

On the Application of the Secant Method
to the Spectral Iterative Approach

C. G. Christodoulou, R. J. Middelveen

University of Central Florida

Orlando, Florida 32816-0993

J. F. Kauffman

North Carolina State University

Department of Electrical and Computer Engineering

Raleigh, North Carolina 27650

ABSTRACT

A new iterative algorithm for calculating the electromagnetic scattering from planar, periodic gratings and grids was developed. Results are compared with the moment method and the Spectral Iteration (S.I.T.) method. The Secant approach is used in conjunction with the Spectral Iteration Approach to achieve convergence. It is shown that the Secant approach, which does not depend on the evaluation of numerical derivatives to achieve convergence like the contraction-corrector S.I.T. method, yields good results. Finally, suggestions for applying this method to two dimensional structures are included and discussed.

1. INTRODUCTION

Over the years, many numerical methods have evolved for solving the problem of electromagnetic scattering from periodic structures. The most popular approach, the method of moments [Harrington, R. F.], usually requires large amounts of computer memory when applied to periodic surfaces, although there are method of moment techniques that use Floquet information which improves this situation [Munk, B. A. and G. A. Burell]. Another technique, the Spectral-Iteration technique (S.I.T.) [Tsao, C. H. and R. Mittra] circumvents this memory requirement, but suffers from convergence problems.

Brand [Brand, J. C. and J. F. Kauffman] applied a corrective scheme that not only solved but accelerated the convergence problem. This method, however, depends on the evaluation of numerical derivatives to generate a series of convergent iterations. In some cases the computation of the derivatives can be so critical that the new corrective scheme fails to converge. This paper offers an improvement over Brand's method by using the Secant method instead of the contraction mapping approach.

The reasons why the contraction corrector-S.I.T. method fails to converge for two dimensional problems are discussed.

Results obtained by this method are compared with other theoretical and experimental data.

2. FORMULATION

It can be shown [Tsao, C. H. and R. Mittra] that the incident \bar{H} field can be expressed in terms of the aperture electric field \bar{E}_a as:

$$\bar{H}^{inc} = -2/j\omega\mu \sum \begin{bmatrix} \alpha_{mn} \beta_{mn} & k^2 - \alpha_{mn}^2 \\ \beta_{mn}^2 - k^2 & -\alpha_{mn} \beta_{mn} \end{bmatrix} \begin{matrix} \cong \\ \cong \\ G E_a \end{matrix} \times \exp[j(\alpha_{mn}x + \beta_{mn}y)] \quad (1)$$

where the tilde (\sim) symbol is used to denote the transformed quantity.

The parameters α_{mn} and β_{mn} are the Floquet modes and are given by:

$$\alpha_{mn} = 2\pi m/a - k \sin\theta \cos \phi \quad (2)$$

$$\beta_{mn} = 2\pi n/b \sin \Omega - 2\pi m/a \cot \Omega + k \sin\theta \sin \phi$$

The Fourier transform of Green's function is given by:

$$\bar{G} = (-j/2) (k^2 - \alpha_{mn}^2 - \beta_{mn}^2)^{-1/2} \bar{I} \quad (3)$$

Equation (1) applies only to the aperture region shown in Figure 1 and in

order to include the contribution of the \bar{H} field along the conducting strips, the current densities have to be added to equation (1) to yield:

$$\text{Tcr}'(\bar{J}) = \hat{n} \times \bar{H}^{inc} + 2/j\omega\mu \sum \begin{bmatrix} \alpha_{mn} \beta_{mn} & k^2 - \alpha_{mn}^2 \\ \beta_{mn}^2 - k^2 & -\alpha_{mn} \beta_{mn} \end{bmatrix} \begin{matrix} \cong \\ \cong \\ G E_a \end{matrix} \times \exp[j(\alpha_{mn}x + \beta_{mn}y)] \quad (4)$$

Because the current density can only be present on the strips, a truncation operator is used which is defined by:

$$\text{Tcr}\{X(\bar{r})\} = \begin{cases} X(\bar{r}) & \text{for } \bar{r} \text{ in the aperture} \\ 0 & \text{for } \bar{r} \text{ in the conducting region} \end{cases}$$

$T_{cr}\{X(\bar{r})\}$ = the opposite of T_{cr}

For gratings, $\Omega = 90^\circ$, $b = \infty$, and one ends up with the one dimensional problem.

