORTRAN standard.

A second advantage of Pascal is the wealth of data types which are both built in and which can be constructed by the user. This has been used in this work in relation to the vectorization of the software. By vectorization, we mean that data is organized in long vectors, the longer, the better. The advantage here is that an array processing or parallel processing computer can very quickly manipulate long vectors. Thus, rather than, for example, placing a multiply followed by an add (D(I) = A(I)*B(I) + C(I)) in a loop and executing the loop N times, we would do all N multiplies (perhaps in parallel), then all N adds. In short, rather than multiplying scalars, one at a time, we would be multiplying entire vectors all at once.

Thus our problem reduces to specifying a vector data type and writing procedures to manipulate (multiply, add, etc.) that data type. Note that now, Pascal and FORTRAN are at equal disadvantage. In both cases, we must now express equations as sequences of procedure calls. However, Pascal has the advantage of being able to create a more flexible (and portable) vector data type as described next.

The vector data type designed for this work is a RECORD. The record has three parts. The first part is an INTEGER which specifies the length of the vector. The second and third parts are pointers. A pointer is simply space to store a memory address. The memory location thus pointed to will be the beginning of an array of floating point numbers which will form the vector. The first pointer in the

record points to the real vector, while the second pointer points to the imaginary part. The real and imaginary vectors are referred to as dynamic arrays, that is, space for the arrays is allocated or deallocated during run time, not at compile time. The pointers (which point to the vectors) are initialized to the standard value NIL (usually 0) which means that no memory is allocated for the array. When necessary, the software allocates memory for the vectors, manipulates the vectors and deallocates the memory when finished, freeing the memory for other uses.

An important advantage here, is that if there is no need for the imaginary (or real) part in a particular problem, the imaginary (or real part) takes no memory. In addition, the vector procedures are easily written so that the additional time required for a full complex manipulation is not used when the arguments have no imaginary (or real) parts. Thus, in the problem considered in this work, the same software can perform a lossless or lossy analysis with little compromise.

Another advantage of Pascal which will be only briefly mentioned is the structure and readability of the software. The group of programming techniques generically referred to as 'structured programming' are well established in the programming community. Ignorance of such structured techniques is extremely hazardous. Pascal not only makes possible many structured techniques (as compared with FORTRAN), it encourages their use. Pascal also tends to be more readable than FORTRAN. This is important in the software maintenance area, especially if a programmer other than the orginal author is to

maintain the software. As one example of readability, consider an example FORTRAN variable: NXP. What does it do? The Pascal programmer would be more likely to write: Num_x_ports. Now, even out of context, we have some idea of what the variable does.

In the course of this work, we have only evaluated Pascal as an alternative to FORTRAN for numerical electromagnetic software. In the future, we also plan to evaluate C to find its advantages and disadvantages. Our increased interest in C is due to its recently increased availablity on personal computers.

At this point, we would recommend FORTRAN over Pascal if:

1) The programmer has an extensive installed base of FORTRAN software.
2) If the software will be predominately numerical in nature (little user interface or input/output).
3) The nature of the task does not justify or conform easily to vectorization.
4) The software will not be ported to different systems (given a significant user interface).
5) Maintenance/update requirements are minimal.

We would recommend Pascal over FORTRAN if:

1) Maintainability and portability are important issues.
2) The specific problem can take advantage of advanced data structures (such as is the case with

vectorization).

3) There is going to be significant input/output or user interface.

4) The size of the task justifies the additional work required to develop a complex (vector) data type.

The above are, of course, subjective guidelines. It is up to the indvidual programmer faced with a specific task to evaluate the importance of each guideline and to determine a weight for each guideline. In fact, given a specific problem, there are quite likely to be other factors which must also enter the objective function.

One word of caution, beware advice stating that one language (or anything else) is better than another. 'Better' must always be measured with respect to the task at hand. Advice which does not take that into consideration should be treated carefully.

Chapter 6

SOFTWARE STRUCTURE

Several programs were written to assist in the verification of the electromagnetic analysis. All the programs were written in Pascal on an IBM-PC. At no time was the PC found to be lacking in memory for the compiled source (limited to 64 K by the particular compiler being used). The upper limit on data array size was reached several times. With all calculations performed in double precision and with the matrices involved being purely imaginary (no real part), it was found that a circuit involving somewhat less than 200 subsections could be handled on an IBM-PC with 640 K of memory.

After developing the complex vector data type, a set of procedures were written to manipulate the vectors. These procedures include add, subtract, multiply, divide, trig and inverse trig functions, square root, dot product, change sign, conjugate and move. In addition, procedures were written to perform scalar-vector operations such as to multiply each element of a vector by a scalar. Another procedure was written to do a pivot (multiply a vector by a scalar and add to a second vector). Finally, procedures were written to allocate and deallocate memory for the vectors as required.

A program was then written to validate the procedures. This program allows the user to exercise any of the procedures and to veiw the results. Validating the procedures during actual use was unacceptable.

One problem to which the programmer should give careful attention is that the vectors passed to these routines may not be distinct from

each other. For example, if we want to multiply vector A by vector B and place the result back into vector A, the multiply routine should not use vector A for intermediate results!

A second common problem is a corrupted heap. The 'heap' is the area of memory assigned to the storage of the dynamic variables (in this case, the vectors). The pointers to vectors must be initialized to NIL (nothing allocated) by the programmer, this is not done by the compiler. If a pointer is set to NIL, the complex vector procedures have been designed to treat the vector as all zeros. If one forgets to initialize a pointer and starts to use it, the pointer is pointing to some random place in memory and the system will usually crash. This can be a difficult problem to isolate.

A third problem occurs when one forgets to deallocate a vector, say, a temporary vector used for intermediate results. Each time the procedure using this temporary vector is called, it will allocate a new area of the heap for the vector. After enough procedure calls, the heap will become full (stack-heap collision) and the program will abort. This is a particularly insidious problem on systems with virtual memory. With virtual memory, the heap never becomes full. The only thing that will happen is that, at some point, the system must initiate substantial swapping and system response time will slow drastically.

As mentioned above, all the complex vector procedures perform only the necessary calculations. If we have two real vectors to multiply, the multiplies involving the imaginary part are not performed. A vector which is not allocated is treated as zero.

In addition to the vector data type, a complex matrix data type was developed. The matrix data type is essentially a one dimension array with each element of the array formed by a complex vector data type. The complex vectors form the rows of the matrix. A set of procedures was also written to manipulate the matrix data type much as was done for the vector data type. Also a program was written to exercise the procedures for validation purposes.

The initial analysis program was used to calculate the current resulting from truncating the modal series. It was later modified to include electric field. The vector length for this program was the length of the summation. Each element of a vector represented one term of the summation. This program was set up in an 'experimental' format, that is, it was designed so that switching between various modes of analysis (e.g., triangle versus parabolic expansion functions) could be accomplished readily. Fast execution time was a secondary consideration.

The second analysis program performed microstrip analysis. It was written much as the above program with each element of the vector representing one term of the summation. Elements of the system matrix were calculated one at a time. As with the other programs described here, it was written in an experimental format.

A third analysis program also performed microstrip analysis. It was written in order to check the results of the other microstrip analysis program. Small discrepancies were found which were traced to a precision problem in the calculation of $\hat{Y}$ for high order modes. This program was written with the vector length equal to the number of

subsections subdividing the microstrip geometry. Each term of the summation was calculated for the entire matrix before proceeding to the next term.

A fourth analysis program implemented the first and second stage summation described in the chapter on efficient calculations. All the analysis programs give identical results.

Finally, a program was written to allow a user to graphically define a microstrip circuit geometry with the aid of a mouse. A mouse is a small hand sized box with buttons on the top. It rests on the desk top (or on a special pad) and is connected to the computer. As the mouse is moved, the cursor on the computer screen is moved. By pressing a button on the mouse, a pop-up menu appears on the screen. The menu may contain options like 'X directed subsection'. The user moves the cursor (by moving the mouse) to the desired option. By pressing the button a second time, the option is selected and the menu disappears. If 'X directed subsection' were selected, a mouse button would now be defined as an X directed subsection button. That is, each time it is pressed, a subsection of X directed current would be laid down at the present location of the cursor.

In this manner, entire microstrip geometries can be specified. In fact, the PC is capable of capturing very complicated circuit geometries, much more complicated than can actually be analyzed on the PC. This suggests that the PC would make a good 'front end' for a larger system. The geometry captured by the PC could be up-loaded to a large system for analysis. The large system need not be tied up with data capture chores.

Figure 6.1 shows an interdigitated microstrip capacitor whose geometry was captured using an IBM-PC. Figure 6.2 shows a magnified view of the same capacitor. The fingers of the capacitor have only x directed current (horizontal lines) while much of the rest of the capacitor is composed of subsections which allow both x and y directed current (both horizontal and vertical lines). The size of the circuit (about 500 subsections) makes it far too large for a PC analysis while it is just the right size for many larger systems.

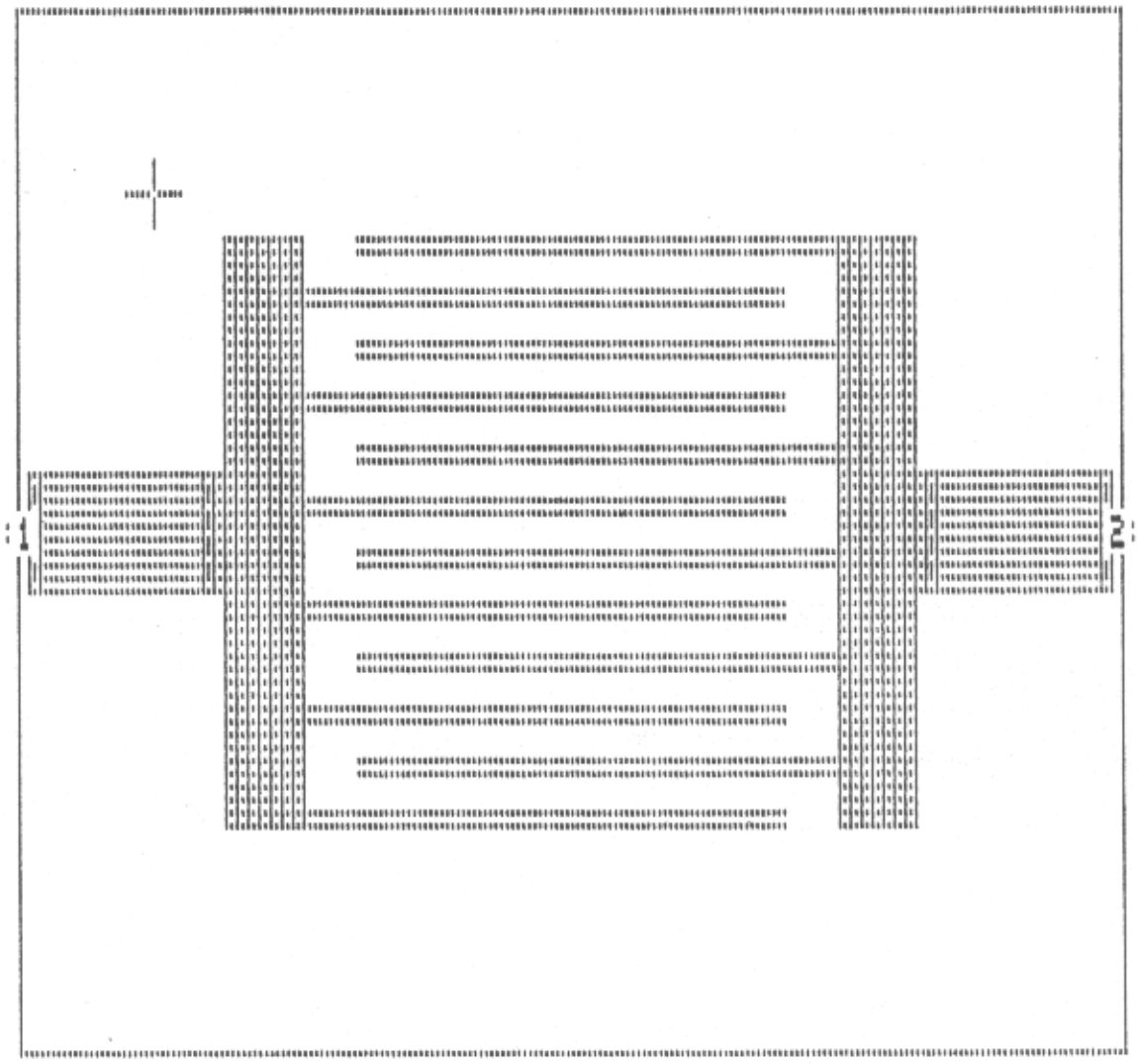


Figure 6.1. Sample graphical output of the microstrip geometry capture program. Horizontal lines indicate areas where x directed current is allowed while vertical lines indicate allowed y directed current. The structure is an interdigitated capacitor.

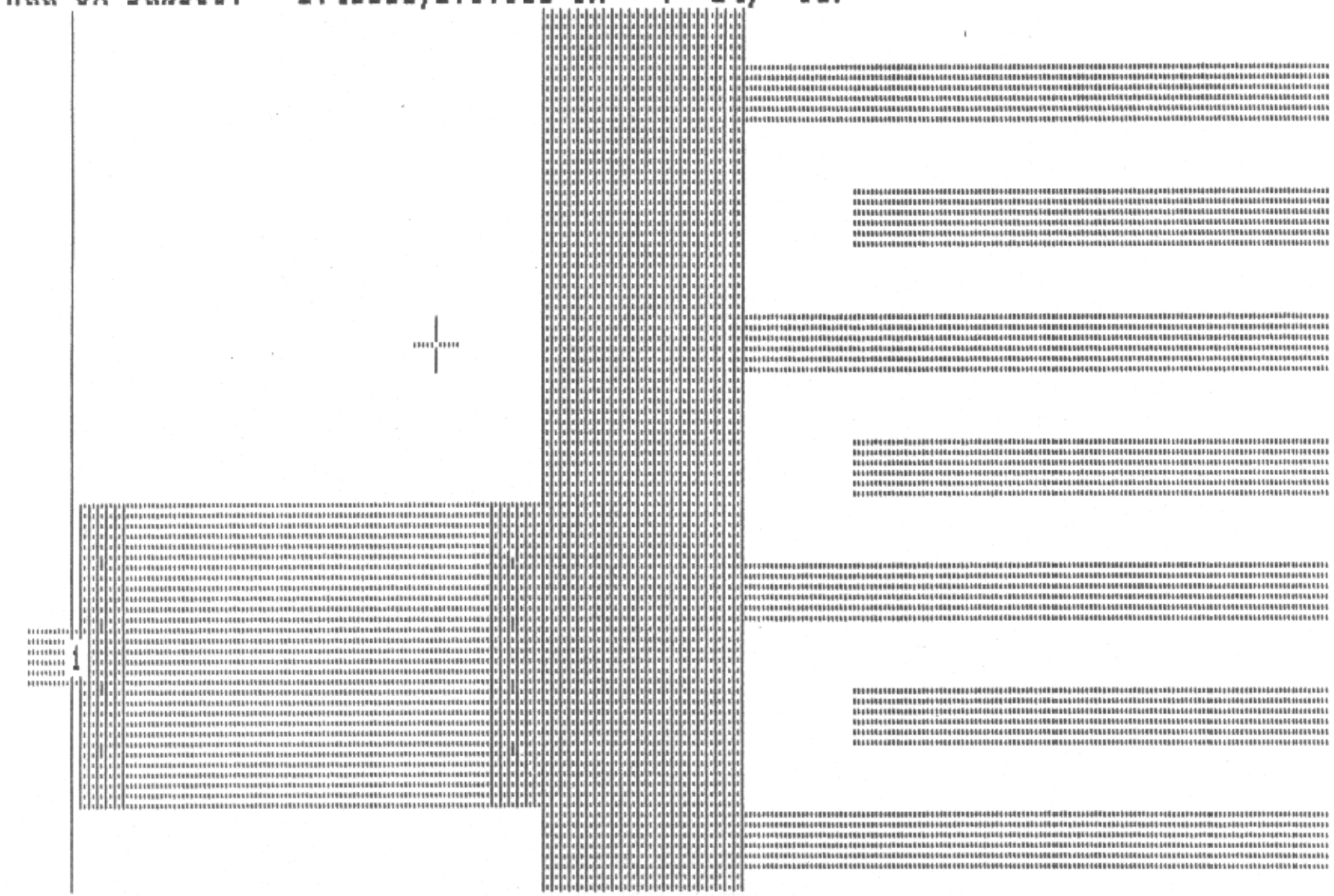


Figure 6.2. A magnified region of the interdigitated capacitor of Figure 6.1.

Chapter 7

RESULTS

To gain confidence in the analysis, several problems were investigated in detail. First, the fields generated by individual current patches were calculated and plotted. Next, an open circuited microstrip stub geometry was selected for analysis. With reasonable results here, a second stub geometry was selected, constructed and measured. The measurements provided excellent agreement with the analysis once an estimate of the fringing capacitance due to the circular coaxial aperture formed by the input SMA connector had been included. To further check the validity, a notch was cut in the stub midway along its length. Again the agreement between measured and calculated data was excellent.

Details of the measurements (performed using an Hewlett Packard 8510 automated network analyzer) and the analyses (all of which were performed on an IBM-PC with numeric coprocessor) will be presented next.

7.1 The Convolved Green's Function

The fields generated by a point source represent the Green's function for a given problem. In this section we will present examples of the fields due to small rectangular patches of surface current. These fields can be viewed as the Green's function convolved with the surface current distribution which generated the field. We do not calculate the Green's function itself.

The analysis calculates the V_i, the modal coefficients, for the desired current distribution. The resulting current and fields are then calculated by the summations in section 3.1. The summations are truncated after a specified number of modes. Modes, numbered by m and n, are included if m is less than an upper limit and n is less than a second upper limit. The upper limits are selected such that a dimension with finer geometries will be represented with more modes. For example, if $\Delta x/A$ is half of $\Delta y/B$, then the upper limit for m will be twice the upper limit for n.

Analyses which selected only the lowest order (in terms of cutoff frequency) modes were also effected. For a given number of modes, there was little difference in the resulting fields. This approach required more time (to find only the lowest modes) and the other technique is more adaptable to the efficient calculation algorithms described previously. Thus, the approach of finding only the lowest order modes was not pursued further.

The geometry of the current patch to be analyzed is shown in Figure 7.1 for x directed current and Figure 7.2 for y directed current. In all subsequent plots we will evaluate the field due to the current patch along a cross section passing through the center of the patch and parallel to the x axis.

Most of the analyses deal with the roof-top function, Figure 3.3. The current resulting from truncating the number of modes in the summation is shown in Figure 7.3 for a roof-top current distribution with the current directed in the x direction. Since the roof-top function has a triangle dependence in the direction of current flow,

we see an approximation to the triangle function. The three plots show the result of truncating the summation at 200, 1000 and 5000 modes respectively. Figure 7.4 shows the detail of the transition into the triangle region for a 5000 mode calculation.

The x directed electric field due to the x directed current, Figure 7.5, shows the increasing singularity of the field at points where the derivative of the triangular current density is discontinuous. This singularity suggests why the weighted integral of the electric field (Galerkin technique) is desired as opposed to point matching at the center of the patch. Again, the figure shows the effect of changing the included number of modes.

