TABLE I
SAMPLING MIXER SPECIFICATIONS

	Sampling Mixers	Specifications	
W-J Model No.	LO Frequency (GHz)	RF Frequency (GHz)	Conversion Loss (dB Max.)
6300–310	.200±.02	2–18	25
6300–340	1.000±.100	2–18	15
6300–370	1.000–1.500	2–18	22

are much longer than would seem appropriate according to the ideal pulse approximation. The 90 ps 0.25 pF circuit evaluated here has a monotonically decreasing output that is usable to 40 GHz (Fig. 4(a)). The same sampler modified for a 35 ps chamber delay exhibits lower conversion efficiency and a deep null at 34 GHz (Fig. 5(b)). Care must be exercised in selecting the sampling frequency of operation. For example, operating the 0.25 pF circuit at the 300 MHz sampling frequency used in [8] would be undesirable since the capacitor value is not suitable to hold the sampled signal over the duration of the sample period. VHF sampling rates require higher sampling capacitor values and faster SRD rise times to achieve Ka band operation.

Sensitivity to the pulse generator rise time for the 90 ps chamber delay sampler was investigated using four different SRD lots from three different manufacturers. The SRDs had 10%–90% transitions times in the 35 to 60 ps range. After the SRD matching network was adjusted, all four versions of the circuit were within the production performance specifications listed in Table I. In contrast, units that employ much shorter round-trip chamber delays often require a number of SRD lot evaluations before finding diodes with acceptable rise times.

V. CONCLUSION

A computer model has been developed to aid in the understanding of sampling mixer conversion loss variations at microwave frequencies. Conversion loss is predicted using the fourier transform of diode conductance waveforms. Experimental verification was given for three different values of sampling capacitors over a 2–40 GHz range. Results have been used to design a product line of integrated sampling mixers that are amenable to production manufacturing and insensitive to variations in SRD parameters.

ACKNOWLEDGEMENT

The authors would like to acknowledge W. Strifler for the fabrication of the GaAs sampling chip, and M. Dezzani, E. Pabonan, and P. Nash for their assistance in developing the integrated sampling downconverter.

REFERENCES

[1] S. E. Moore, B. E. Gilchrist, and J. Galli, "Microwave sampling effective for ultrabroadband frequency conversion," *MSN & Commun. Technol,*. pp. 113–126, Feb. 1986.
[2] S. R. Gibson, "Gallium Arsenide lowers cost and improves performance of microwave counters," *Hewlett-Packard Journal*, pp. 4–10, Feb. 1986.
[3] B. E. Gilchrist, R. D. Fildes, and J. G. Galli, "The use of sampling techniques for miniaturized microwave synthesis applications," in *IEEE Int. Microwave Symp. Dig.*, June 1982, pp. 431–433.
[4] S. A. Maas, *Microwave Mixers*. Waltham, MA: Artech House, 1986, pp. 15–28.
[5] G. E. Forsythe *et al.*, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice Hall, 1977, pp. 156–157.
[6] A. Miura *et al.*, "Full monolithic sampling head IC," in *IEEE Int. Microwave Symp. Dig.*, May 1990, pp. 845–848.
[7] K. Madani and C. S. Aitchison, "A 1 to 20 GHz sampler," *Proc. 20th European Microwave Conf.*, Sept. 1990, pp. 617–622.
[8] W. C. Whiteley, W. E. Kunz, and W. J. Anklam, "50 GHz sampler hybrid utilizing a small shockline and an internal SRD," in *IEEE Int. Microwave Symp. Dig.*, June 1991, pp. 895–898.

Parameter Extraction of Microwave Transistors using a Hybrid Gradient Descent and Tree Annealing Approach

Steven G. Skaggs, Jason Gerber, Griff Bilbro, and Michael B. Steer

***Abstract*—Tree annealing is a robust optimization scheme which can be used to find the "valleys" of an error surface. The problem of entrapment in local minima is not a factor with this type of optimization, however, it is much slower than gradient-based techniques. The method presented here attempts to take advantage of the speed of gradient-based methods and of the efficient pseudo-random searching abilities of tree annealing. The result is a technique which behaves as a directed multi-start gradient method. All minima encountered during optimization are recorded, thus providing alternatives in case of a non-physical final solution. The technique is used in the extraction of a modified Materka-Kacprzak model of a GaAs MESFET.**

I. INTRODUCTION

Circuit designers rely on models of active devices to design and simulate active microwave circuits prior to fabrication. The ability to determine accurately the parameters of these models has been limited by the lack of a sufficiently accurate equivalent circuit and a tool that extracts consistently reliable model parameters from device measurements and model predictions. Gradient-descent algorithms are most commonly used but results are seldom satisfactory unless the initial estimate of parameter values is very good. Measurement error, coupled with the large number of elements of a physically-based equivalent circuit, leads to an error function having many minima. If there are few local minima, gradient-based algorithms can be conveniently restarted from many different (often randomly chosen) initial points and the best solution taken as the approximation of the global minimum. However, the number of additional minima grows rapidly as the number of equivalent-circuit elements is increased.