To solve equation (4), Tsao and Mittra developed an iterative scheme which basically can be expressed in terms of the electric field only as:

$$\bar{E}_t^{(i+1)} = F^{-1} \left[\overset{\cong}{G}_0^{-1} F \left\{ (j\omega\mu/2) \left[T_{cr}\{\bar{H}_t^{inc} + (2/j\omega\mu) \right. \right. \right. \\ \left. \left. \left. F^{-1} \left[\overset{\cong}{G}_0 F T_{cr}(\bar{E}_t^i) \right] \right] \right\} - \bar{H}_t^{inc} \right] \quad (5)$$

where:

$$\overset{\cong}{G}_0 = \begin{bmatrix} \alpha_{mn} \beta_{mn} & k^2 - \alpha_{mn}^2 \\ \beta_{mn}^2 - k^2 & -\alpha_{mn} \beta_{mn} \end{bmatrix} \quad \overset{\cong}{G} \quad (6)$$

F stands for the Fourier transform and F^{-1} for its inverse. In this form, equation (5) does not converge for strip spacings of one wavelength or less. Brand imposed a corrective scheme to avoid this convergence problem. An alternative technique was developed, called the Secant Spectral Iterative method [Middelveen, R., C. G. Christodoulou, and J. F. Kauffman]. This method avoids the calculation of numerical derivatives which are required in Brands's technique. To see how the Contraction and Secant methods are used, equation (5) is expressed in an operator form as:

$$\bar{E}_t^{i+1} = L(\bar{E}^i) \quad (7)$$

where L is an operator representing the right hand side of equation (5).

For equation (5) to converge the spectral radius of L should be:

$$\rho(L) \leq 1 \quad (8)$$

One way to achieve this is to cast equation (7) as:

$$g(x^i) = x^{i+1} = L(x^i) \quad (9)$$

by letting $E = x$.

Define a new mapping $G(x^i)$ so that

$$G(x^i) = \theta x^{i+1} + (1-\theta)g(x^i) \quad (10)$$

According to the contraction mapping theory [Rus, I.A] the transformation G of a metric space X onto itself is Lipschitz continuous if there exists a ρ , independent of x and y such that

$$d(G(x), G(y)) < \rho d(x, y) \text{ for all } x, y, \in X.$$

where $d(x, y)$ is a proper metric in X. For strictly contractive mapping ρ is less than one.

One Dimensional Case

For the one dimensional case the simplest possible metric d that can be used to obtain the optimum θ is chosen as follows:

$$|G(y) - G(x_0)| < \rho |y - x_0| \text{ for } \rho < 1 \quad (11)$$

Let $y=x_0 + \delta$ then

$$|G(x_0 + \delta) - G(x_0)| < \rho |\delta| \quad \text{or} \quad \left| \frac{G(x_0 + \delta) - G(x_0)}{\delta} \right| < \rho \quad (12)$$

So the necessary and sufficient condition for contraction mapping becomes:

$$\frac{d(G(x))}{dx} < \rho \quad (13)$$

Now substitute (10) in (12) to obtain:

$$\begin{aligned} &|\theta(x_0 + \delta) + (1-\theta)g(x_0 + \delta) - \theta x_0 - (1-\theta)g(x_0)| < \rho |\delta| \quad \text{or} \\ &|\theta + (1-\theta) \frac{dg(x)}{dx}| \leq \rho \end{aligned}$$

Setting $\rho=0$ in the above equation and solving for θ yields:

$$\theta = (dg(x)/dx) / (dg(x)/dx - 1) \quad (14)$$

This value of θ is called the "contraction" factor since it will yield a convergent scheme even in those cases where the basic iterative scheme of equation (7) fails to converge. It should be noted here that in the above analysis θ is treated as a constant when in fact it is a function of x . The reason for that treatment is that θ is solved in the neighborhood of a solution (root) x_0 where the values that θ acquires are approximately equal. Therefore, θ can be assumed to be constant within that particular neighborhood.