The final figure for x directed current, Figure 7.6, shows the y directed electric field along the same cross section. As the number of included modes goes to infinity, these fields should go to zero. This is indeed what appears to be happening. The peak field goes from just under 4 volts per meter for 200 modes to about 0.3 volts per meter for 5000 modes.

The next figures concern y directed current. Since we will be taking the same cross section parallel to the x axis, we will see the rectangle dependence in the current density. This is shown in Figure 7.7 for a range of included modes. Note that while the approximation improves for an increasing number of modes, the magnitude of the peak ripple (sidelobe) shows little improvement. In fact, the peak ripple corresponds closely to the classic 13 dB first sidelobe of the sin(x)/x function. Returning to the transition into the triangle region of the roof-top pulse (Figure 7.4), we note that the peak

ripple is about 26 dB down from the peak current. This is similar to the peak sidelobe of the square of the sin(x)/x function. We will return to this topic when we discuss the parabolic pulse function.

Figure 7.8 looks at the y directed electric field. In the vicinity of the patch this field shows a large and increasing value as we increase the number of modes. What we are seeing is a broadside view of the singularity in the electric field at the peak of the triangle noted earlier. As the number of modes is increased, the singularity is more closely approximated. Again, this is why we are interested in a weighted integral of the electric field. Once outside the patch and away from the singularity, the electric field becomes better behaved.

The x directed electric field due to y directed current should go to zero along this cross section as the number of modes goes to infinity. Figure 7.9 illustrates the trend.

As mentioned above, for a given number of modes the peak ripple is about 13 dB less for a triangle than it is for the rectangle pulse. Since the triangle is the convolution of a rectangle with another rectangle, one might try convolving a triangle with a rectangle to get a piecewise parabolic pulse. The piecewise parabolic pulse might then have peak ripple down an additional 13 dB for a total of 39 dB down. The advantage being realized here is that the parabolic pulse could be more accurately approximated with fewer modes.

Figure 7.10 shows a parabolic x directed current pulse approximated with 1000 modes. A rectangular dependence is used in the y direction. The x and y directed electric fields are also shown.

At this scale no ripple is even seen in the current pulse; it is well approximated with 1000 modes. Thus we might be tempted to substitute the parabolic pulse for the triangle pulse in our roof-top functions. Since the parabolic pulse can be represented with fewer modes, we might expect a more efficient and more accurate analysis to result. This was attempted and while the resulting analysis was more efficient, it was not more accurate. The reason for this is that the parabolic pulse is restricted to zero derivative at the center and at the ends. Since the actual current distribution on a microstrip circuit rarely has zero derivative at the centers of whatever small subsections we might choose, the parabolic pulse does not represent the actual current distribution as well as the piecewise linear approximation provided by the triangle function. For this reason, the parabolic pulse function was not investigated further.

7.2 Initial Open Circuited Microstrip Stub Analysis

The dimensions selected for this and the other analysis are comparable to dimensions usually selected for "scaled" circuits. Since, in this case, we have nothing from which to scale, referring to these structures as scaled circuits is not appropriate. However, the reason that these dimensions (on the order of centimeters) were chosen is the same reason that the dimensions of scaled circuits are chosen: ease of construction and measurement. While the circuit in this section was not actually built, several others were. Microstrip circuits designed for actual applications typically have dimensions on

the order of $1/10^{th}$ to $1/100^{th}$ of the dimensions given here. Thus when attempting to determine the range of the analysis' validity for actual microstrip design, the frequencies which follow should be multiplied by a factor of 10 to 100.

The first circuit analysis was a microstrip open circuited stub. Dimensions were set at 1.0 cm wide and 2.81 cm long. It was analyzed inside a box 4.0 cm long, 2.0 cm wide and 5.0 cm high. The 1.0 cm thick substrate was specified with a dielectric constant of 10.0. The purpose of the analysis was to check for reasonable analysis results and to take the analysis high enough in frequency so that it would break down.

The input impedance of the structure was calculated over a range from 50 to 4000 MHz with the results presented in Figure 7.11. The stub input impedance appears to be the classical cotangent function. Closer inspection shows that the impedance plot is not symmetrical about zero, especially at the higher frequencies. This is explained, at least in part, by the fringing capacitance between the base of the stub and the adjacent sidewall.

Above 3.0 GHz, we seem to have taken the analysis past a point of failure. However, when the analysis is repeated over that range with a step of 10 MHz, rather than 50 MHz, we see that reasonable data is still obtained, made reasonable by the observation that at 3.0 GHz, the substrate is electrically about one half wavelength thick. What we are seeing, in Figure 7.12, is a spectrum of higher order microstrip modes. The 50 MHz step size was insufficient to resolve the higher order modes.

Caution should be exercised here as the higher order modes predicted by the analysis have not been checked for validity by measurements. Several other stubs have been constructed and checked for accuracy as will be described in the next several sections.

7.3 Microstrip Open Circuited Stub Measurements

A second microstrip stub was built and measured as well as analyzed. This stub was 2.54 cm wide and 10 cm long contained in a box 13.0 cm long, 7.9 cm wide and 5.0 cm high. To ease fabrication requirements, air was used as a dielectric throughout the box.

Two measurements were made using a Hewlett Packard 8510 automated network analyzer. The first measurement was of the stub itself. Since the network analyzer was calibrated using standards external to the circuit, the measurement reference plane was just outside the box. In order to transfer the reference plane just inside the box, the end of the input connector was shorted to the adjacent wall of the box with a short copper strap and a second measurement was taken. The phase of this measurement was, in effect, the phase length of the input connector.

In both cases, the magnitude and angle of the reflection coefficient was measured. The magnitude information (which was close to 1.0) was not used. To use the angle information, the phase length of the connector was subtracted out. From the internal short circuit measurement, it was determined that the connector phase length was 3.60 +/- 0.05 degrees per 100 MHz. It was later found that a phase

length of 3.28 +/- 0.05 degrees per 100 MHz provided a better fit. The difference of 0.32 degrees per 100 MHz can be attributed to the small inductance in the shorting copper strap. After the connector phase length was subtracted out, the phase information was converted to reactance and plotted.

The initial comparison between measured and calculated data is shown in Figure 7.13. The difference between measured and calculated increases with increasing frequency and at high impedance levels. The direction of the discrepancy suggests that a shunt capacitance at the base of the stub is unaccounted for in the analysis. This shunt capacitance can be attributed to the fringing capacitance of the circular coaxial aperture formed by the input SMA connector.

A single frequency was selected, 1.5 GHz, and the required shunt capacitance calculated, 0.56 pF. This capacitance was then placed in shunt with the calculated stub input impedance at all frequencies and replotted. The result is shown in Figure 7.14. The agreement is substantially improved, especially below 2.3 GHz. The discrepancy above 2.3 GHz may be due, in part, to what may be viewed as the input connector fringing capacitance increasing with frequency. Several other possibilities are explored next.

7.4 Sources of Error at High Frequencies

After the above measurements had been made, it was found that the stub (which was formed of copper tape, cut to size with a razor knife) was nearly a millimeter longer than 10.0 cm. To check the sensitivity

of the calculations to a small change of length, the stub was then reanalyzed with a length 10.1 cm. The best fit with measured data was obtained with a connector phase length of 2.94 degrees per 100 MHz and fringing capacitance of 0.67 pF. Note that the calculated data, Figure 7.15, now fits well to 3.5 GHz. The changes in the best fit connector phase length and connector fringing capacitance suggest that the evaluation of these quantities is strongly dependent on the precise dimensions of the microstrip circuit.

A second potential source of error is in the accuracy of the analysis itself. To check this possibility, two additional analyses were performed. The first simply increased the number of subsections. The original analysis subdivided the stub into 5 by 9 subsections for x directed current and nearly an equal number for y directed current. A second analysis subdivided the stub into 7 by 14 subsections for x directed current. The results are shown in Figure 7.17. The crosses indicate the selected frequencies at which the more detailed analysis was performed. Due to the time required to calculate each point (several hours on an IBM-PC), the analysis was not performed at all frequencies. As can be seen, there is some difference at higher frequencies. The difference is in the right direction to account for some of the observed discrepancies.

Another possibility for error is that an insufficient number of waveguide modes were included to properly represent the roof-top current distribution on each subsection. To check this possibility, the number of included modes was increased by a factor of 25. This was done by changing the upper limit on the first stage summation (see

the section on efficient computation) from one term (or 'cycle') to five terms. The number of included modes goes with the square of that upper limit. This particular analysis was for a notched stub (discussed in a subsequent section), but the result still holds. Figure 7.17 shows the two analyses and we find almost no difference between the two. Thus the number of included modes appears to be a second order factor.

A check for another kind of problem was made using this geometry. It was suspected that there might be a numerical problem at the resonant frequency of the empty (no microstrip circuit) box. With the two largest dimensions of 13.33 cm by 7.257 cm, the resonant frequency is 2351.79998 MHz. An analysis was performed from 2350.8 to 2353.55 every 50 KHz with no numerical problem becoming evident. The resulting input impedance was on the order of 276 Ohms and was a smooth function of frequency.

As mentioned in a previous section, we are using dimensions corresponding to a scaled circuit. Thus, for practical microstrip design, the frequency range for valid analysis will be much larger than that cited for the above analysis.

7.5 Microstrip Stub Current Distributions

The current distribution on the microstrip stub was plotted at a number of frequencies. The current distribution is shown conceptually in Figure 7.18. The source current is injected (by a coaxial connector penetrating the shielding box) into the base of the stub.

From that point, the current immediately flows out to the edge of the microstrip stub. Now, with most of the current concentrated on the edge of the stub, the current flows down to the end of the stub, slowly and smoothly tapering off to zero as it reaches the end of the stub. Since the current distribution is symmetrical about the center line of the stub, the subsequent plots will show only the lower half of the stub. To assist orientation, the edge of the microstrip stub is also outlined in the plots.

The current distribution is indicated on the plots by a set of arrows. The length of the arrow indicates the magnitude of the current. There is one arrow per subsection. Since a subsection will have either x or y (and not both) directed current, each arrow is directed in either the x or the y direction. Since the x directed subsections are offset by $\Delta x/2$ and $\Delta y/2$ with respect to the y directed subsections, we will find the arrows have the same offset. Any arrow which is less than four printer dots long is left without an arrowhead. All subsections are plotted with at least one dot, no matter how small their current.

Figure 7.19 shows the current distributions for frequencies from 500 MHz to 3000 MHz. At 500 MHz, we see the current injected into the center of the stub in the upper left corner of the plot. The current then proceeds down (and up, not shown due to symmetry) to the edge of the microstrip stub. Then the current propagates along the edge of the stub, slowly and smoothly tapering off to zero as it reaches the end of the stub. At this frequency, the stub is a little less than $1/8^{th}$ of a wavelength long.

There is an unusual lateral current near the source. It is believed that this current is a numerical artifact due to the modeling of the source.

At 1000 MHz, we see much the same distribution except that the current does not go quite as smoothly to zero. This appears to be the result of the increasing influence of fringing capacitance off the end of the stub. The current does actually go to zero at the exact end of the stub because we use a triangle dependence for the current density in the direction of current flow and the end of the triangle is at the end of the stub.

At 1500 MHz, the stub is now a little over one half wavelength long. Note that on the edge of the stub, the current reversal occurs near the corner of the stub while on the interior of the stub, the current reversal occurs a full subsection further down the length of the stub.

At 2000 MHz, the current reversal has moved down the length of the stub and the current reversal now takes place at about the same point on both the edge and the interior of the stub.

At 2500 MHz, the stub is now very close to one wavelength long. The first current reversal actually occurs before the current reaches the edge of the stub. The second current reversal again occurs sooner on the edge of the stub than in the interior. Also note that the effect of fringing capacitance off the end of the stub is becoming more pronounced.

At 3000 MHz, the effect of fringing capacitance is even more pronounced. An additional effect is also starting to appear at this

point. The lateral current at the end of the stub is starting to grow. At higher frequencies, the current will actually wrap around the corner of the stub, further increasing the electrical length of the stub.

The first current reversal on the edge of the stub has now rounded the first corner. The first current reversal on the interior of the stub occurs further down than on the edge. Surprisingly, the second current reversal is aligned across the entire width of the stub.

Figure 7.20 shows the current distributions of the same stub now analyzed with 7 by 14 (x directed) subsections instead of the previous 5 by 9 subsections. The distributions are essentially the same as above only represented in finer detail.

7.6 Notched Stub Measurement

The microstrip stub of the previous section was modified by cutting a notch midway along the length of the stub. The notch was 2.0 cm long and 0.5 cm deep, changing the width of the line from 2.54 cm to 1.54 cm. Both the stub and a short circuit were then measured. The internal short provided an estimate of 3.54 +/- 0.05 degrees/100 MHz for the phase length of the connector. The best fit between calculated and measured data occurred with a connector phase length of 3.24 +/- 0.05 degrees/100 MHz indicating a shorting strap inductance of 0.30 degrees/100 MHz. This agrees well with the results from the previous stub (3.60 and 3.28 degrees/100 MHz).

Figure 7.21 shows the initial comparison between measured and calculated data with no compensation for the connector fringing capacitance. The shunt capacitance required to bring the calculated data in line with the measured data at 1500 MHz is now 0.29 pF. This value is significantly smaller than the previous case. As already pointed out, the discrepancy is probably due to dimensional inaccuracies in stub fabrication.

The result of including the connector fringing capacitance at all frequencies is shown in Figure 7.22. Agreement is excellent up to 2300 MHz. Inaccuracies above 2300 MHz could be explained in much the same manner as for the previous case.

Figure 7.23 shows the calculated data for both the notched and unnotched stub. We see that at very low frequencies (less than 400 MHz), where the stub may be viewed as a parallel plate capacitor, introducing the notch decreased the capacitance and made the reactance more negative. At about 1000 MHz, we see that the first resonance has moved lower in frequency, due to the expected inductive nature of the narrower width in the region of the notch. We can see that at many frequencies, the differences between the notched and unnotched stub calculations, while small, are greater than the differences between measured and calculated data in either case.

Figure 7.24 shows the current distribution on the notched stub at various frequencies. The current is seen to flow around the notch in all cases. An unusual lateral current, smaller but similar to that in the vicinity of the source, is also seen near the beginning and end of the notch. The locations of the current reversals on the interior of

the stub are also affected by the notch.

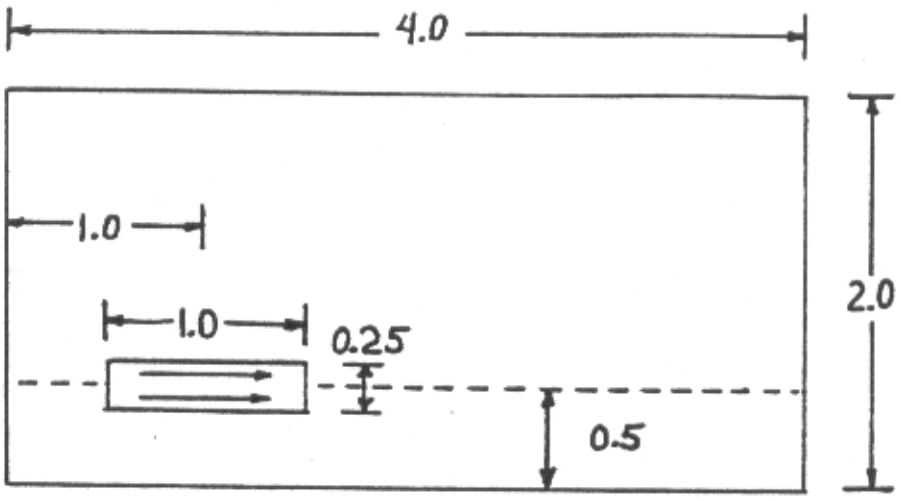


Figure 7.1. Geometry of the x directed current patch used for subsequent calculations of the resulting electric field. The dashed line indicates the cross section along which the analyses will be performed. All dimensions are in cm.

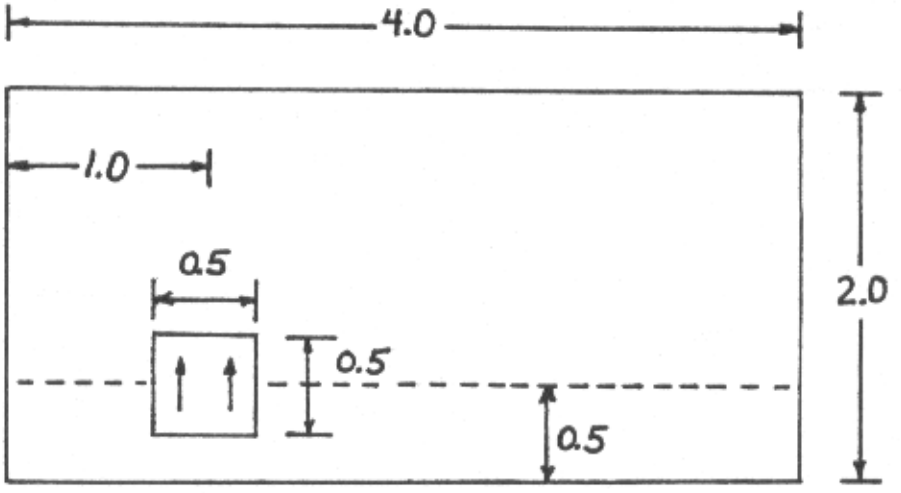
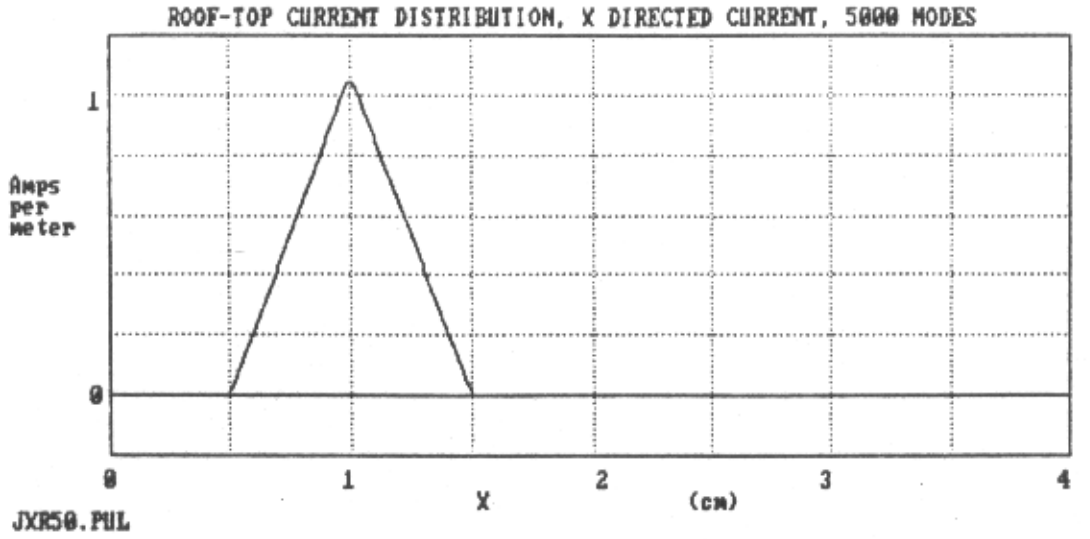
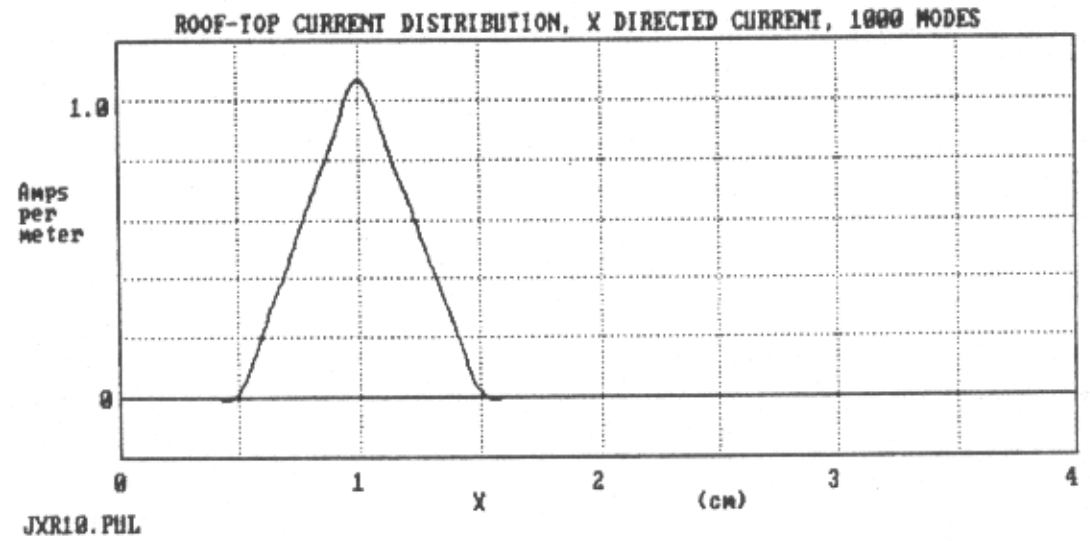


Figure 7.2. Geometry of the y directed current patch used for subsequent calculations of the resulting electric field. The dashed line indicates the cross section along which the analyses will be performed. All dimensions are in cm.