Alternatives to gradient-descent-based parameter extraction have recently been proposed by Vai *et al.* [1]–[3] who used simulated annealing (SA) [4] and Bilbro *et al.* [5] who used tree annealing (TA)—a modified form of simulated annealing. These are global optimization techniques. SA differs from descent algorithms by continually accepting some proportion of random moves up the error surface. SA is more efficient than random multistart descent algorithms if the local minima tend to be shallow and if the error

Manuscript received October 28, 1991; revised August 3, 1992. This work was supported in part by the National Science Foundation through Grant No. ECS-8657836 and by Compact Software, Inc.

S. G. Skaggs, G. Bilbro and M. B. Steer are with the High Frequency Electronics Laboratory, North Carolina State University, Raleigh, NC 27695-7911.

J. Gerber is with Compact Software, Inc., 483 McLean Boulevard, Peterson, NJ 07504.

IEEE Log Number 9206289.

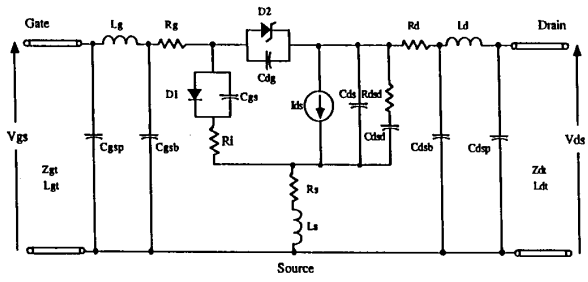

Fig. 1. Modified Materka–Kacprzak nonlinear MESFET model.

surface is not dominated by deep winding grooves or valleys, so that a feasible proportion of uphill and downhill moves can be generated. This is a consequence of SA using perturbation methods to generate new solution candidates. TA differs from SA in that information about the error surface is recorded and used to generate new solution candidates and so TA tolerates deep local minima. The robustness of the TA parameter extraction procedure is not free as it is more intensive computationally than gradient-based algorithms. However the major cost of gradient-based algorithms is the engineering time required to examine the extracted circuit, decide on its merits and to choose an alternate starting point if a satisfactory minimum has not been obtained.

We report a parameter-extraction procedure that combines tree annealing optimization with gradient descent. The result is a robust extractor which matches or exceeds the performance of either extraction technique on its own. This hybrid parameter extraction procedure requires only that the upper and lower limits of the element values be specified by the user. The only restriction enforced on the optimization process is that the number of function evaluations is limited so that the process stops when this limit is exceeded. All minima encountered in the optimization process are recorded, so that the most desirable solution can be chosen from among them.

II. Method

A modified Materka–Kacprzak of a MESFET model [6], [7] is shown in Fig. 1 where the branch R_{DSD}–C_{DSD} was added across the current source to account for the frequency sensitivity of the output conductance. The I–V and C–V expressions for the nonlinear elements, shown in Table I, and the linear elements were extracted using both I–V and S-parameter data to obtain all model parameters simultaneously. Non-unique solutions are avoided by the use of sufficient data taken in all regions of device operation. The large amount of data strengthens the ability to identify each model parameter and reduces the region in parameter space that yields an acceptable solution.

The TA optimization procedure was previously reported in [5]. The gradient-descent method used here is a variable-metric optimization process generally called a *quasi-Newton* optimizer. The mainstay of the quasi-Newton method is an approximation of the second derivatives of the error function. The TA procedure was merged with the gradient descent procedure as follows. First, an initial starting point is chosen (this can be chosen at random or by the user), and a gradient-descent optimization is performed. Optimization using tree annealing then proceeds. Whenever the error decreases below the previous best error obtained during tree annealing, a gradient-descent search is initiated. At any time, if the residual error is below a user-specified tolerance, parameter extraction is halted. In this way several local minima can be detected before the error is suitably low.