Two Dimensional Case

In two dimensions, the basic iterative scheme of equation (10) is

given by:

$$\begin{bmatrix} x^{n+1} \\ y^{n+1} \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix} \begin{bmatrix} x^n \\ y^n \end{bmatrix} + \begin{bmatrix} (1-\theta_{11}) & -\theta_{12} \\ -\theta_{21} & (1-\theta_{22}) \end{bmatrix} \begin{bmatrix} g(x^n, y^n) \\ h(x^n, y^n) \end{bmatrix} = \begin{bmatrix} G(x^n, y^n) \\ H(x^n, y^n) \end{bmatrix} \quad (15)$$

Now it is easy to set all four partial derivatives G_x , G_y , H_x and H_y equal to zero to obtain:

$$\begin{aligned} G_x &= \theta_{11} + (1-\theta_{11}) g_x - \theta_{12} h_x = 0 \\ G_y &= \theta_{12} + (1-\theta_{11}) g_y - \theta_{12} h_y = 0 \\ H_x &= \theta_{21} - \theta_{21} g_x + (1-\theta_{22}) h_x = 0 \\ H_y &= \theta_{22} - \theta_{21} g_y + (1-\theta_{22}) h_y = 0 \end{aligned} \quad (16)$$

Solving this system of equations for θ_{11} , θ_{12} , θ_{21} , θ_{22} yields:

$$\theta_{11} = \frac{h_x g_y + g_x (1-h_y)}{h_x g_y - (1-g_x)(1-h_y)} \quad (17)$$

$$\theta_{12} = \frac{g_y}{h_x g_y - (1-g_x)(1-h_y)} \quad (18)$$

$$\theta_{21} = \frac{h_x}{h_x g_y - (1-g_x)(1-h_y)} \quad (19)$$

$$\theta_{22} = \frac{h_x g_y + (1-g_x)h_y}{h_x g_y - (1-g_x)(1-h_y)} \quad (20)$$

Again, this choice of θ 's works very well for the one dimensional problem

but it does not lead to convergence for the two dimensional wire mesh problem.

To explain why this method does not work for the two dimensional problem the theory for constructing convergent iterations for a pair of transcendental equations is invoked. According to this theory the original system of equations

$$\begin{aligned}
 x^{n+1} &= G(x^n, y^n) \\
 y^{n+1} &= H(x^n, y^n) \quad \text{can be written as:} \\
 x^{n+1} &= x^n + \alpha[g(x^n, y^n) - x^n] + \beta[h(x^n, y^n) - y^n] = G(x^n, y^n) \\
 y^{n+1} &= y^n + \gamma[g(x^n, y^n) - x^n] + \delta[h(x^n, y^n) - y^n] = H(x^n, y^n)
 \end{aligned} \tag{21}$$

Note the similarity of the above equation with equation (15). The parameters α , β , γ and δ play the same role in equation (21) as the relaxation factors θ_{11} , θ_{12} , θ_{21} and θ_{22} in equation (15). To find the root of equation (21) it is desired to determine α , β , γ and δ , by the four conditions that the first partial derivatives of G and H are zero at some point (x,y) that hopefully is near the root. Note that the unknown parameters enter linearly in the same way as θ 's do in equation (15), so the calculation of the partial derivatives G_x, G_y, H_x and H_y poses no problem. For the case of transcendental equations, it is known that this method of constructing convergent schemes works provided that the partial derivatives g_x, g_y, h_x and h_y DO NOT vary very rapidly in the neighborhood of the root (x_0, y_0) . Thus, although it is easy to produce a G and an H that are well behaved at the root (x_0, y_0) they may behave quite badly a small distance away. If this strategy is to be successful, G and H must not only have small partial derivatives in some region, but this region must also include the desired root. For the two dimensional wire mesh it

was found that the derivatives g_x, g_y, h_x and h_y vary very rapidly, especially at points close to the edges of the wire. So this fact, and the lack of knowledge of the region within which a root exists, causes this method to fail.

For the Secant Method a different approach is taken. First we define a residue vector F defined by:

$$F(E) = L(E) - E \quad (22)$$

The value of F is actually the error in the solution of equation (7). The secant method technique is applied to equation (22) as:

$$E^{i+1} = E^i - [F(E^i)(E^i - E^{i-1}) / (F(E^i) - F(E^{i-1}))] \quad (23)$$

It should be noted that the form of this equation is good only for one dimensional periodic structures (gratings) and that it has to be applied at each sampling point. Numerically, the secant method has an order of convergence of 1.62 [Traub, J. F.].