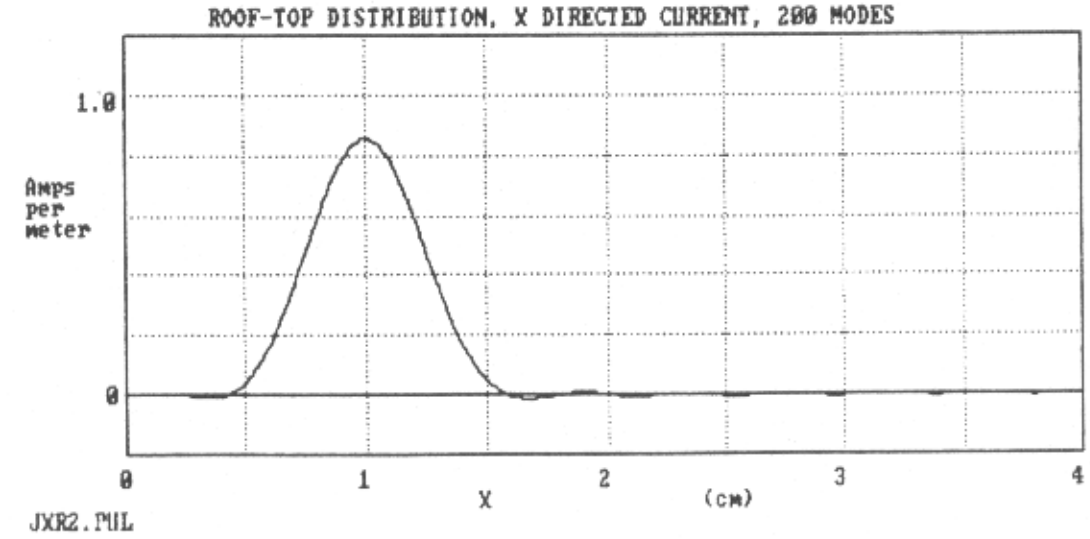


Figure 7.3. Effect on resulting surface current of truncating the summation.

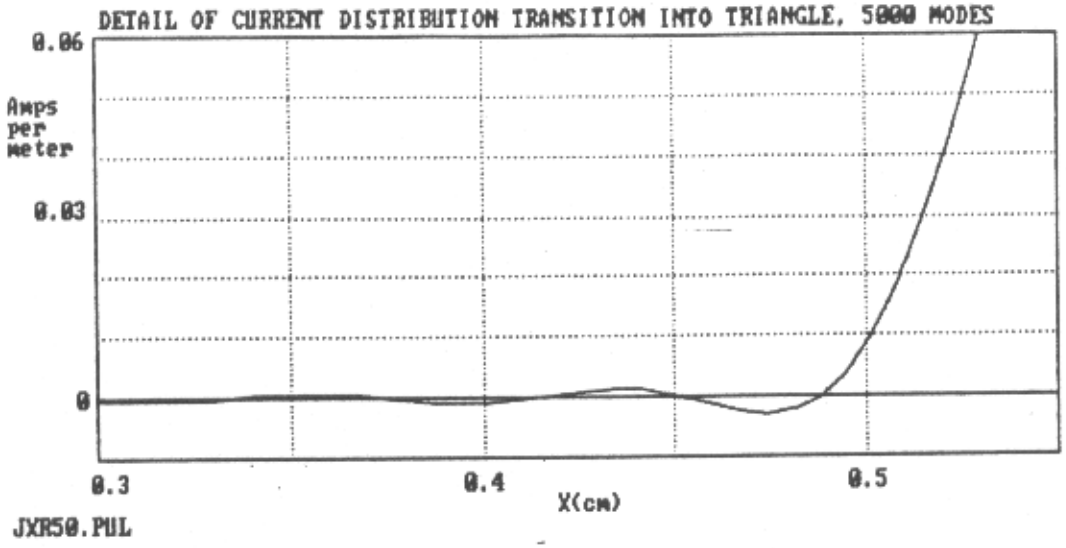
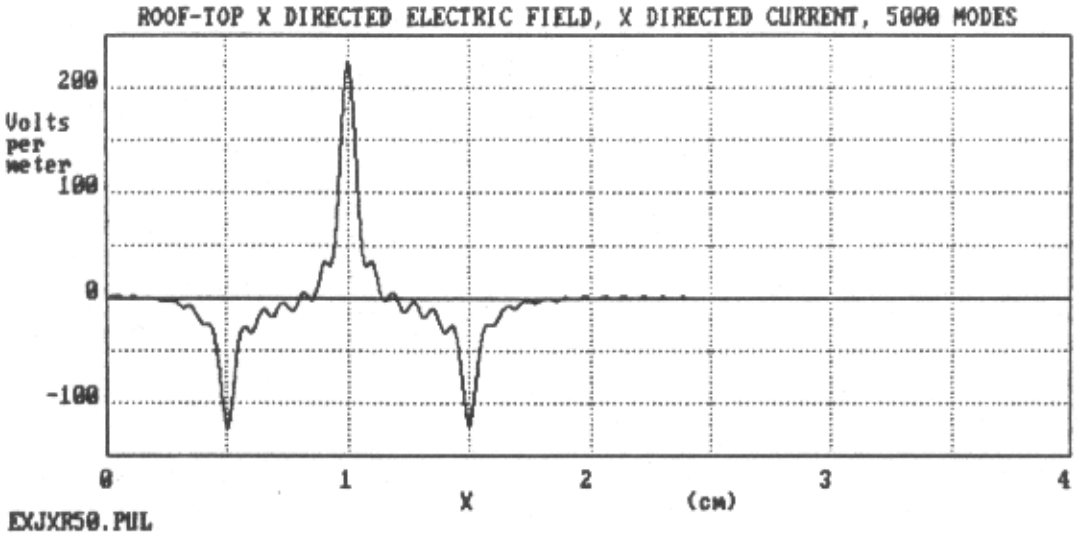
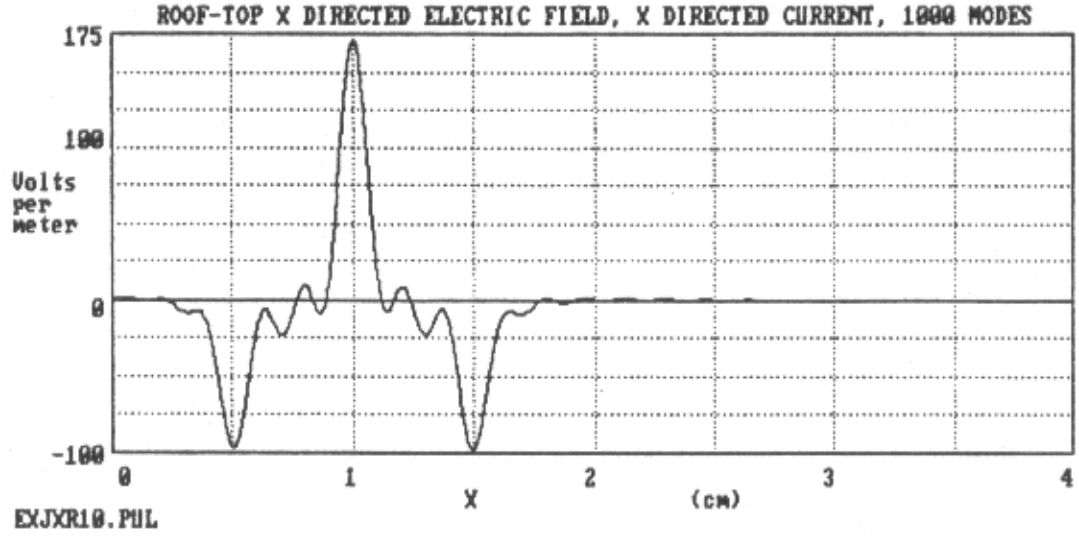


Figure 7.4. Detailed view of the transition region of the triangle function with 5000 included modes.





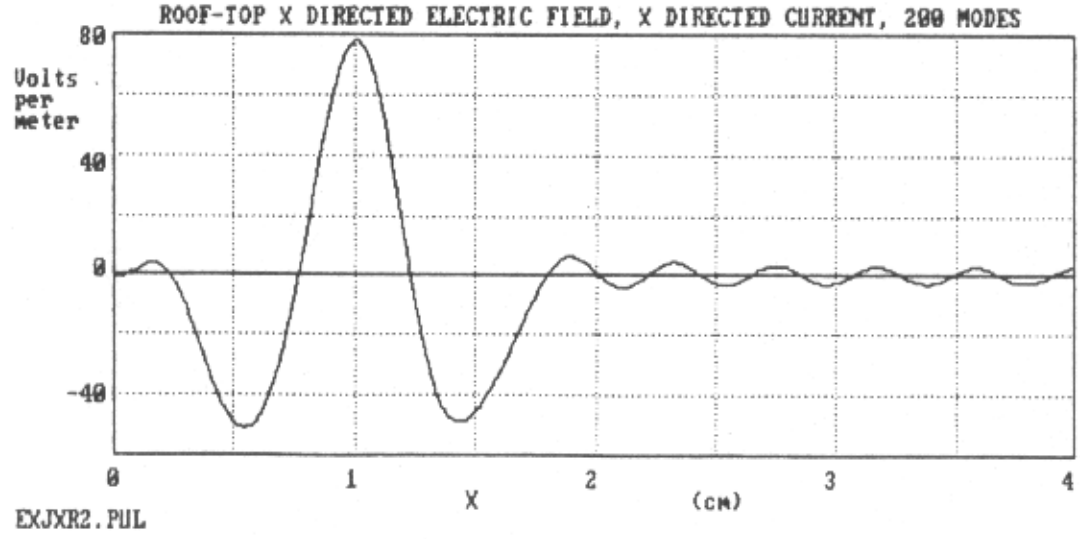
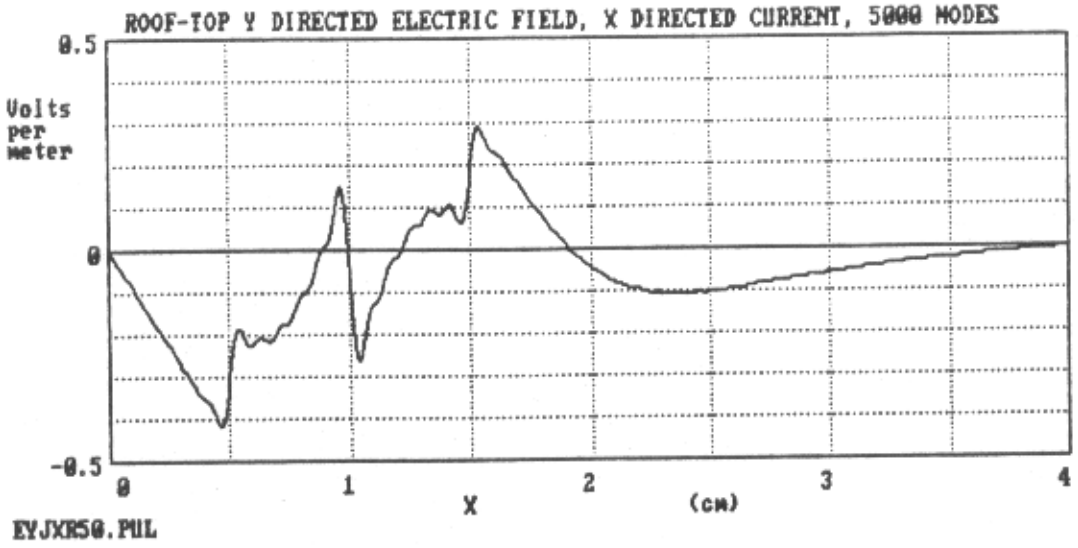
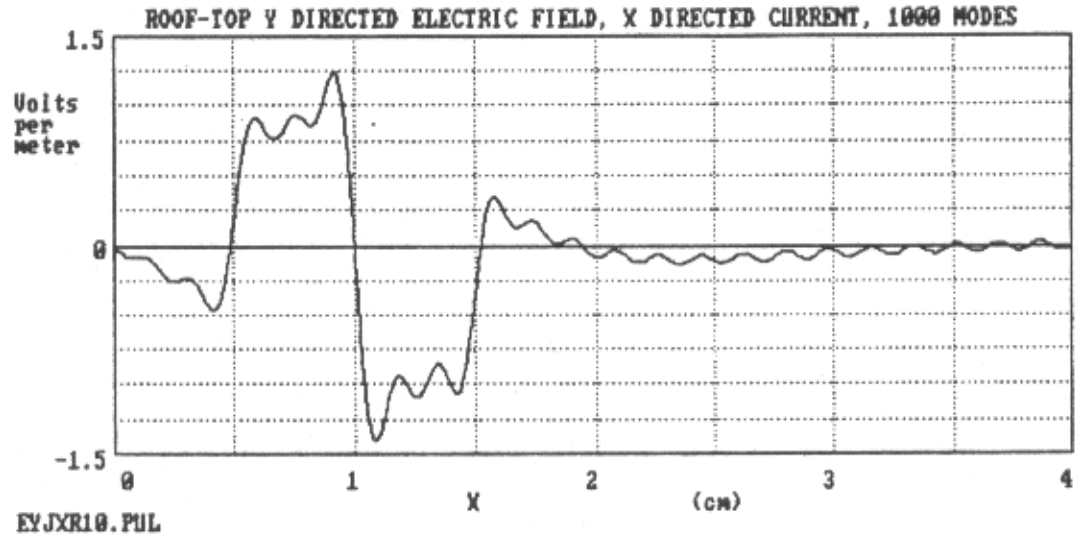


Figure 7.5. Electric field corresponding to the current in Figure 7.3.





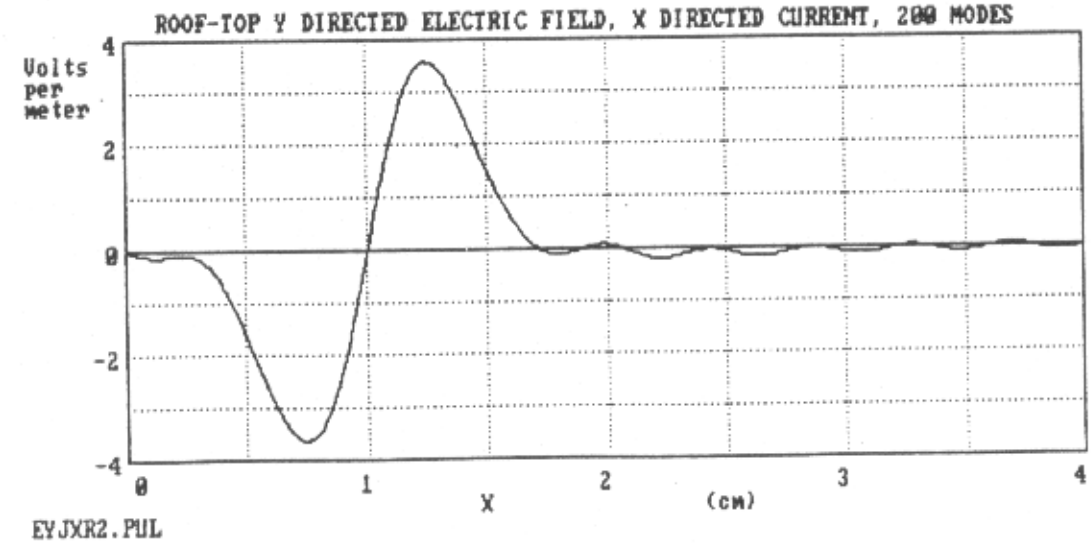
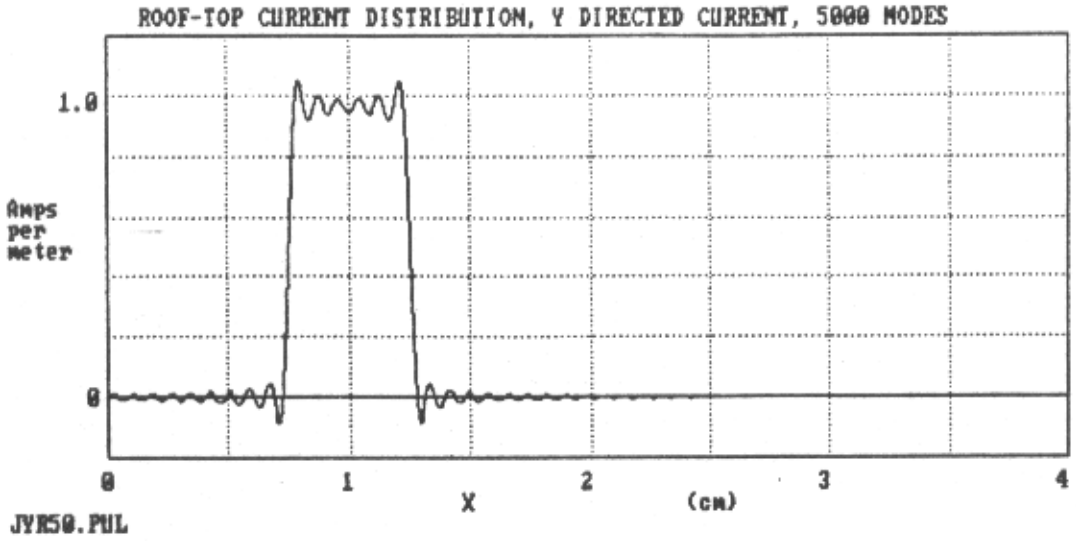
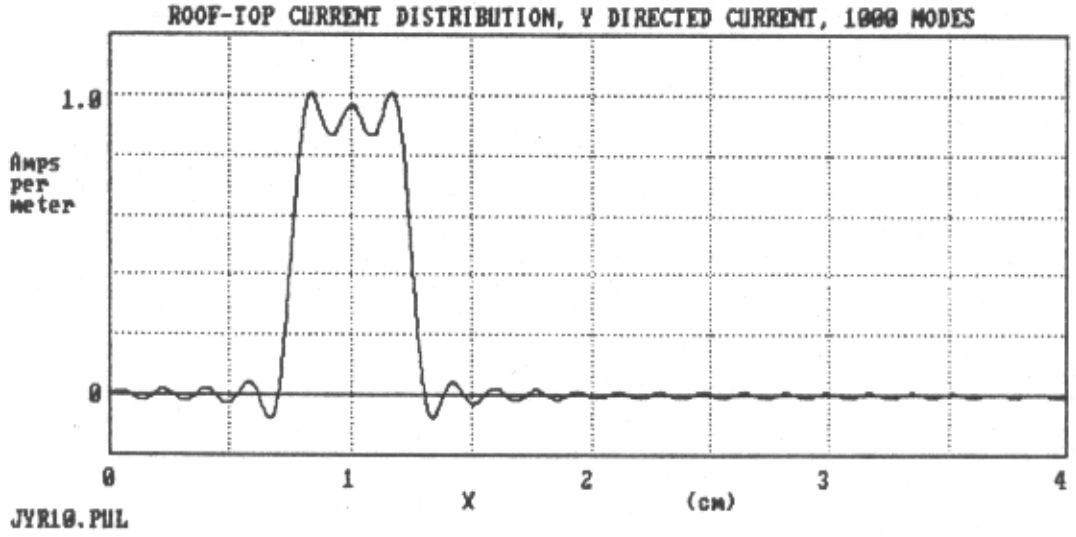


Figure 7.6. This electric field approaches zero as accuracy increases.