TABLE I
Device Equations for the Materka-Kacprzak MESFET Model (The Gate-Source Voltage Across C_{GS} is Defined as V_{GSI}, the Intrinsic D rain-Source Voltage A cross the Current Source I_{DS} is V_{DSI} and the Voltage Across C_{GS} and R_I is V_1.)

$$I_{DS} = I_{DSS}\left(1 + S_S\frac{V_{DSI}}{I_{DSS}}\right)\left(1 - \frac{V_{GSI}(t-T)}{V_{PO}+\gamma V_{DSI}}\right)^{(E + K_E V_{GSI}(t-T))} \cdot \tanh\left(\frac{S_L V_{DSI}}{I_{DSS}(1 - K_G V_{GSI}(t-T))}\right)$$

$$C_{GS} = \frac{C_{10}}{\sqrt{1 - K_1 V_{GSI}}}$$

$$C_{GS} = C_{10}\sqrt{5} \quad \text{if } K_1 V_{GSI} \geq 0.8$$

$$C_{DG} = \frac{C_{F0}}{\sqrt{1 - K_F(V_1 - V_{DSI})}}$$

$$C_{DG} = C_{F0}\sqrt{5} \quad \text{if } K_F(V_1 - V_{DSI}) \geq 0.8$$

$$R_I = R_{I0}(1 - K_R V_{GSI})$$

$$R_I = 0 \quad \text{if } K_R V_{GSI} \geq 1.0$$

$$I_{GS} = I_{g0}(\exp(\alpha_g V_{GSI}) - 1)$$

$$I_{DG} = I_{B0}(\exp(\alpha_b(V_{DSI} - V_1 - V_{BC})) - 1)$$

All minima encountered throughout the optimization process are recorded along with their errors. Since the quasi-Newton method is used at the outset, the hybrid method is guaranteed to perform at least as well as the quasi-Newton method alone.

III. The Error Function

The error function for this problem is directly related to the measured and calculated drain and gate terminal currents (I_D and I_G, respectively) as well as measured and calculated S -parameters. The error vector

$$\boldsymbol{E} = [\boldsymbol{M} - \boldsymbol{C}]^T \tag{1}$$

where $\boldsymbol{M}$ is a vector containing measured S-parameters and dc quantities, and $\boldsymbol{C}$ is a vector containing the corresponding calculated quantities. Because measurements can range several orders of magnitude, some weighting must be given to measurements of smaller magnitude so that the error related to them will be minimized as effectively as the error related to measurements of much larger magnitude. In this implementation, the weighting is determined automatically by the software. The weighted error is then

$$\boldsymbol{E}_W = \boldsymbol{W}[\boldsymbol{M} - \boldsymbol{C}]^T \tag{2}$$

where $\boldsymbol{E}_W$ is now the weighted error function and $\boldsymbol{W}$ is a diagonal matrix whose elements are the weights necessary to force all of the errors to roughly the same order of magnitude. All measurements of the same type (e.g., I_D, I_G, S_{11} magnitude and phase, etc.) are assigned the same weight. For the ith measurement, the weight is given by

$$w_{ii} = \frac{1}{m_i n_i} \tag{3}$$

where $m_i$is the largest error of the same type as measurement i, and n_i is the number of measurements of that type. Finally, the total error is

$$F = \boldsymbol{E_W}^T \boldsymbol{E_W} \tag{4}$$

where F is a scalar figure of merit.

IV. Results

Fifteen parameters of the modified Materka–Kacprzak MESFET model were optimized. The parameters that were not optimized were established by previous parameter extractions. dc I-V measurements

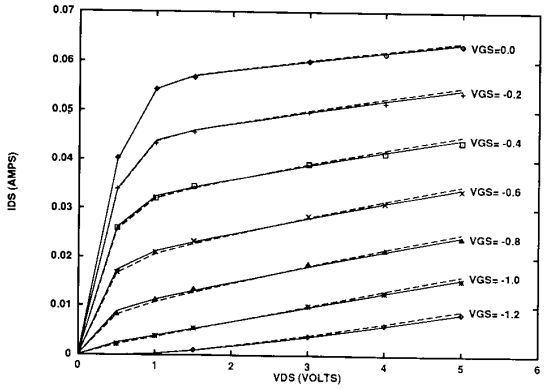

Fig. 2. Current versus voltage characteristics of the solutions given by gradient descent (dotted lines) and the hybrid method (solid lines). Points represent measured data.