3. TWO DIMENSIONAL PROBLEM

For two dimensional problems, such as grids, the formulation of the Secant method is more complex. In this case the problem becomes the fitting of a surface at three points [Acton, F. S.]. We actually seek a common root for $E_x(x,y)$ and $E_y(x,y)$. In such a case this method is summarized as follows:

1. Given three points in (x,y) plane, fit planes to the two surfaces, E_x and E_y , and determine the point of intersection, point (x_4, y_4) .

2. Substitute one of the first three points by (x_4, y_4) .

3. Repeat steps 1 and 2 to convergence.

It should be noted that there are several degrees of freedom in choosing which point to substitute at each iteration. Geometrically, this method fails if all points are collinear due to instability. So one should choose the three points that will maximize some measure of noncolinearity [Froberg, C. E.].

Some Computational Details:

To carry out step one of the above mentioned process we start with the general plane equation:

$$a_1x + a_2y + z = m \quad (24)$$

Substituting the n^{th} point on the E_x surface we get:

$$a_1x_n + a_2y_n + E_{xn} = m \quad (25)$$

Subtracting this equation from the general plane equation on the (x, y) plane yields:

$$a_1(x - x_n) + a_2(y - y_n) = E_{xn} \quad n = 1, 2, 3 \quad (26)$$

Similarly, for the E_y surface we obtain:

$$a_3(x - x_n) + a_4(y - y_n) = E_{yn} \quad n = 1, 2, 3 \quad (27)$$

After some manipulation these equations can be decoupled and end up with

the following conditions for solving for x and y:

$$\begin{vmatrix} x-x_1 & x-x_2 & x-x_3 \\ E_{x1} & E_{x2} & E_{x3} \\ E_{y1} & E_{y2} & E_{y3} \end{vmatrix} = 0 \quad (28)$$

and

$$\begin{vmatrix} y-y_1 & y-y_2 & y-y_3 \\ E_{x1} & E_{x2} & E_{x3} \\ E_{y1} & E_{y2} & E_{y3} \end{vmatrix} = 0 \quad (29)$$

Also, to avoid having colinearity the following condition should be satisfied:

$$a_1 a_4 - a_2 a_3 \neq 0 \quad (30)$$

In general, stability in two dimensions is a much more difficult quality to handle than in one.

4. NUMERICAL RESULTS

A couple of numerical results are shown here to validate the one dimensional algorithm. Figure 2 depicts the electric field across a unit cell with a large strip size. The result is in good agreement with Brand's result. In Figure 3, the reflection coefficient predicted with this method is compared with those reported by Brand and Wait [Wait, J. R.] for various cell widths. In both examples 128 samples were used.

These results were run on a VAX 11/750 machine. Table 1 shows the number of iterations required for different numbers of sampling points.

The 32 sample points with 8 iterations require 5.25 seconds of CPU time while with 512 sampling points and 20 iterations it takes 38.58 seconds to converge. It was found that a sampling number larger than 128 will not lead to any significant change in the value of the reflection coefficient.

It should be mentioned here that this method solves for both induced currents and aperture fields whereas the FFT conjugate gradient method can only solve for one variable at a time (current or aperture field).

5. SUMMARY AND CONCLUSIONS

An alternative derivative-free method was developed to ensure the convergence of the spectral iteration approach as applied to the electromagnetic scattering from periodic gratings. This method derived from the original S.I.T. method to which a modification was made via the secant method. The two dimensional problem was formulated and its conditions for convergence stated.

6. REFERENCES

- Acton, F. S., Numerical Methods that Work. Harper and Row. New York, N.Y., 1970.
- Brand, J. C. and J. F. Kauffman, "The application of contraction theory to an iterative formulation of electromagnetic scattering", IEEE Trans. Antennas Propag., vol. AP-33, pp. 1354-1362, Dec. 1985.
- Froberg, C. E., Introduction to Numerical Analysis, Addison-Wesley Co., Reading, Mass. 1965.
- Harrington, R. F., Field Computation by Moment Methods, MacMillan, New York, 1968.
- Middelveen, R. J., G. G. Christodoulou, and J. F. Kauffman, "The secant-corrector spectral iteration method", IEEE/AP-S and URSI Symp., Philadelphia, June 1986, pp. 60. (digest).
- Munk, B. A. G. A. Burrell, "Plane Wave Expansion for Arrays of Arbitrary Oriented Piecewise Linear Elements and its Applications in Determining the Impedance of a Single Antenna in a Lossy Half-Space", IEEE Trans. Antenna Propag., Vol. AP-27, No.3, May 1979.
- Rus, I. A., "Approximation of Fixed Points of Generalized Contraction Mapping", from Topics in Numerical Analysis II, edited by John J. Miller, Academic Press, New York, 1974, pp. 157-161.
- Traub, J. F., Iterative Methods for the Solution of Equations. Chelsea Publishing Co., New York, N.Y., 1982.
- Tsao, C. H. and R. Mittra, "A spectral-iteration approach for analyzing scattering from frequency selective surfaces", IEEE Trans. on Antennas Propag., vol. AP-30, no. 2, pp. 303-308, Mar. 1982.
- Wait, R. J., "Theories of scattering from wire grid and mesh structures", in Electromagnetic Scattering, P.L.E. Uslenghi, Ed., New York: Academic Press, 1978, pp. 253-287.