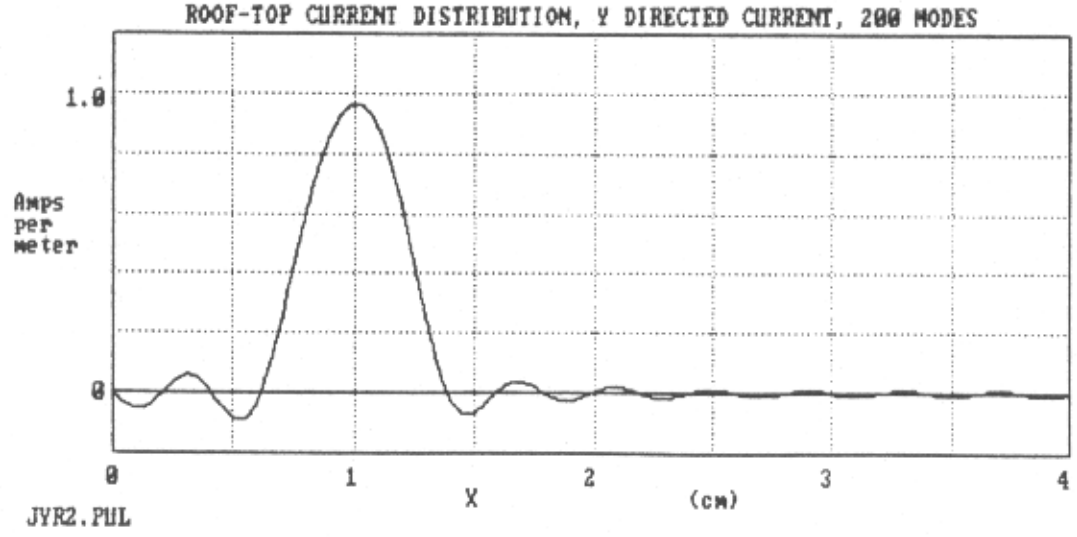
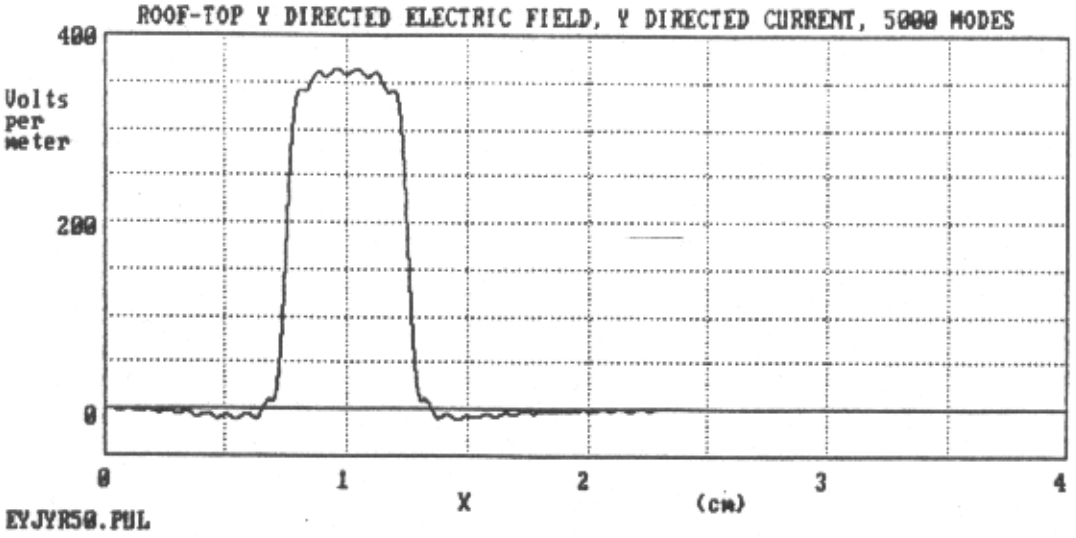
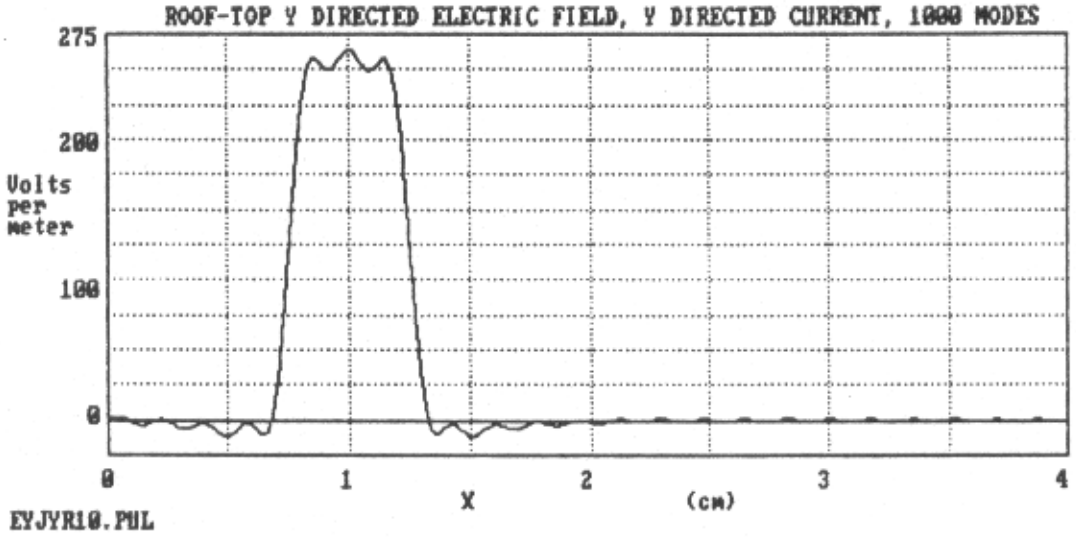


Figure 7.7. Effect on resulting surface current of truncating the summation.





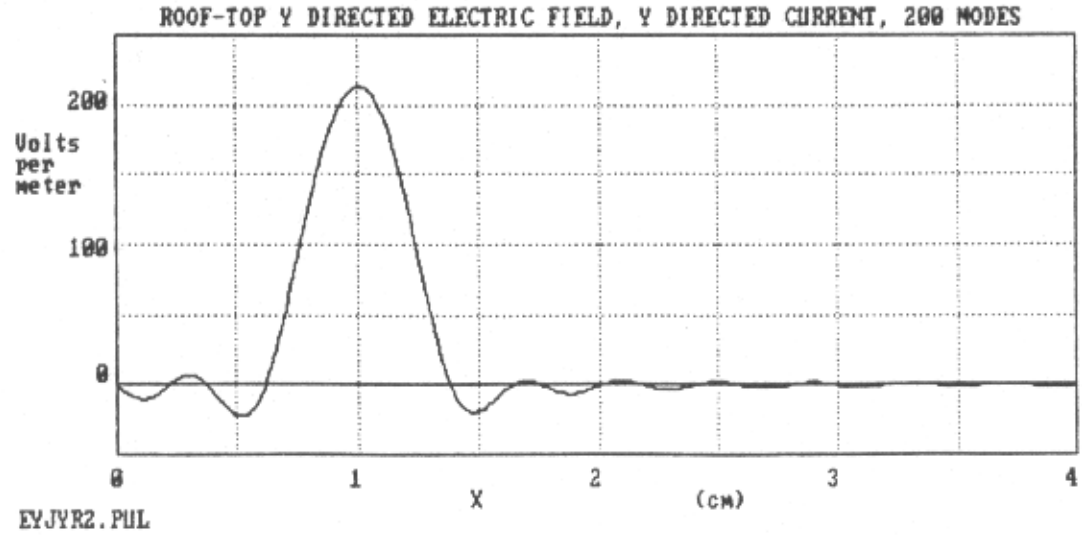
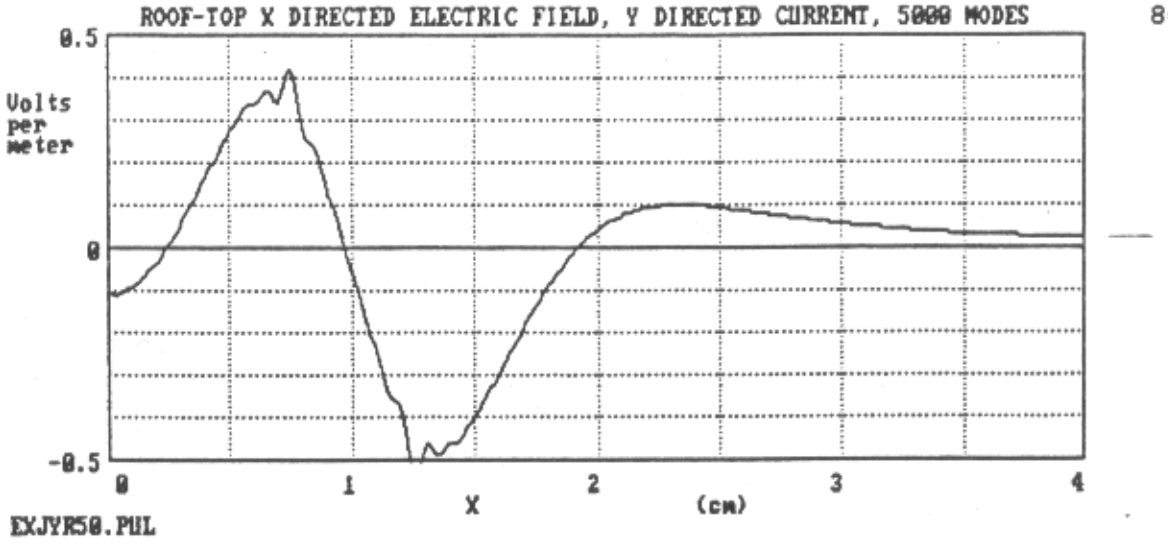
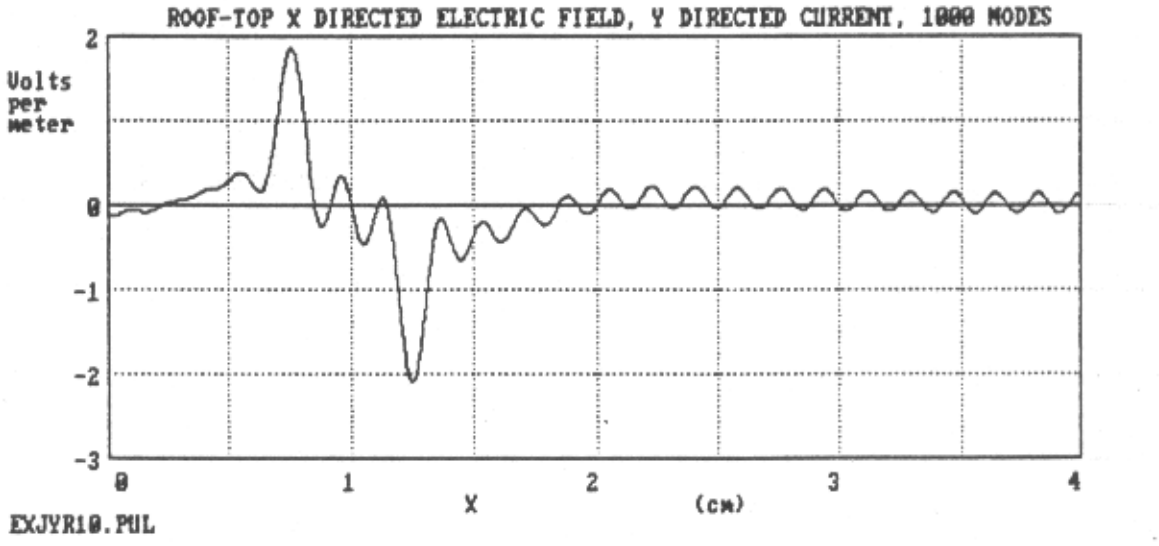


Figure 7.8. Electric field corresponding to the current in Figure 7.7.





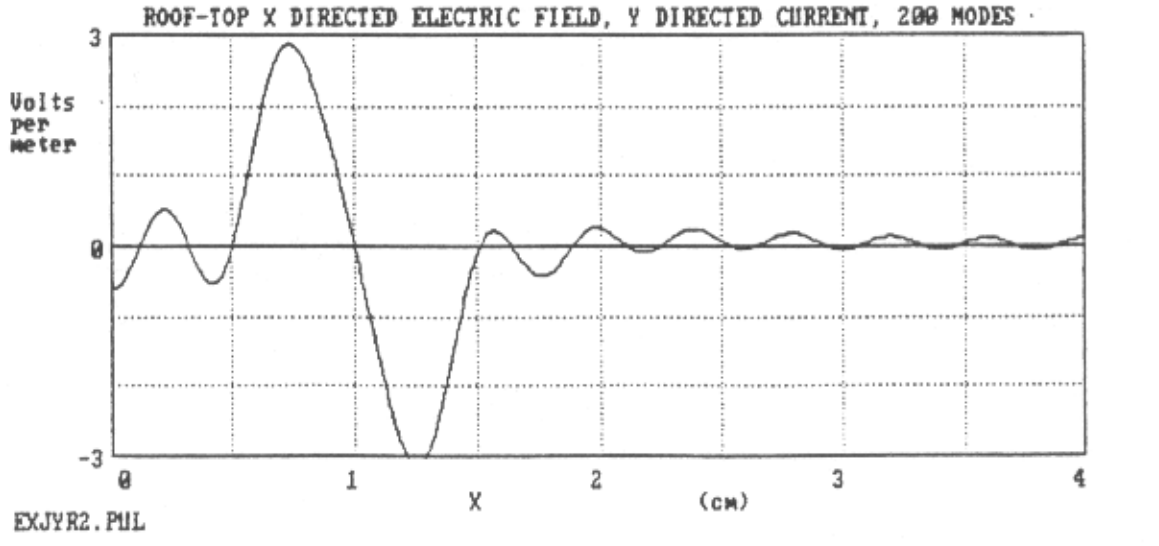
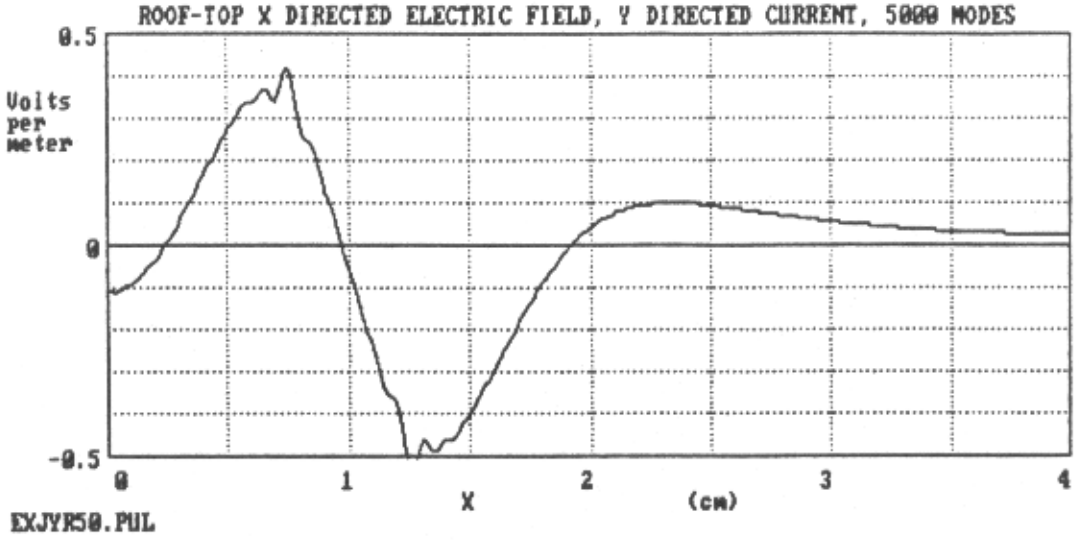
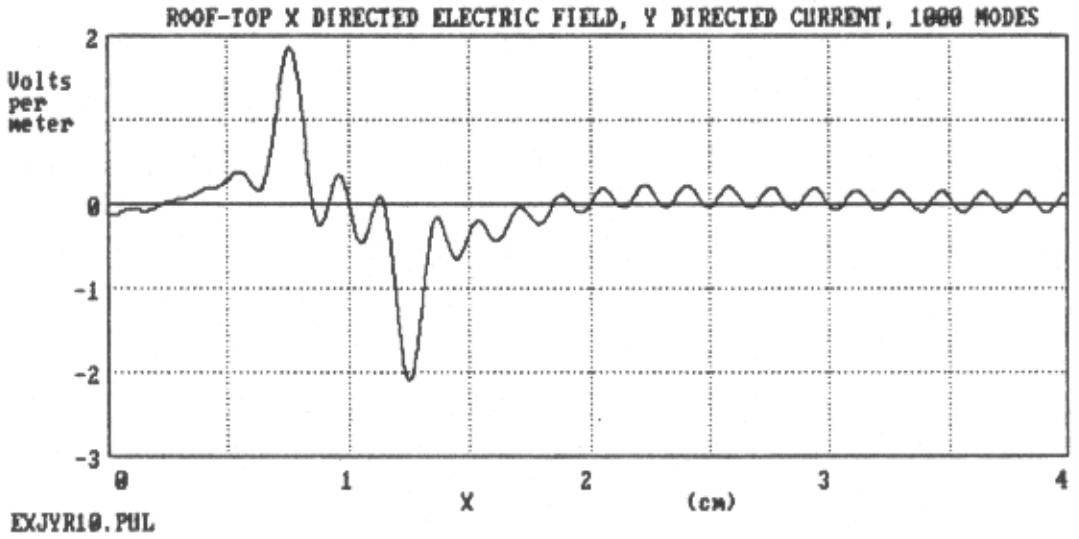


Figure 7.9. This electric field approaches zero as accuracy increases.