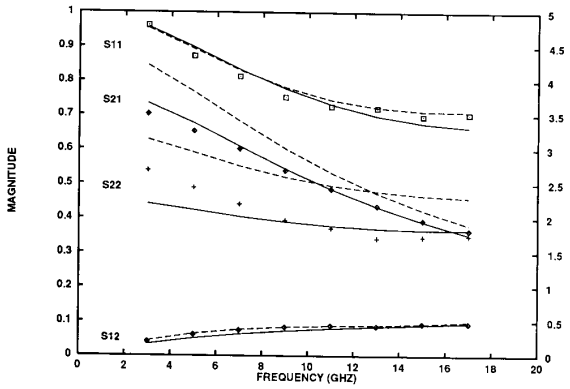

Fig. 3. Magnitude of S-parameters as predicted by gradient descent (dotted lines) and hybrid optimization (solid lines). The scale on the right corresponds to S_{21} only. Points represent measured data.

of a Toshiba 8818a GaAs MESFET were made at 49 bias points. These results are shown in Fig. 2. S -parameters were measured from 3 to 17 GHz at 2 GHz intervals at two different bias points, one with $V_{GS} = -0.6$ V and $V_{DS} = 1.5$ V and one with $V_{GS} = 0.0$ V and $V_{DS} = 1.5$ V. The resulting dc solutions for both the quasi-Newton and the hybrid optimization techniques are shown in Fig. 2, and the resulting S -parameter solutions are shown in Figs. 3 and 4. The initial parameter values and the solutions given by both optimization methods are shown in Table II. For comparison, results from a tree annealing optimization alone (with a 13 000 function evaluation limit) are also shown here. A list of some of the minima encountered in hybrid optimization along with the errors is given in Table III.

V. Discussion and Conclusion

Both the gradient-descent and hybrid optimization techniques performed well on the $I_{DS} - V_{DS}$ curves, as seen in Fig. 2. Clearly there is some discrepancy in the S-parameter data, most likely due to measurement error. However, the hybrid method clearly outperformed the gradient method on S-parameter data, particularly S_{21} and S_{22} as shown in Fig. 3 and Fig. 4. Most of the difference between the weighted error given by the two methods can be accounted for in the magnitude of S_{21}. It is interesting to note that tree annealing has proven to be particularly effective in reducing S-parameter error. One possible explanation is that the S-parameter error surface is more irregular than the error surface of the I–V data and thus contains more local minima.

The main drawback to the hybrid method is the computing time necessary for the optimization. For the example discussed here, the hybrid method required four hours of computing time on a VAXstation 3200 (rated at 3 to 5 VAX 11/780 MIPS) to extract the 15 parameter equivalent circuit. However, about 90% of the optimization

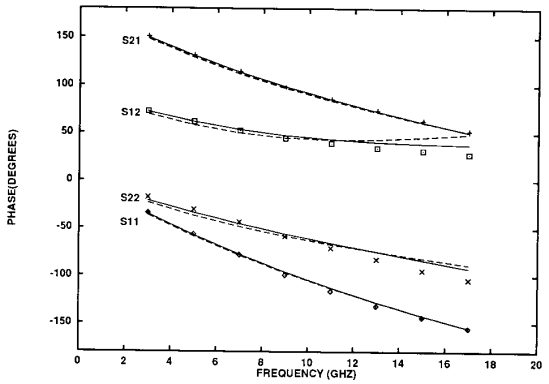

Fig. 4. Phase of S-parameters as predicted by gradient descent (dotted lines) and hybrid optimization (solid lines). Points represent measured data.

TABLE II
Extracted Parameter Values Obtained by the Hybrid Method and by Gradient Descent Alone

Parameter	Minimum Value	Nominal Value	Maximum Value	Hybrid Method	Tree Annealing	Gradient Descent
γ	-0.2	0.1539	0.2	-0.1378	0.0569	-0.1638
KE	-1.3	0.7596	1.3	-0.6884	0.1458	-0.8237
R_{10} (Ω)	0.01	5.8401	10	5.2983	9.3413	0.24
KR	-3	-2.4443	3	2.2234	0.8439	-0.035
C_{10} (pF)	0.05	0.282	10	0.2435	0.1322	0.208
$K1$	1	31	40	0.9025	33.856	0.9184
C_{F0} (pF)	0.05	0.169	1	0.0497	0.5266	0.1448
KF	1	32.914	40	1.94	20.491	12.094
L_G (nH)	0	0.166	1	0.1750	0.1824	0.1779
L_D (nH)	0	0.119	1	0.0654	0.0606	0
L_S (nH)	0	0.0404	1	0.0514	0.09354	0.0838
C_{DS} (pF)	0	0.119	2	0.1259	0.0353	0.1471
R_{DSD} (Ω)	100	4225.16	5000	211.6	142.91	4842
C_{GSB} (pF)	0	0.0705	1	0.0929	0.0616	0.1057
C_{DSB} (pF)	0	0	1	0.00006	0.017529	0
WEIGHTED ERROR, F				0.4308	1.6242	1.6669