Table 1.

Sampling points Vs. Iterations required

Sampling points	Iterations
32	8
64	8
128	16
256	14
512	20

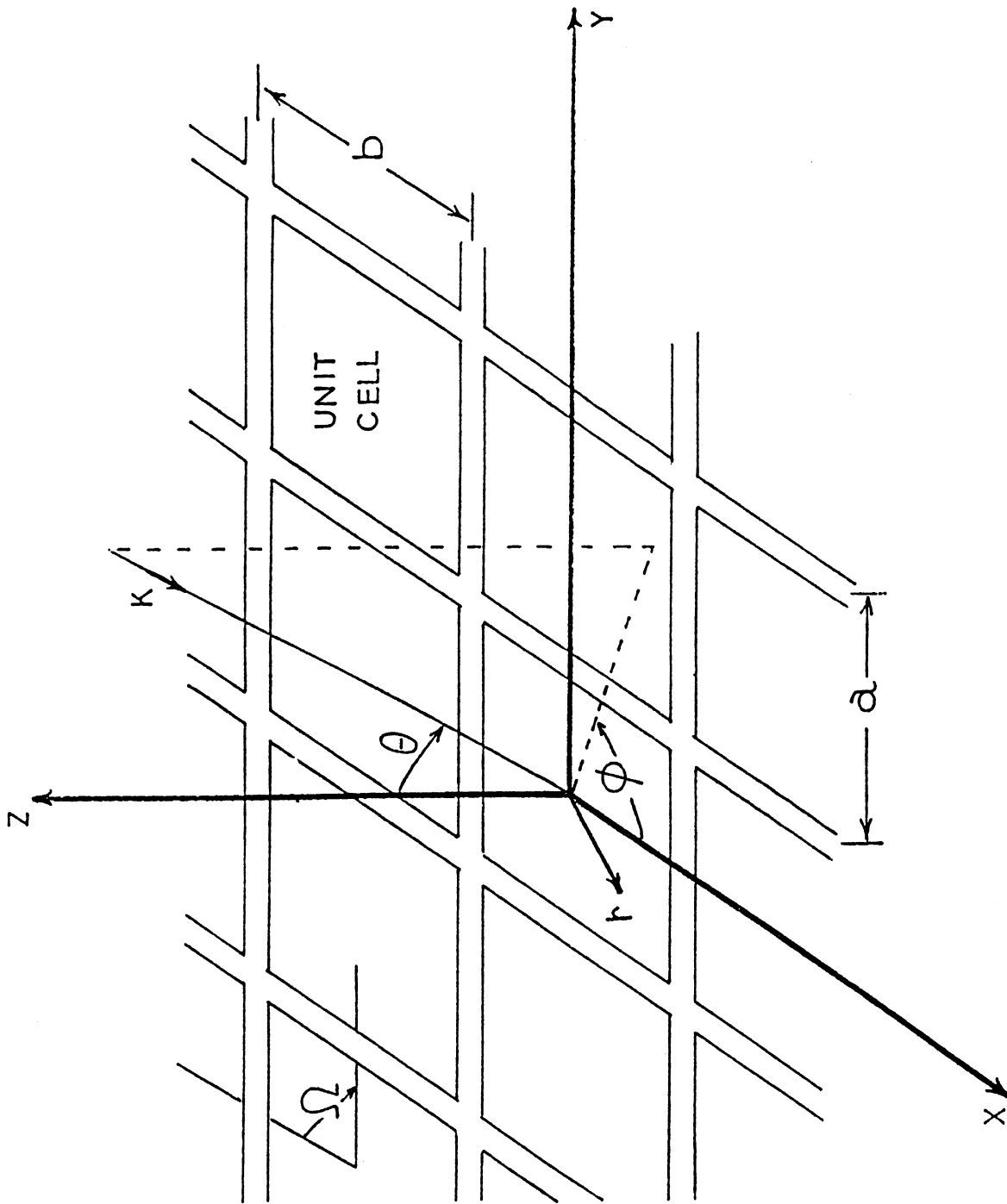