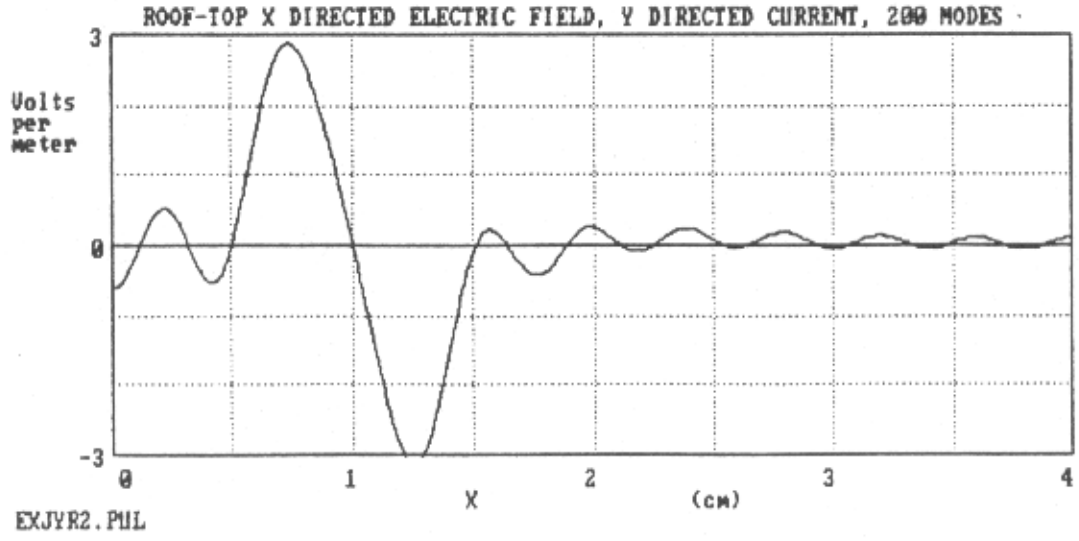
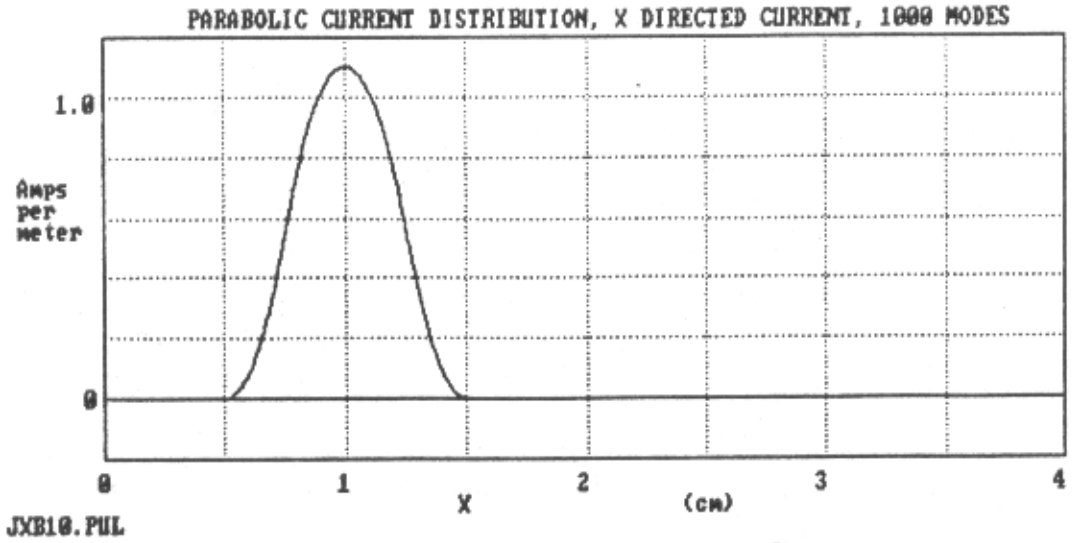
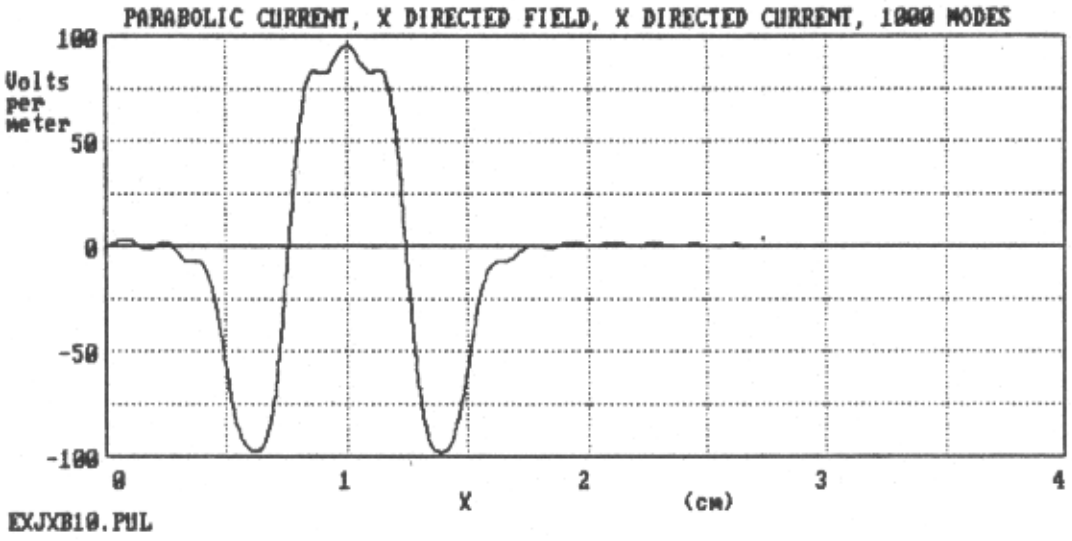


Figure 7.9. This electric field approaches zero as accuracy increases.





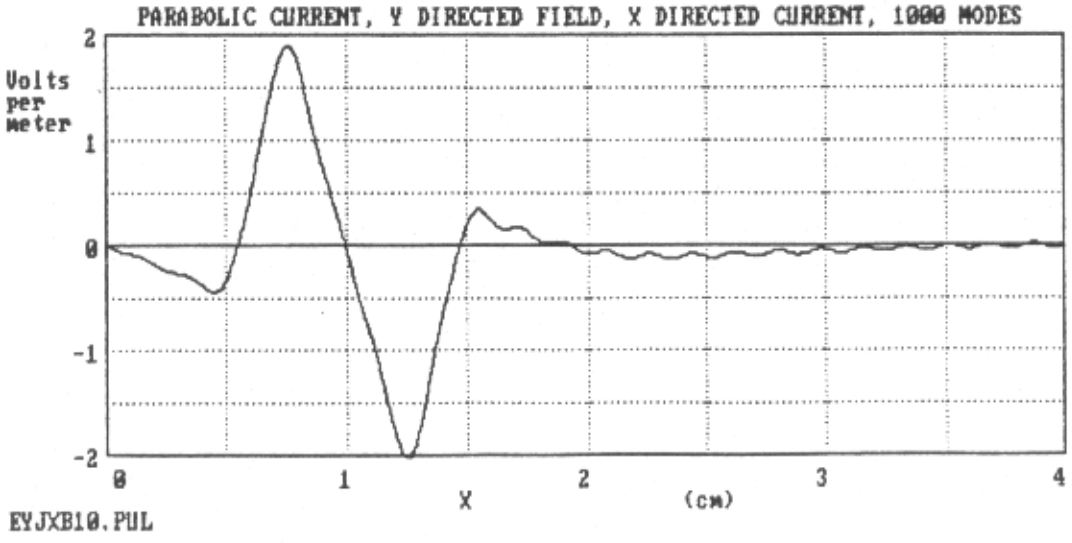
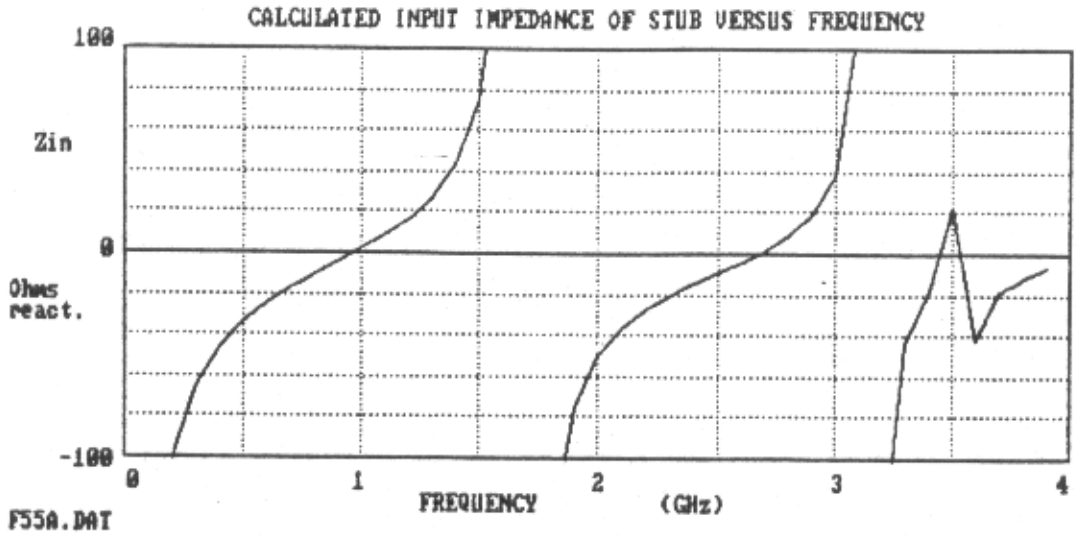


Figure 7.10. Parabolic pulse and associated fields.



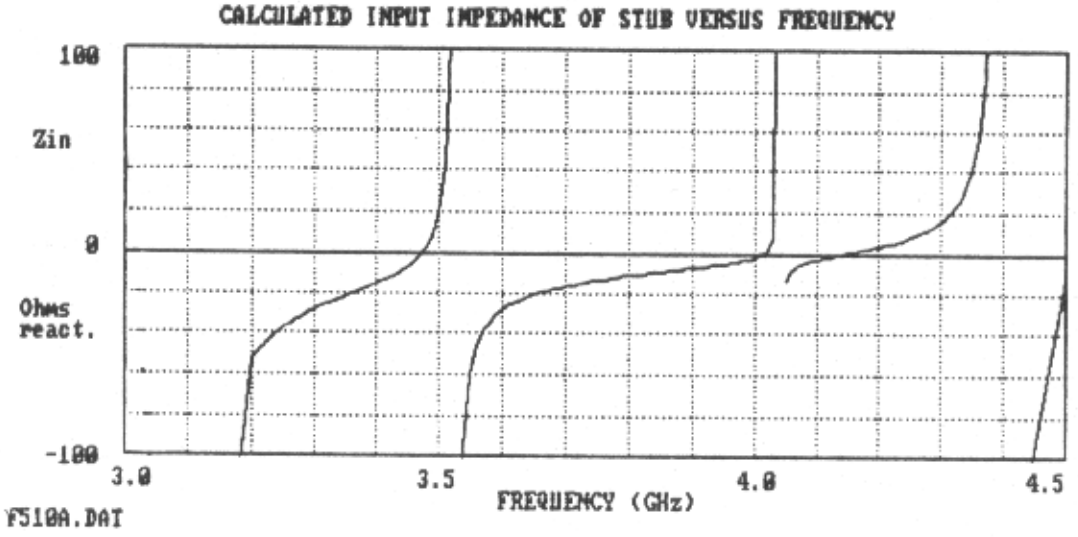
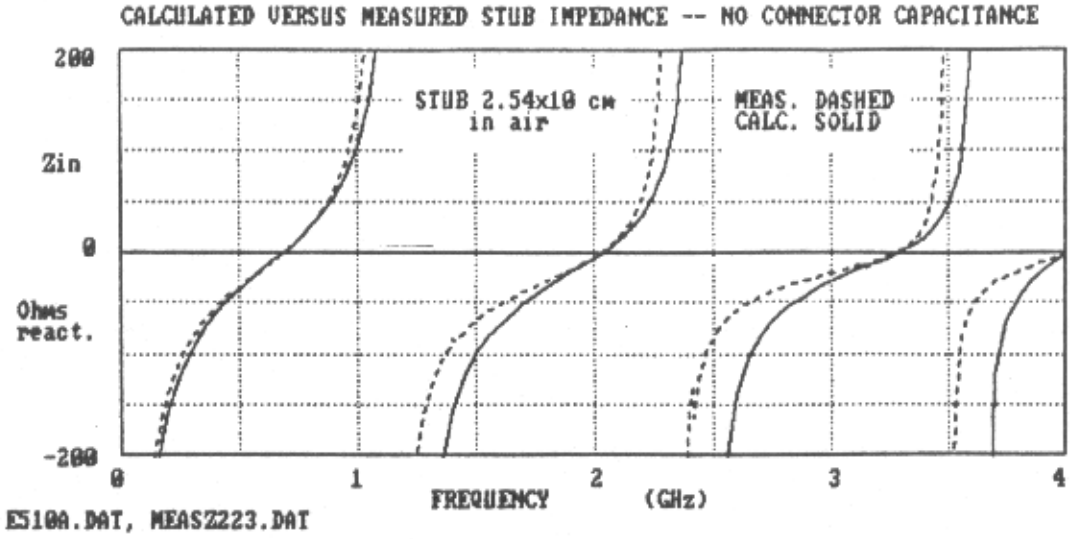


Figure 7.11. (TOP) Initial analysis of an open circuited stub shows siginficant asymmetry in the calculated input impedance as compared with the usual cotangent function predicted by circuit theory.

Figure 7.12. (BOTTOM) A detail of the higher frequencies of Figure 7.11 shows what appear to be a wealth of higher order microstrip modes. We see that the analysis is still providing reasonable results, once we select a frequency step size small enough to display the detail.



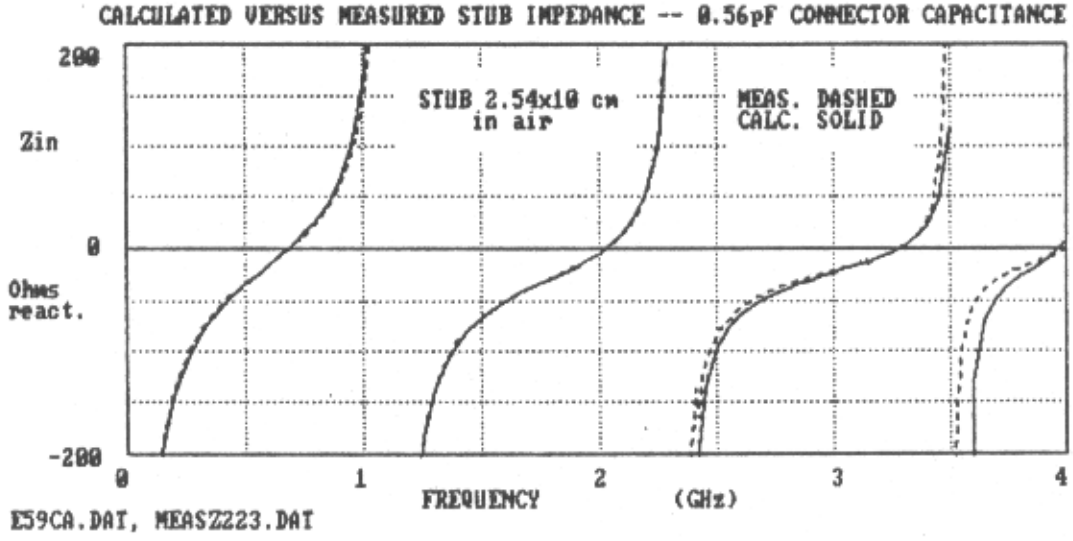


Figure 7.13. (TOP) The initial comparison between measured and calculated for an open stub shows error at high frequencies and at high impedance levels. Fringing capacitance from the input SMA connector is suspected.

Figure 7.14. (BOTTOM) After adding a shunt capacitor to the calculated data, significant improvement in the agreement between measured and calculated data is realized.

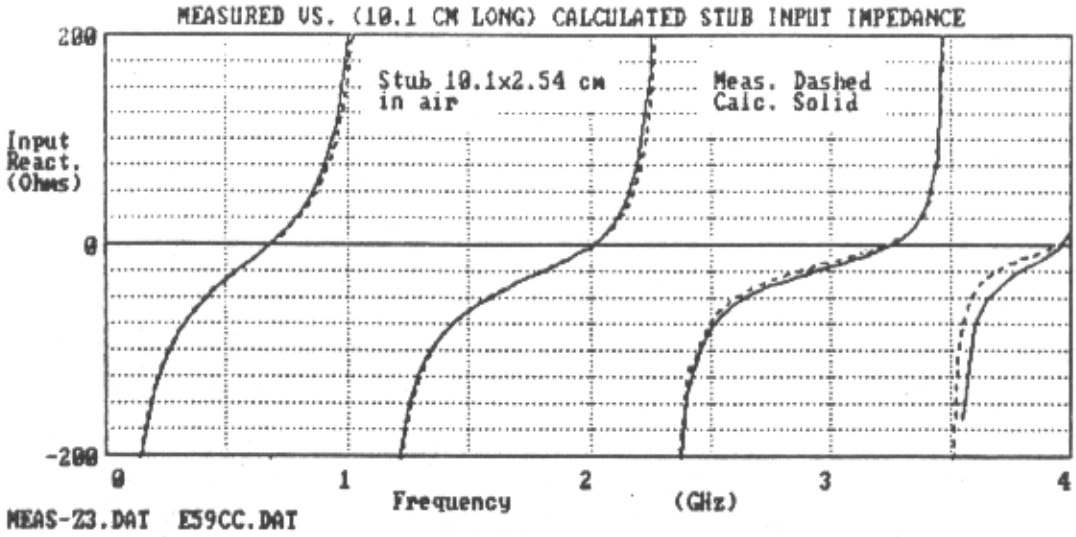
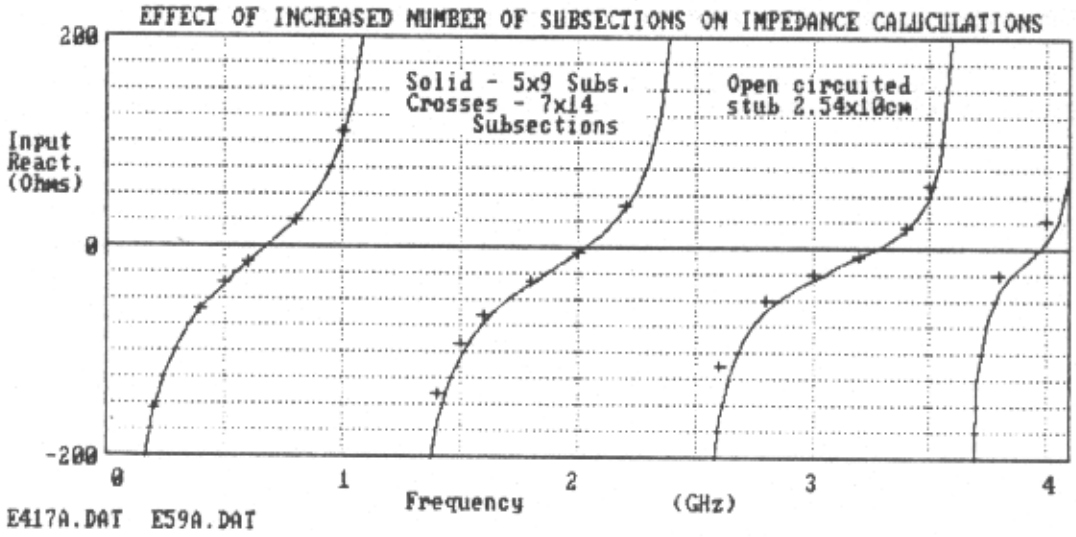


Figure 7.15. Changing the length of the stub in the calculated data by 0.1 cm to compensate for an error in measuring the physical length of the stub provides better agreement.