TABLE III
Minima Encountered During Hybrid Optimization and Total Related Weighted Error

Parameter	1^{st} Minimum	2^{nd} Minimum	3^{rd} Minimum	7^{th} Minimum	11^{th} Minimum	12^{th} Minimum
γ	-0.1638	-0.1396	-0.1395	-0.1394	-0.1378	-0.1396
KE	-0.8237	-0.704	-0.7027	-0.7023	-0.6884	-0.7034
R_{10} (Ω)	0.24	8.116	6.5664	6.1128	5.2983	7.6728
KR	-0.035	1.1924	1.6054	1.7368	2.2234	1.2795
C_{10} (pF)	0.208	0.2352	0.2377	0.2389	0.2435	0.2365
$K1$	0.9184	0.8975	0.8941	0.8925	0.9025	0.8967
C_{F0} (pF)	0.1448	0.1804	0.1856	0.1835	0.0497	1372
KF	12.094	36.856	37.414	36.09	1.94	20.86
L_G (nH)	0.1779	0.1786	0.1777	0.0177	0.1750	0.1782
L_D (nH)	0	0.0583	0.0574	0.0567	0.0654	0.0589
L_S (nH)	0.0838	0.0493	0.0517	0.0524	0.0514	0.4995
C_{DS} (pF)	0.1471	0.1267	0.1281	0.1282	0.1259	0.1271
R_{DSD}(Ω)	4842	208.6	211.1	211.8	211.6	209.3
C_{GSB} (pF)	0.1057	0.1024	0.0992	0.0978	0.0929	0.1011
C_{DSB} (pF)	0	0	0	0.0003	0.0006	0
WEIGHTED ERROR, F	1.6669	0.4752	0.4634	0.4598	0.4308	0.4704

time was spent in the many gradient descent optimizations initiated in the process. In our implementation, the tree annealing optimization serves in a limited way in that the tree structure described in [5] does not become very large. However, tree annealing is still effective in determining starting points for quasi-Newton optimizations and each minimum is evaluated at approximately the cost of a single quasi-Newton optimization.

REFERENCES

[1] M.-K. Vai and S. Prasad, "Computer-aided design of monolithic MESFET distributed amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 38, no. 4, pp. 345–349, Apr. 1990.

[2] M.-K. Vai, S. Prasad, N. C. Li, and F. Kai, "Modeling of microwave semiconductor devices using simulated annealing optimization," *IEEE Trans. Electron Devices*, vol. ED-36, pp. 761–762, Apr. 1989.

[3] M.-K. Vai, J.-S. Lin, and S. Prasad, "Acceleration of simulated annealing and its application to microwave device and circuit optimization," *IEEE MTT-S Int. Microwave Symp. Dig.*, June 1991, pp. 1213–1216.

[4] S. Kirkpatrick, C. Gelatt, and M. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 13, 1983.

[5] G. L. Bilbro, M. B. Steer, R. J. Trew, C. R. Chang, and S. G. Skaggs, "Extraction of the parameters of equivalent circuits of microwave transistors using tree annealing," *IEEE Trans. Microwave Theory Tech.*, vol. 38, 1990, pp. 1711–1718.

[6] A. Materka and T. Kacprzak, "Computer calculation of large-signal GaAs amplifier characteristics," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, Feb. 1985, pp. 129–135.

[7] J. Gerber and R. Gilmore, "Parameter extraction of MESFETS for use in oscillator design," in *1989 WESCON/89 Conf. Rec.*, pp. 83–87.