Fig. 1. Geometry of plane

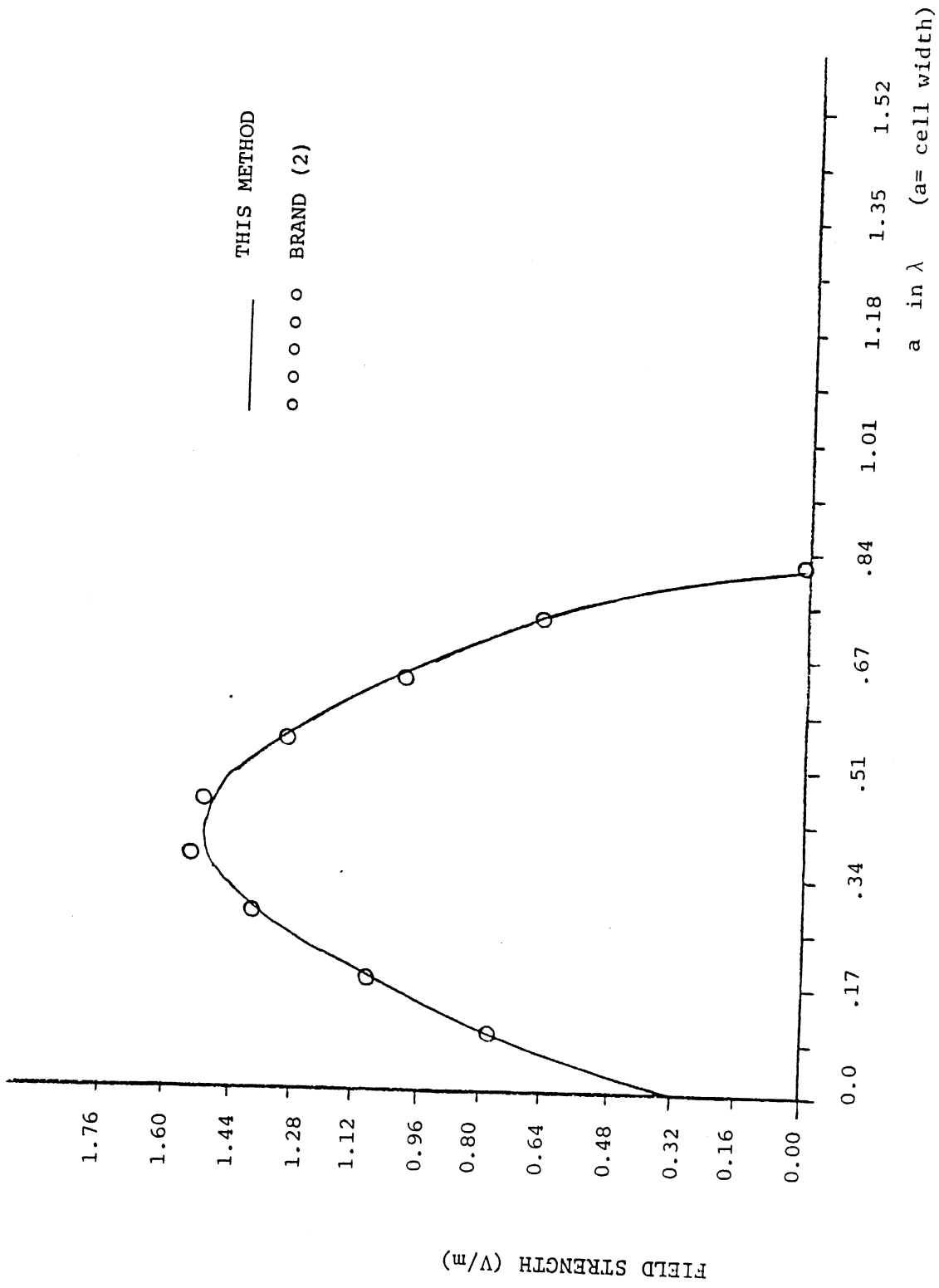


Fig. 2. Electric field across a unit cell with $a=1.4\lambda$,
 strip width=0.6 a, $b=\infty$, $\theta=0$.