Figrue 7.16. Increasing the number of subsections for the stub representation provides little change in the resulting data suggesting that the analysis has converged. The small changes at higher frequencies are in a direction to reduce the differences between measured and calculated data in Figure 7.14.

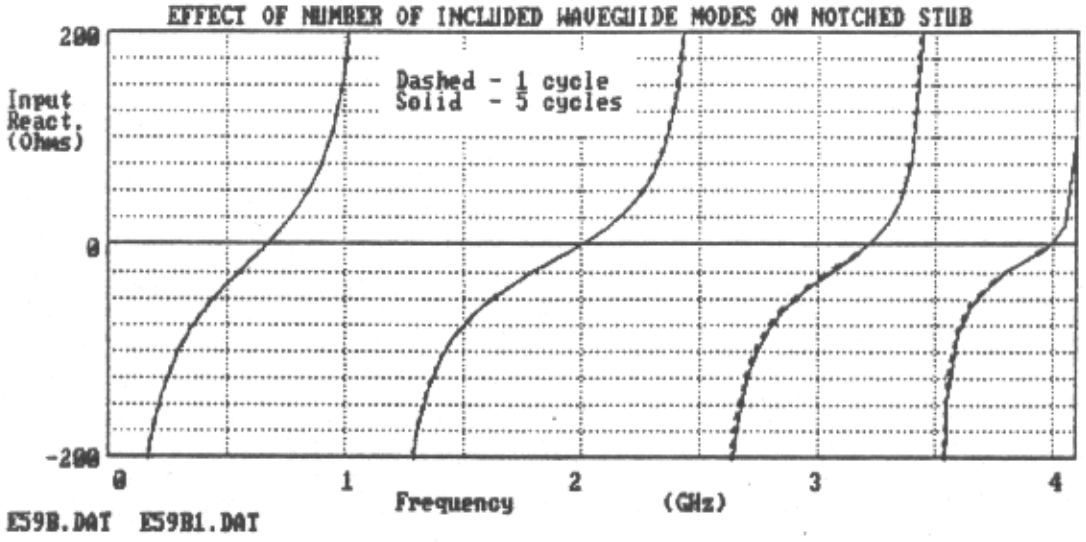


Figure 7.17. Increasing the number of modes used to represent the roof-top current distributions makes little difference in the resulting data.

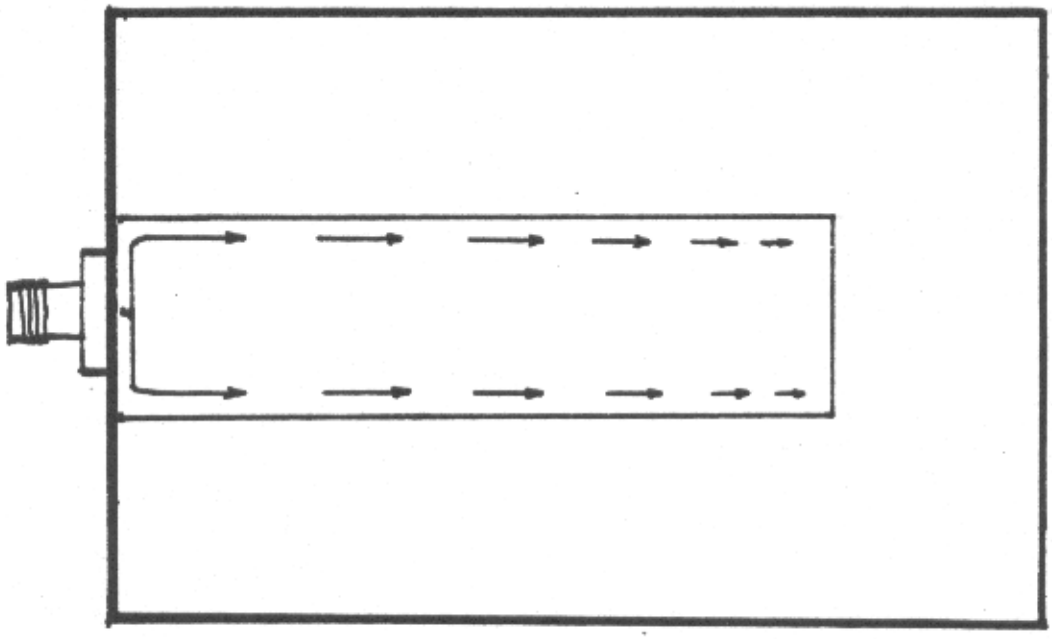
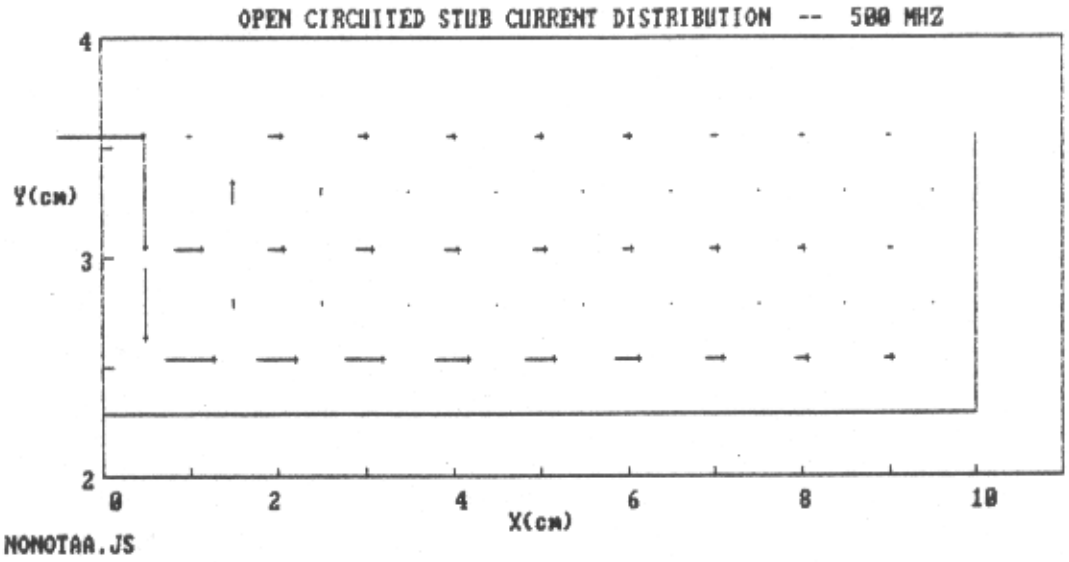
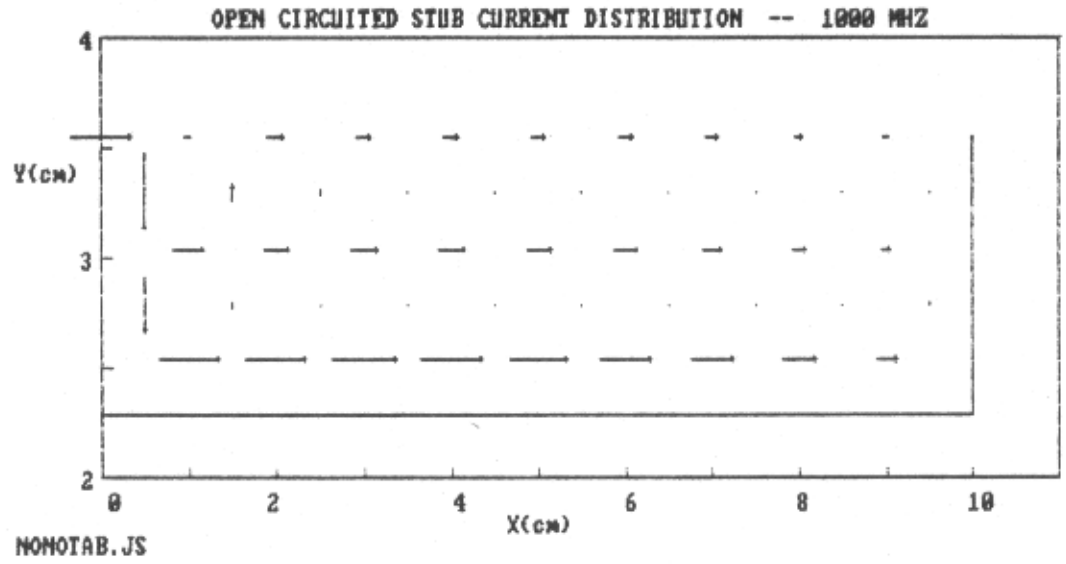


Figure 7.18. Conceptual representation of the current distribution on a stub. The current is injected into the center of the stub by the connector. The current then flows immediately to the edge of the stub then propagates down along the edge smoothly going to zero at the end of the stub.





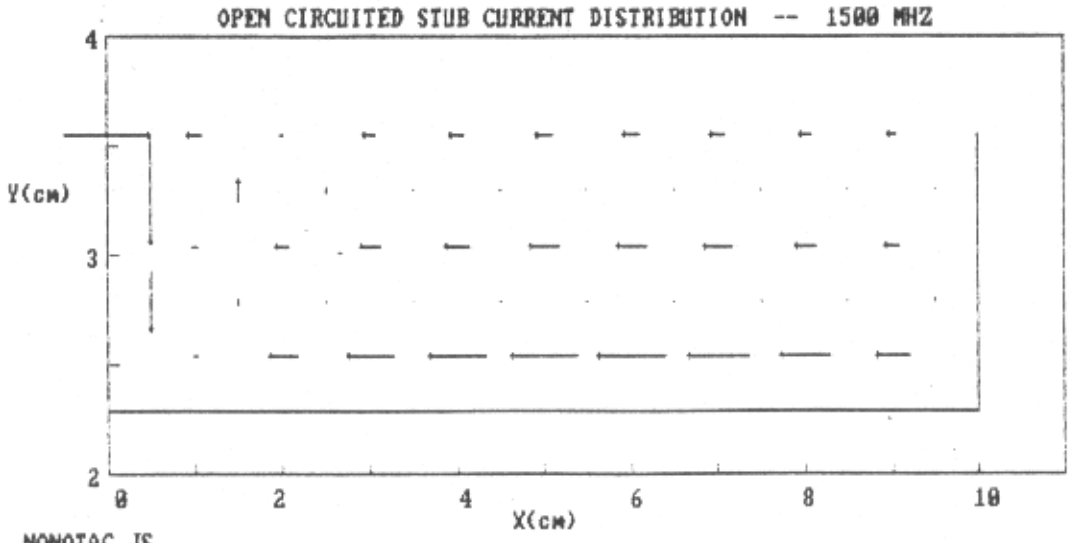
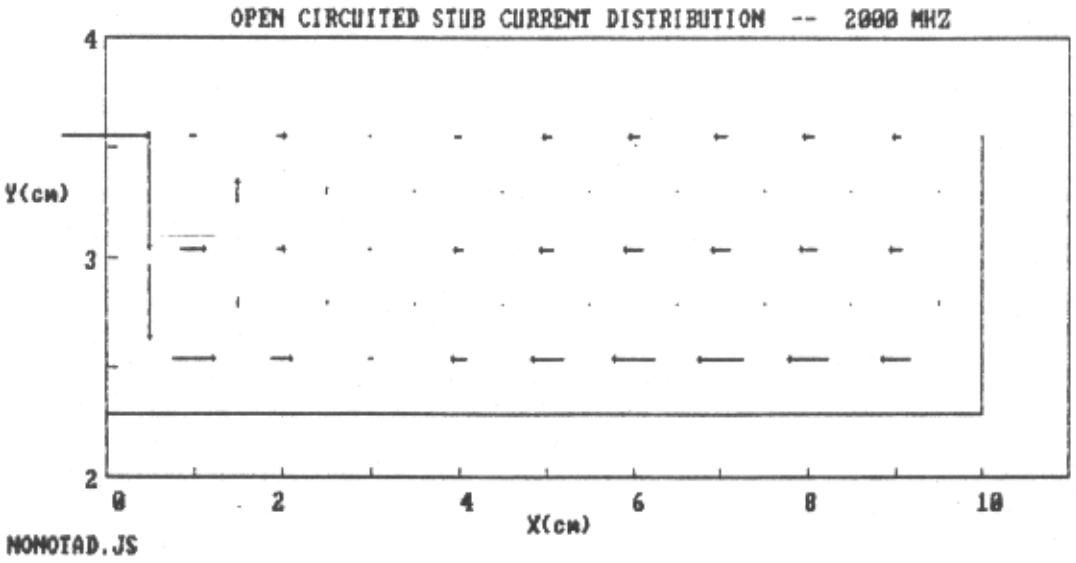
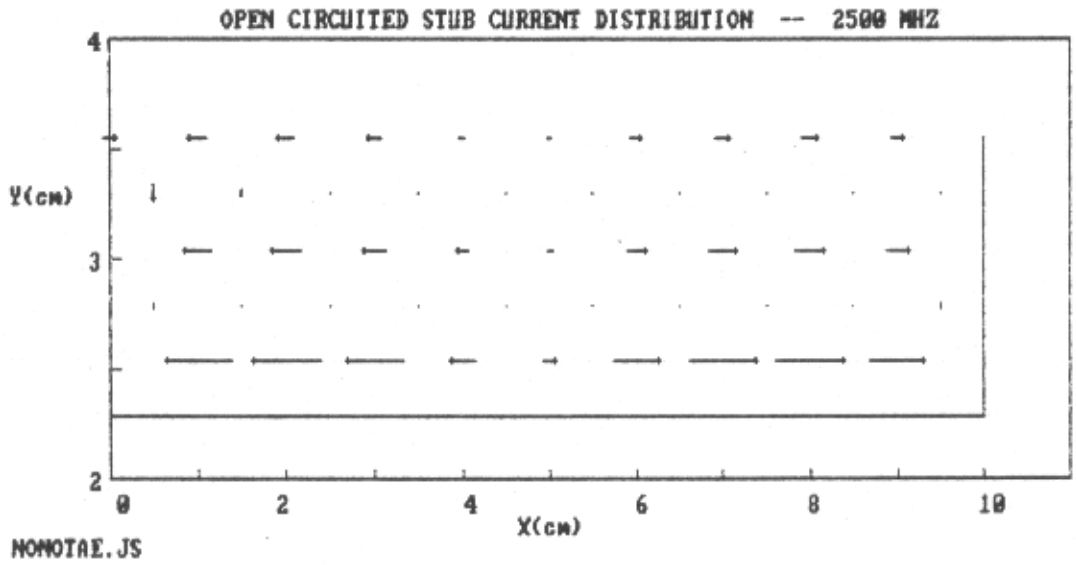


Figure 7.19. Stub current distribution (part 1).





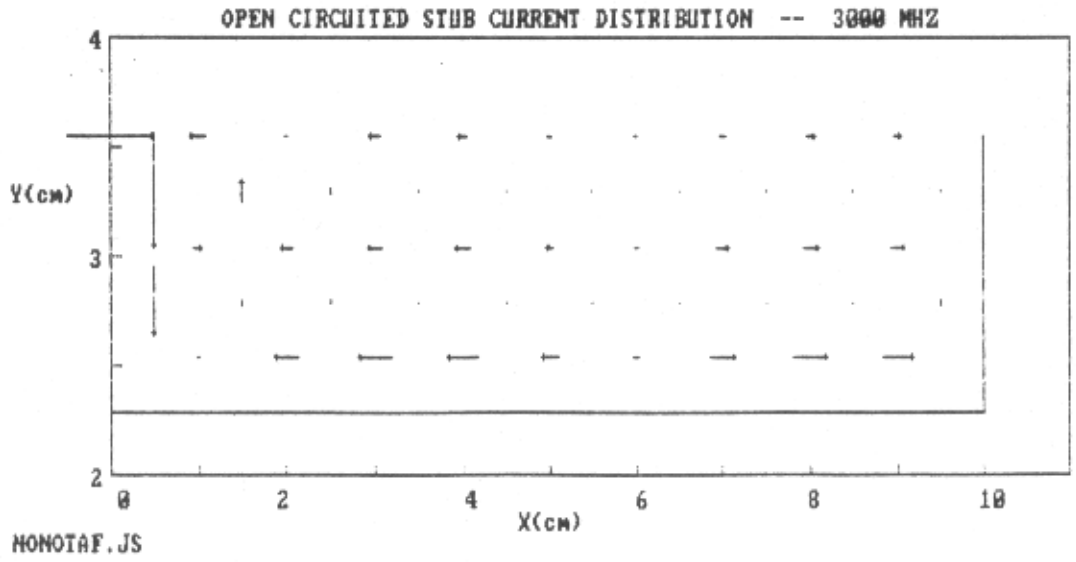
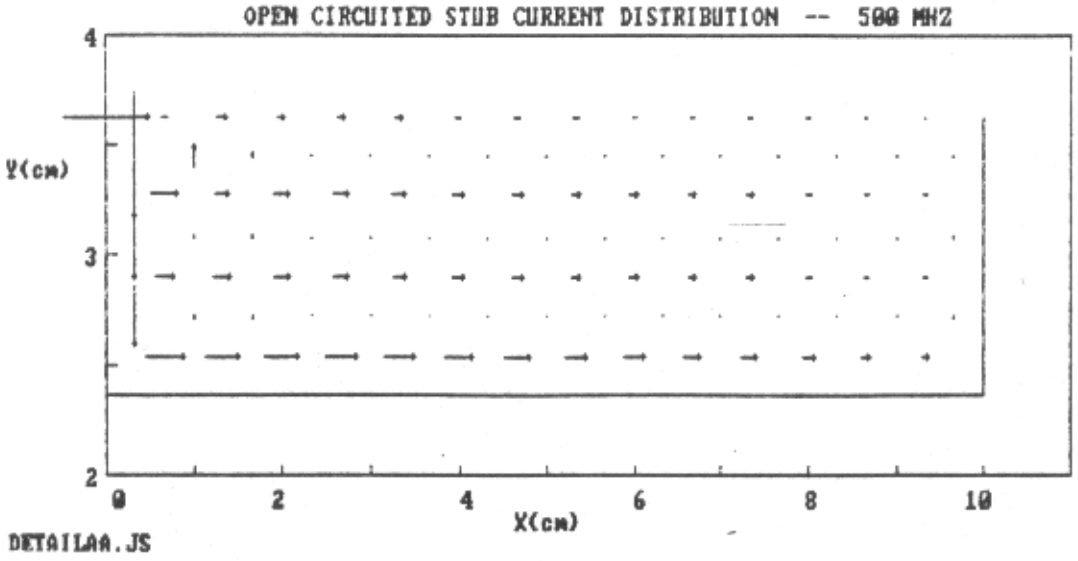
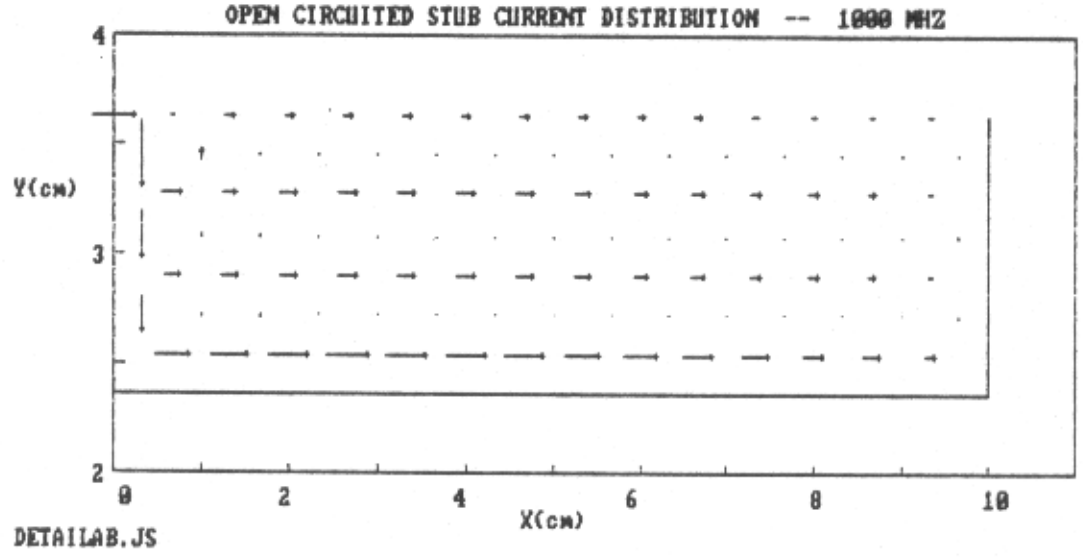


Figure 7.19. Stub current distribution (part 2).