Quasi-Static Analysis of Shielded Microstripline by a Modified Boundary Element Method

T. N. Chang and Y. T. Lin

***Abstract*—This paper presents a modified boundary element method for analyzing the shielded microstrip-like structure. The boundary integral equations are derived via the Green's second identity with the adjoint fields chosen to satisfy the boundary conditions along the outside shielding conductor. Numerically, these result in a considerably reduced matrix size compared to that using free space Green's functions as the adjoint fields. The computation time for off-diagonal element of the matrix can be decreased by taking the Maclaurin series expansion forms of the infinite sums. Results for microstrip line are found in good agreement with those in the literature.**

I. INTRODUCTION

The boundary element method has previously been applied to the analysis of lines in a shielding box with and without a dielectric substrate [1]–[2]. An analysis of microstrip line with finite thickness was presented in [3]. In this method, the wave equation is converted to an integral over the boundary of the region of interest by way of Green's second identity. Although the free space Green's function is chosen as the adjoint field in [1]–[3], it is by no means the only choice. In [4], the Green's function for the classical image problem was employed. For a shielded structure, a unified approach to determine the required Green's function was suggested [5].

One disadvantage of the choice of free space Green's function is that the matrix size formulated by discretizing the boundary integral equation is generally large since the whole shielding conductor should be discretized [3]. In this paper, a modified boundary element method is presented wherein the Green's function is forced to satisfy the boundary conditions along the shielding boundary. A boundary integral results which is performed merely along the line where the strip is located. Therefore, the required memory size will largely be reduced. This newly adopted Green's function involves a slowly-convergent infinite series. However, the computation time is reduced by application of the geometric-series method [6].

Manuscript received June 5, 1992; revised October 30, 1992. This work was supported by the Tatung Company, Taipei, Taiwan Republic of China.

The authors are with Tantung Institute of Technology, 40 Chungshan North Road, 3rd Sec., Taipei, Taiwan 10451, Republic of China.

IEEE Log Number 9206288.

II. FORMULATION

The cross-section of a microstripline shielded by a perfect conductor is considered. The subdomain S_i with contour $\Gamma_i\,(i = 1, 2)$ is homogeneously filled with a loss-free dielectric medium. Inside each region S_i, Laplace's equation

$$\nabla^2\phi_i = 0 \tag{1}$$

holds, where ϕ_i denotes the electrostatic potential. Green's second identity over S_i can be expressed as

$$\int\int[\varphi_i(\nabla^2\phi_i) + \phi_i(\nabla^2\varphi_i)]ds_i = \int[\varphi_i(\partial\phi_i/\partial n_i) - \phi_i(\partial\varphi_i/\partial n_i)]\,d\Gamma_i \tag{2}$$

where φ_i is a suitable adjoint field and $\partial/\partial n_i$ the derivative in the positive normal direction. The free space Green's function is chosen for φ_i in [1]–[3] for boundary element formulation. The disadvantage of this choice is that the boundary integral equation must be performed on both contours $\Gamma_1\,(= ABCDEFA)$ and $\Gamma_2\,(= ABCGHFA)$. There results in a large memory size if the boundary scale is considerably large. In this paper, each Γ_i is divided into two parts Γ_i^g and Γ_i^r, where the superscript g and r stand for the ground and the remainder parts respectively. For example, we have $\Gamma_1^g = CDEF$ and $\Gamma_1^r = FABC$. The suitable adjoint field φ_i now can be chosen to satisfy the required boundary condition on the ground plane. In the present case, the Green's function for a rectangular trough region [6] was chosen as the candidate. With the coordinates shown in Fig. 2, two Green's functions for the homogeneous rectangular trough regions are needed. They satisfy the following differential equations:

$$[\partial^2\varphi_i/\partial x^2] + [\partial^2\varphi_i/\partial y^2] = -(1/\epsilon_i)\delta(x - x_i)\delta(y - y_i) \quad i = 1, 2 \tag{3a}$$

where δ is the Dirac Delta function, $\epsilon_1 = \epsilon_0\epsilon_r, \epsilon_2 = \epsilon_0$, and (x_i, y_i) is the source point in the ith region.

To facilitate understanding, φ_i is given by

$$\varphi_i(x, y/x_i, y_i) = (2/\pi\epsilon_i)\sum_{n=1}^{\infty}(1/n)\sin(n\pi x_i/b)\sin(n\pi x/b) \cdot \sinh[(n\pi/b)f_i]\exp[(n\pi/b)g_i] \tag{3b}$$

where $(f_1, g_1) = (y_1, -y)$ for $y > y_1$; $(f_1, g_1) = (y, -y_1)$ for $y \le y_1$; $(f_2, g_2) = (h + h' - y, y_2 - h - h')$ for $y > y_2$; and $(f_2, g_2) = (y_2 - h - h', y - h - h')$ for $y \le y_2$).