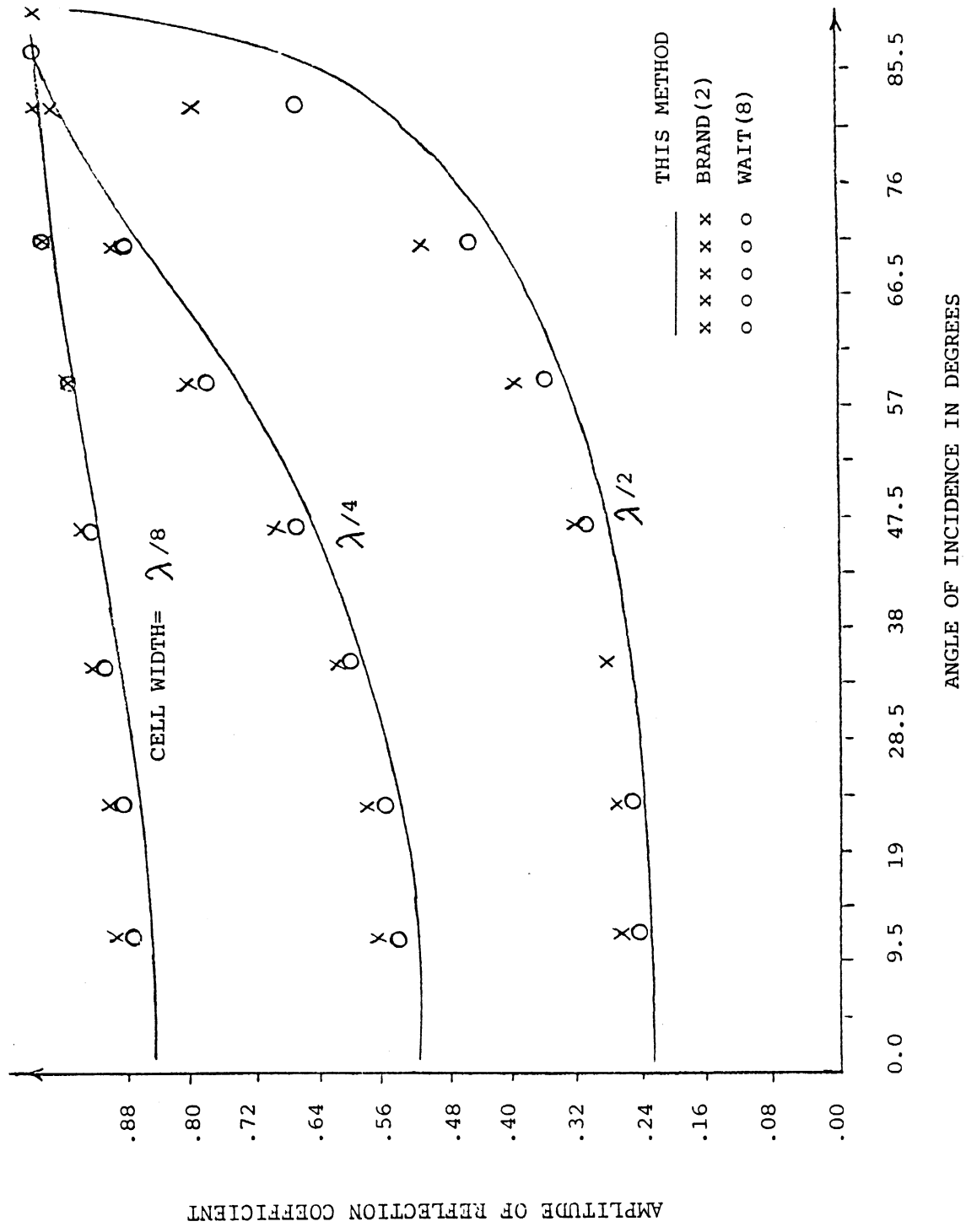


Fig. 3. Reflection coefficient for various angles of wave incidence. Width of strip = $4/600 \lambda$, $b = \infty$ and ($\phi = 0$).