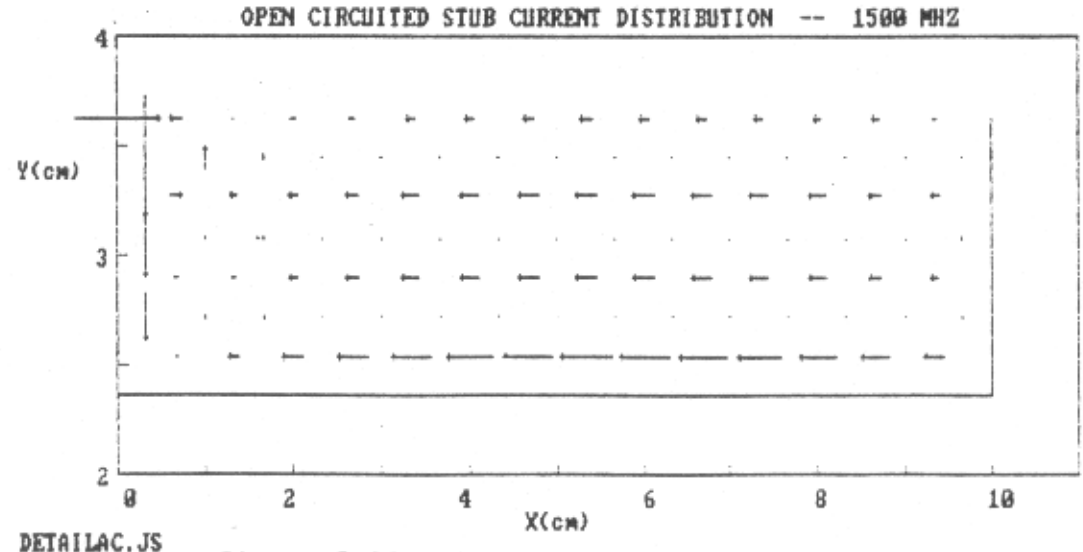
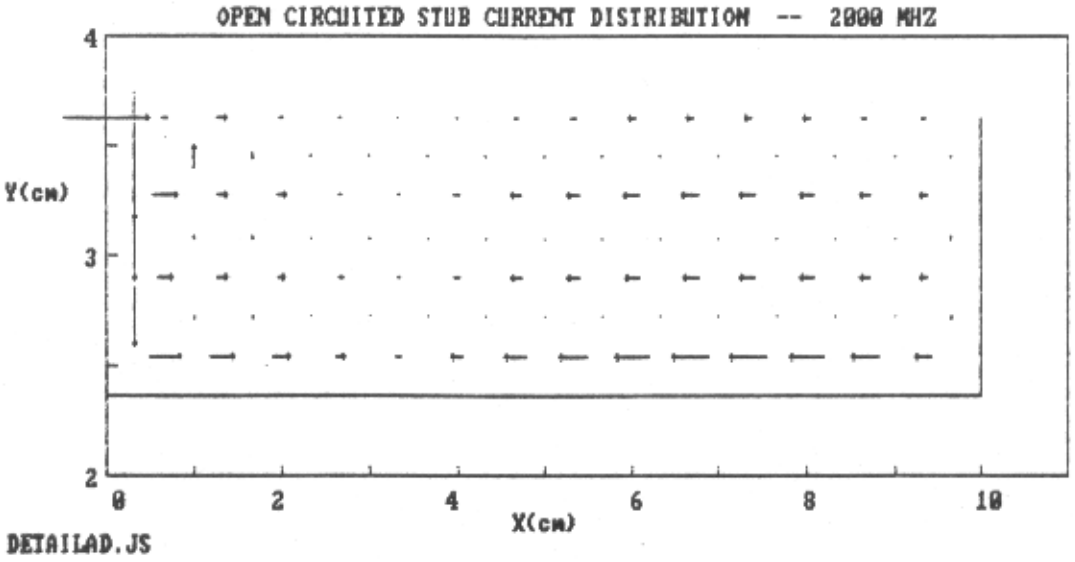
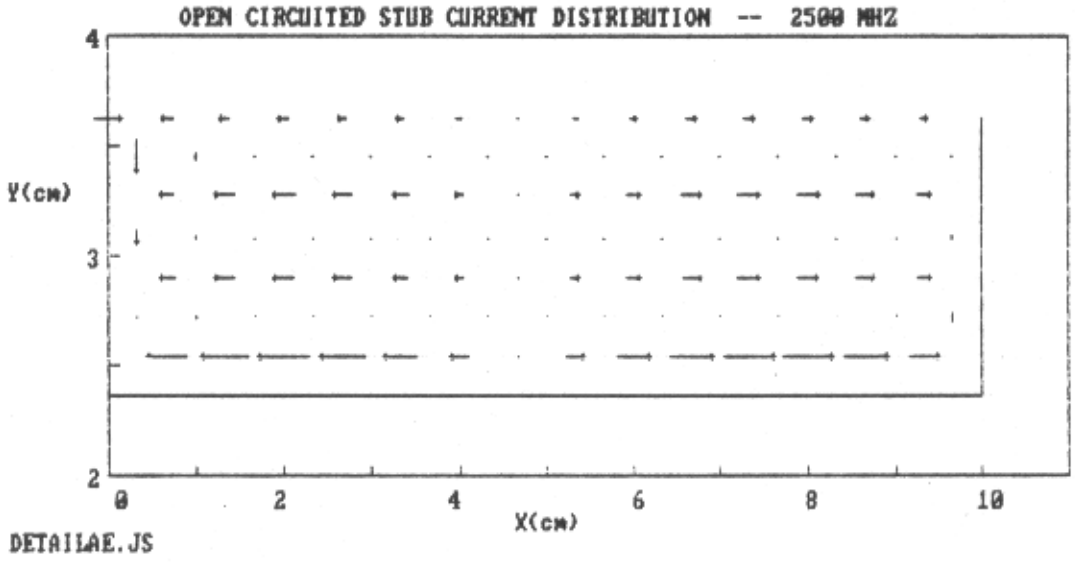


Figure 7.20. Stub current distribution (part 1).





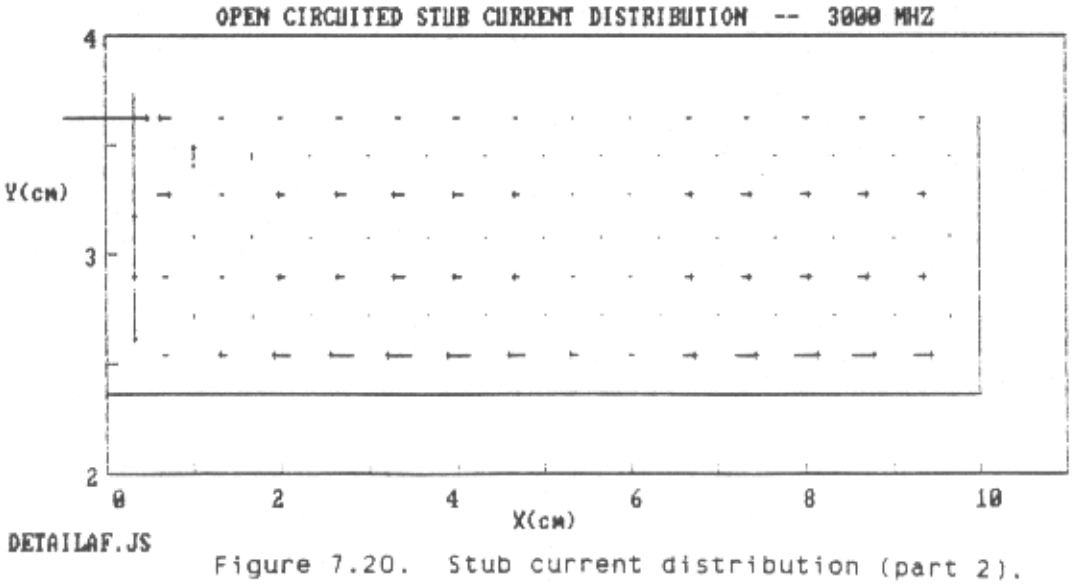


Figure 7.20. Stub current distribution (part 2).

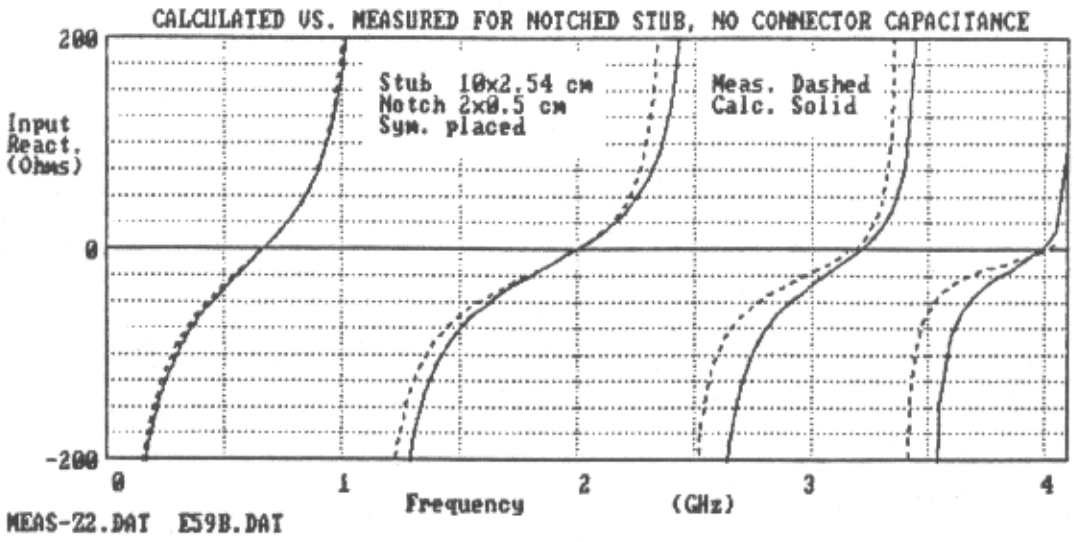


Figure 7.21. Measured versus calculated data for the notched microstrip stub before input connector fringing capacitance is included.

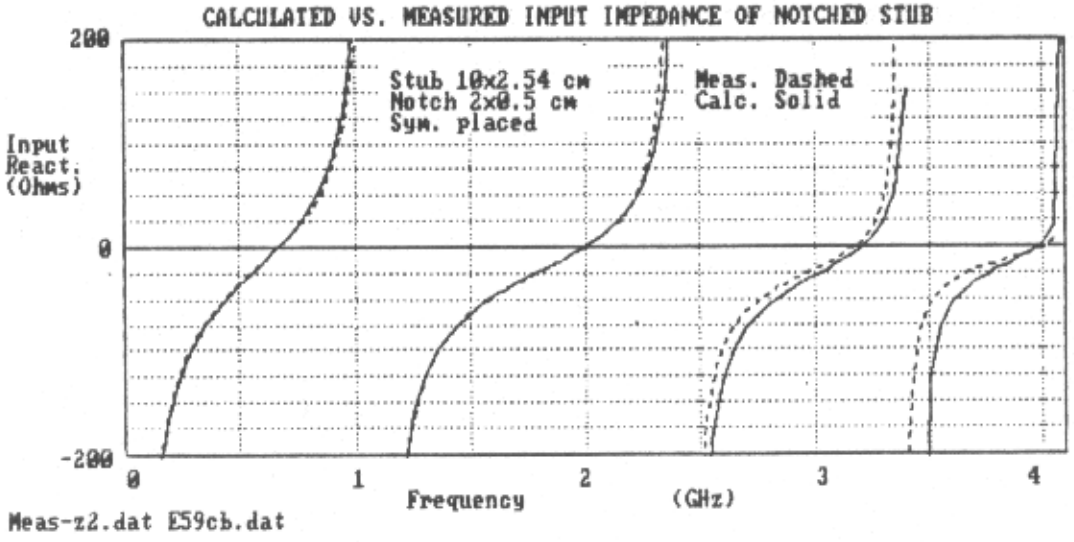


Figure 7.22. The calculated and measured notched stub data agree well once the input connector fringing capacitance is included.

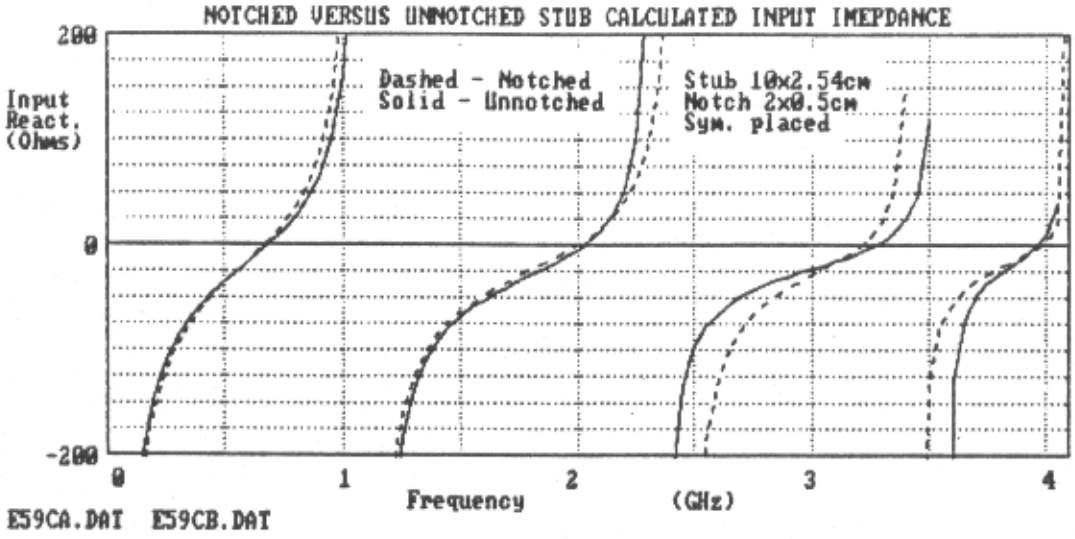
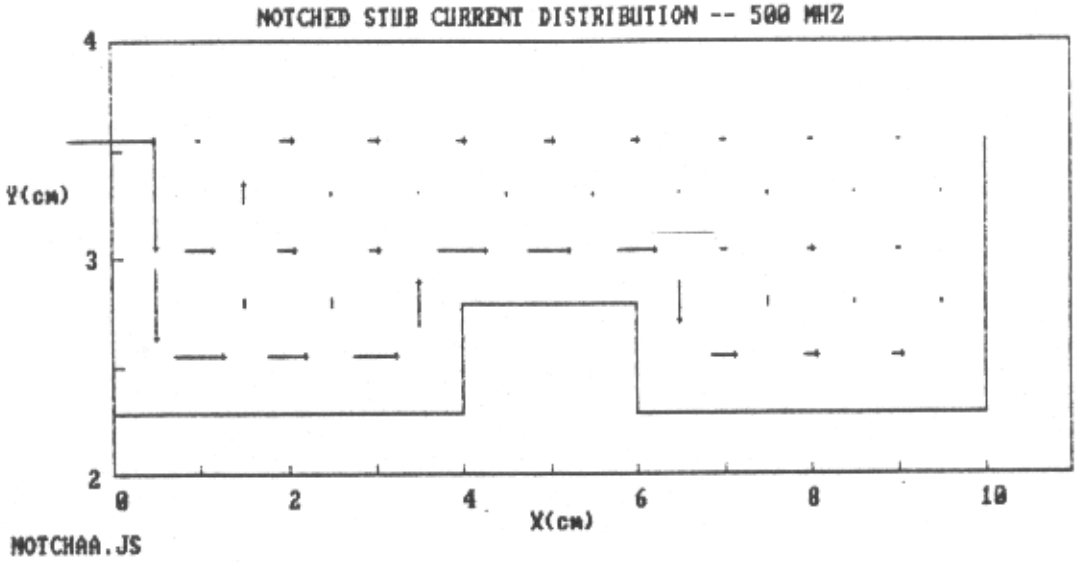
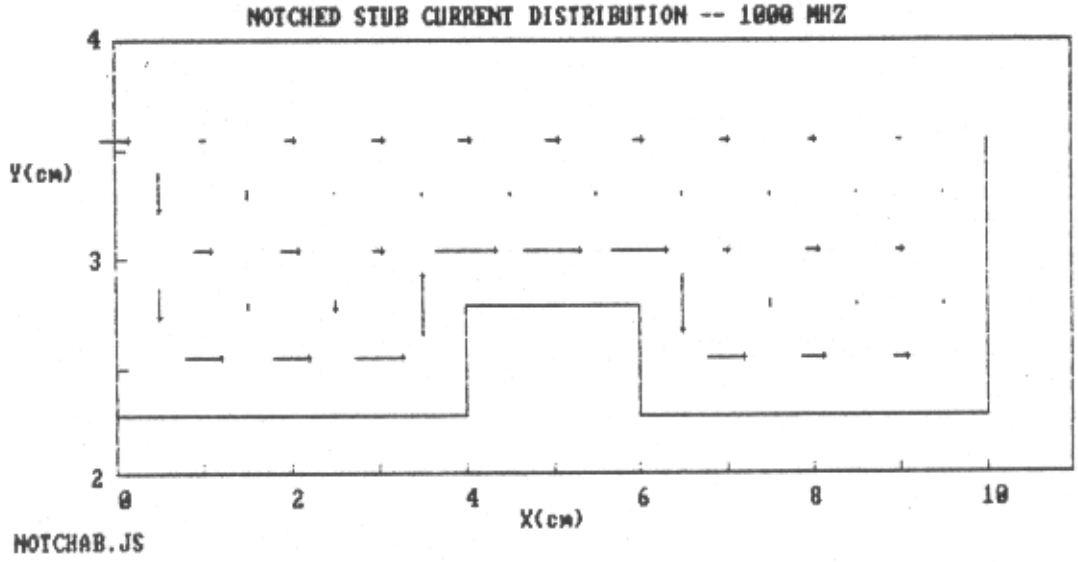


Figure 7.23. The notched and unnotched stub calculated data.





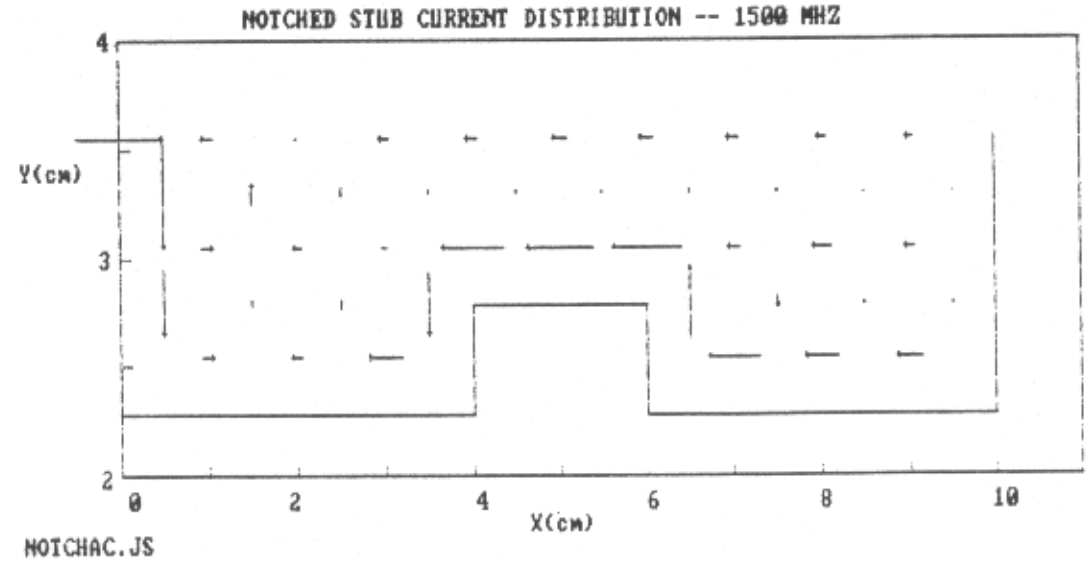
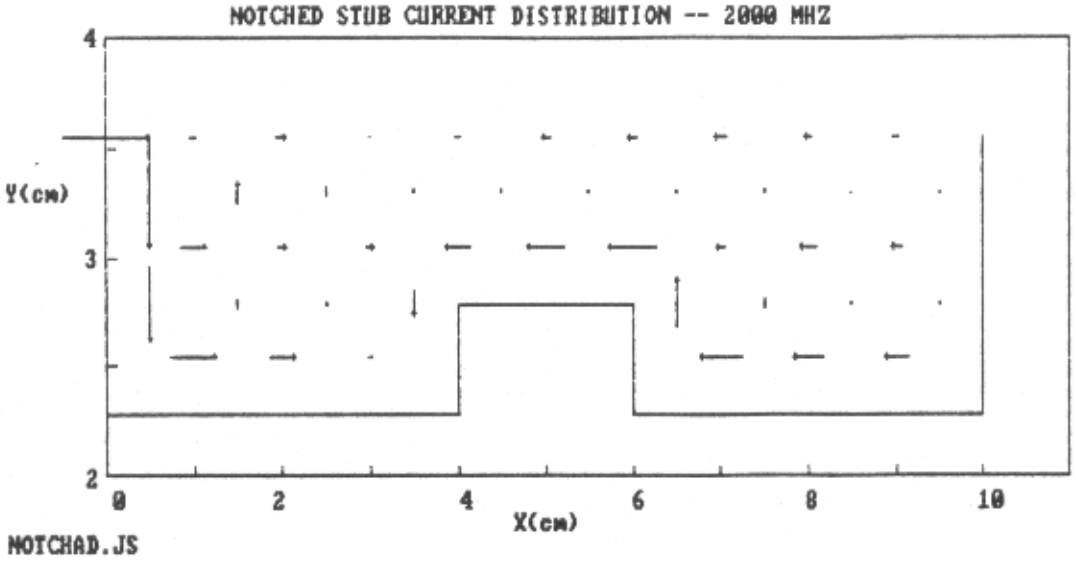
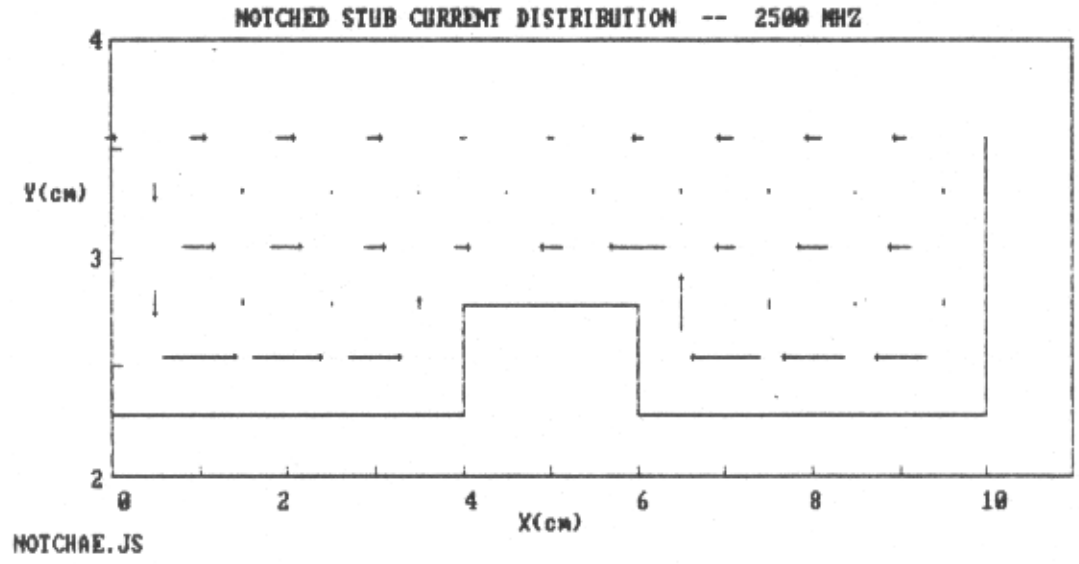


Figure 7.24. Notched stub current distribution (part 1).





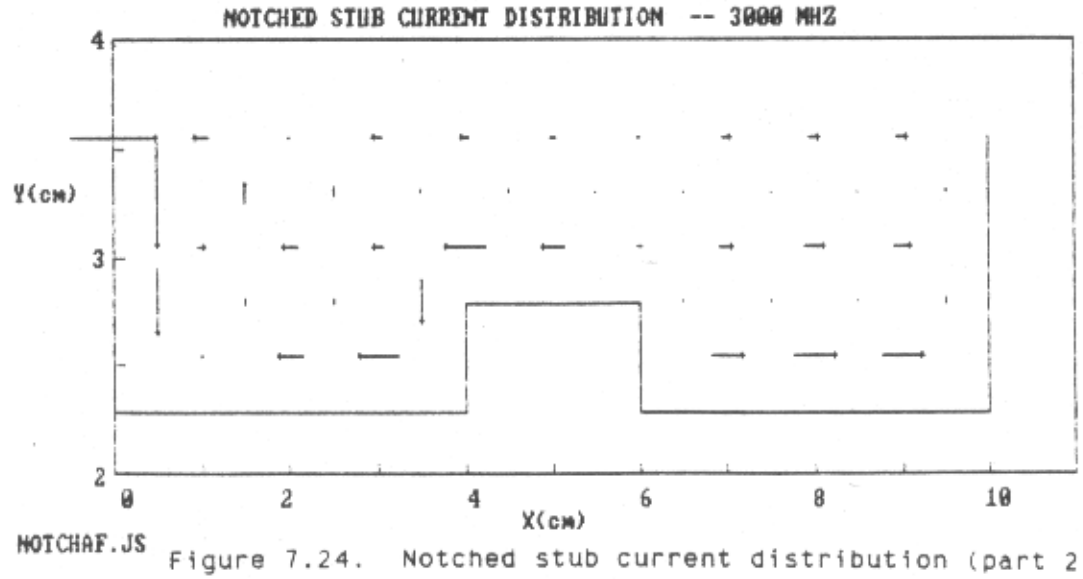


Figure 7.24. Notched stub current distribution (part 2).

Chapter 8

CONCLUSION

We have described a technique for the electromagnetic analysis of shielded microstrip circuits. The technique consists of subdividing the microstrip metalization into small rectangular subsections. There is one set of rectangles for x directed current and a second, offset, set of rectangles for y directed current. A 'roof-top' distribution of current is assumed in each rectangle. The magnitude of the current on each rectangle is adjusted so as to meet boundary conditions (zero tangential electric field) on the metalization. Once the currents are determined, the problem is solved and the N-port circuit parameters follow immediately. The electric fields due to current in each small rectangle are determined by expanding the fields as a sum of rectangular waveguide modes. This representation of the fields is closely related to the spectral domain approach.

The technique has been implemented in Pascal on an IBM-PC with an 8087 numeric coproccessor. Pascal was selected for its structure, maintainability and portability. The software makes extensive use of advanced data structures such as records, dynamic arrays, pointers and linked lists. The software is heavily vectorized so as to quickly take advantage of any array or parallel processing capability which may become available. The existing software is capable of the analysis of small circuits (10 to 20 subsections) in several minutes, while larger circuits (100 to 200 subsections) require several hours per frequency. There is still considerable room for improving the efficiency of the software. In addition, a mouse based data capture

program has been written which allows the specification of complex microstrip geometries graphically.

Two microstrip circuits have been built, an open circuited stub and a notched open circuited stub. Both stubs were 10 cm long and 2.54 cm wide. The large dimensions were chosen to ease fabrication requirements and to minimize measurement errors. Both circuits were measured on an HP8510 automated network analyzer. Comparison of computed results with the measurements shows a high degree of accuracy.

This analysis is initially expected to be useful in the creation of data bases of S-parameters of specific microstrip discontinuities. The effect of the port connecting transmission lines can be removed by the analysis of appropriate shorted microstrip stubs. This is a task which, for many discontinuities, is even within the capability of the IBM-PC.

As fast computers become more available, it is not unreasonable to consider the analysis of entire microstrip circuits. It is with this application in mind that the software was vectorized. With array processing or parallel processing machines, the analysis of circuits containing several thousand subsections becomes a real possibility.

REFERENCES

[1] K. C. Gupta, R. Garg and R. Chadha, **Computer-Aided Design of Microwave Circuits**. Dedham, MA: Artech House, 1979.

[2] K. C. Gupta, R. Garg, and I. J. Bahl, **Microstrip Lines and Slotlines**. Dedham, MA: Artech House, 1979.

[3] R. Jansen, "A novel CAD tool and concept compatible with the requirements of multilayer GaAs MMIC technology," **Microwave Theory Tech. Symposium**, St. Louis, pp. 711-714, 1985.

[4] R. H. Jansen and N. H. L. Koster, "A unified CAD basis for the frequency dependent characterization of strip, slot and coplanar MIC components," **Proc. 11th Eur. Microwave Conf.**, pp. 682-687, 1981.

[5] R. Sugarman and P. Wallich, "The limits to simulation," **Spectrum**, pp. 36-41, April 1983.

[6] R. F. Harrington, **Time-Harmonic Electromagnetic Fields**. New York: McGraw-Hill, 1961.

[7] R. F. Harrington, "Matrix methods for field problems," **Proc. IEEE**, vol. 55, pp. 136-149, 1967.

[8] R. F. Harrington, **Field Computation by Moment Methods**. New York: Macmillan, 1968. Reprinted Malabar, Fla.: Krieger Publishing, 1982.

[9] R. H. Jansen, "The spectral-domain approach for microwave integrated circuits," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-33, pp. 1043-1056, 1985.

[10] J. S. Hornsby and A. Gopinath, "Fourier analysis of a dielectric loaded waveguide with a microstrip line," **Electron. Lett.**, vol. 5, pp. 265-267, 1969.

[11] T. Itoh and R. Mittra, "Spectral-domain approach for calculating the dispersion characteristics of microstrip lines," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-21, pp. 496-499, 1973.

[12] T. Itoh, "Analysis of microstrip resonators," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-22, pp. 946-952, 1974.

[13] A. K. Sharma and B. Bhat, "Analysis of triangular microstrip resonators," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-30, pp. 2029-2031, 1982.

[14] A. K. Sharma, W. J. R. Hoefer, "Spectral Domain Analysis of a hexagonal microstrip resonator," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-30, pp. 825-828, 1982.

[15] K. Kawano, "Spectral domain analysis of a coupled microstrip resonator," **Trans. IECE Japan**, vol. E66, pp. 390-394, 1983.

[16] A. K. Sharma and B. Bhat, "Analysis of interacting rectangular microstrip resonators," **International Electrical, Electronics Conference**, pp. 38-39, 1981.

[17] A. K. Sharma and B. Bhat, "Spectral domain analysis of interacting microstrip resonant structures," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-31, pp. 681-685, 1983.

[18] I. Wolff, N. Knoppik, "Rectangular and circular microstrip disc capacitors and resonators," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-22, pp. 857-864, 1974.

[19] R. H. Jansen, "Shielded rectangular microstrip disc resonators," **Electron. Lett.**, vol. 10, pp. 299-300, 1974.

[20] R. Jansen, "Computer analysis of edge-coupled planar structures," **Electron. Lett.**, vol. 10, pp. 520-522, 1974.

[21] T. Itoh and R. Mittra, "Analysis of a microstrip disk resonator," **Arch. Elek. Ubertragung**, vol. 27, pp. 456-458, 1973.

[22] S. G. Pintzos and R. Pregla, "A simple method for computing the resonant frequencies of microstrip ring resonators," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-26, pp. 809-813, 1978.

[23] K. Kawano and H. Tomimuro, "Spectral domain analysis for slot lines," **Int. J. Hybrid Microelectron.**, vol 6, pp. 16-22, 1983.

[24] K. Kawano and H. Tomimuro, "Spectral domain analysis of slot Resonator," **Trans. IECE Japan**, vol. E65, pp. 480-484, 1982.

[25] A. Farrar and A. T. Adams, "Matrix methods for microstrip three-dimensional problems," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-20, pp. 497-504, 1972.

[26] R. H. Jansen, "Numerical computation of the resonance frequencies and current density distribution of arbitrarily shaped microstrip structures" (in German), Ph.D. dissertation, Univ. Aachen (RWTH), West Germany, 1975.

[27] R. H. Jansen, "Computer analysis of shielded microstrip structures" (in German), **Arch. Elek. Ubertragung**, Vol. 29, pp. 241-247, 1975.

[28] R. H. Jansen, "Finite-element current density representation in the numerical solution of microstrip problems," (in German) **Arch. Elek. Ubertragung**, Vol. 29, pp. 477-480, 1975.

[29] R. H. Jansen, "High-order finite element polynomials in the computer analysis of arbitrarily shaped microstrip resonators" **Arch. Elek. Ubertragung**, Vol. 30, pp. 71-79, 1976.

[30] U. Schulz, "The method of lines - a new technique for the analysis of planar microwave structures" (in German), Ph.D. disertation, Univ. of Hagen, West Germany, 1980.

[31] S. B. Worm, "Analysis of planar microwave structures of arbitrary shape" (in German), Ph.D. dissertation, Univ. of Hagen, West Germany, 1983.

[32] S. B. Worm and R. Pregla, "Hybrid-mode analysis of arbitrarily shaped planar microwave structures by the method of lines," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-32, pp. 191-196, 1984.

[33] H. Diestel and S. B. Worm, "Analysis of hybrid field problems by the method of lines with nonequidistant discretization," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-32, pp. 633-638, 1984.

[34] S. Akhtarzad and P. B. Johns, "Three-dimensional transmission-line matrix computer analysis of microstrip resonators," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-23, pp. 990-997, 1975.

[35] S. Akhtarzad, "Analysis of lossy microwave structures and microstrip resonators by the TLM method," Ph.D. dissertation, Univ. of Nottingham, England, July 1975.

[36] N. Yoshida and I. Fukai, "Transient analysis of a stripline having a corner in three-dimensional space," **IEEE Trans. Microwave Theory Tech.**, vol MTT-32, pp. 491-498, 1984.

[37] W. K. Gwarek, "Analysis of an arbitrarily-shaped planar circuit - a time-domain approach," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-33, pp. 1067-1072, 1985.

[38] V. Rizzoli and A. Lipparini, "A resonance method for the broadband characterization of general two-port microstrip discontinuities," **IEEE Trans. Microwave Theory Tech.**, vol. MTT-29, pp. 655-660, 1981.

[39] T. Vu Khac and C. T. Carson, "m=0 and n=0 mode and rectangular waveguide slot discontinuity." **Electron. Lett.**, vol. 9, pp. 431-432, 1973.

[40] J. Van Bladel, "Contribution of the ψ = constant mode to the modal expansion in a waveguide." **IEE Proc.**, vol. 128, pp. 247-251, 1981.

[41] A. W. 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, pp. 593-603, 1980.

[42] J. R. Weale, "Use of FFT in microstrip capacitance calculations," **Electron. Lett.**, Vol. 21, pg. 86, 1985.

[43] F. H. Irons, N. C. Deo, "The use of triangular basis functions for hybrid mode analysis of open microstrip lines," Proc. 1978 **IEEE Symposium on Circuits and Systems**, pp. 776-780, 1978.

BIOGRAPHY

James C. Rautio received a BSEE from Cornell in 1978, an MS in Systems Engineering from University of Pennsylvania in 1981. In 1978 he joined General Electric Space Division where he designed microwave circuits, wrote microwave design software and developed millimeter wave network analyzers. In 1982 he transferred to the GE Electronics Laboratory where he designed GaAs monolithic microwave integrated circuits. In 1984 he entered full time studies toward a doctoral degree.