MICROWAVE MODELING USING NEURAL NETWORKS AND IIR FILTERS IN TIME AND FREQUENCY DOMAINS

A. Gati*, M. F. Wong (Senior IEEE)** , G. Alquie (Member IEEE)* ,
V. Fouad Hanna (Fellow IEEE)*
* Laboratoire instruments et systèmes - Case 252 - Université Pierre et marie Curie
4, place jussieu - 75252 Paris cedex 05
** CNET - France Télécom - Département DMR/IIM
38-40, Rue du Général Leclerc - 92794 Issy Les Moulineaux cedex 09
Fax : 00 33 1 44 27 75 09

Abstract¾A method combining Infinite Impulse Response (IIR) filters and Neural Networks (NN) is proposed for precise electromagnetic (EM) of complex structures true the use of efficient optimization and parameterization techniques. This method is based on the characterization of the time response as a transfer function using a digital filter. The use of the neural network allows then the modeling of the geometric variation of the studied structure.

The method is then also expanded in the frequency domain for studying the variation of the scattering parameters for the studied structures as a function of frequency. This is achieved by building a spectral transfer function. The main difference between time and spectral analyses is discussed and it is shown that it is more advantageous to interpolate the transfer function's poles and zeros rather than the polynomial coefficients The validity of our proposed technique is demonstrated on a microstrip step in width discontinuity and a microstrip filter, both analyzed in the time domain and also on a chamfered bend analyzed in the frequency domain.

Key words : Em modeling, Optimization, Neural networks, Numerical filters

I. INTRODUCTION

EM-field Optimization requires to perform several simulations using an electromagnetic solver based on rigorous numerical methods like Finite Elements (FE) [1], or Finite Difference Time Domain (FDTD) [2]. However, each EM simulation is highly time consuming and optimizing using traditional techniques becomes a heavy task.

Conventional Neural Networks, especially Multilayer Perceptron (MLP) have been introduced for optimal modeling and simulation of microwave structures [3]. This method proved its efficiency for reducing the CPU time consumed for optimization by mapping a relationship between input and output (I/O) parameters [4-5]. However, the major drawback of NN is that it costs many computer calculating hours to achieve a good accuracy due to the high complexity of time signals.

The Infinite Impulse Response filters (IIR) which are largely used in signal processing, have been successfully introduced for field simulation in the time domain [6]. The filter approach is well adapted to time domain algorithms because of its iterative construction. However, the IIR filter is not multidimensional and thus, cannot take in consideration the geometric variation of structures.

In this paper we propose to combine neural networks and IIR filters for determining the time behavior of the studied structures. The parameters are separated in two categories. The predefined time domain responses in one side and the geometric structure parameters in the other side. This method is adopted after observing that the variation of the geometric parameters does not hardly affect the response of the structure. Thus, time is nearly the only responsible of the important variations of the system outputs. Hence, we use an IIR-filter for studying the time response and the NN-MLP for modeling geometric variation of the structure dimensions. In this manner, we reduce the size of the samples and increase the accuracy of the interpolated points.

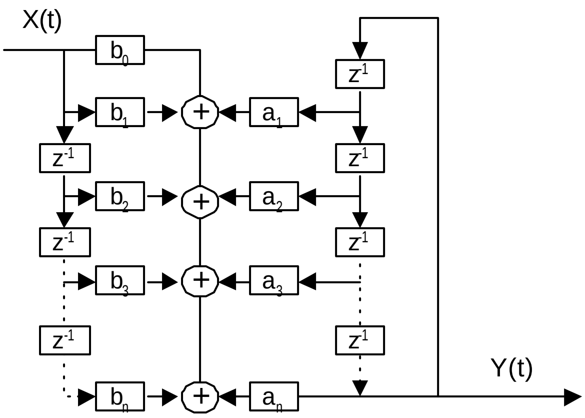
In the frequency domain, the transfer function is built from a data set representing the spectral response of the studied system. Then, by applying a simple algorithm we extract the poles and zeros of the transfer function. The poles and zeros are then learned using neural networks. Hence, for any value of the geometrical parameters, the NN gives the poles and zeros of the transfer function. And thus, the complete response is known.

The paper is organized as follows. Section II starts describing the IIR-filter approach in time domain. Section III contains a description of MLP-NN. Then in section IV, we present the proposed method and especially the connection between the two techniques. Section V describes the domain extension of the method. Results are described in section VI and general conclusion in section VII.

II. IIR FILTERS

The responses provided by the FDTD method can be interpreted as a high order digital filter with input and output signals (x(t), y(t)) respectively. The filter is based on a finite difference equation which can basically be written as :

$$\hat{y}(n) = \sum_{i=1}^{N} a_i y(n-i) + \sum_{j=0}^{M} b_j x(n-j)\ . \qquad (1)$$



This equation represents the transfer function of the IIR filter described in Fig.1, where a_i and b_i are the feedback and feedforward coefficients respectively. y(n-j) are the output samples and x(n-i) are the input samples.

Figure1 IIR filter structure

Equation (1) can be written in a matrix form as:

$$\hat{y}(n) = \left[Y_{n-1}^T X_n^T\right]\begin{bmatrix} a \\ b \end{bmatrix} = z_n^T \boldsymbol{q} \qquad (2)$$

where :

- ♦ $Y_{n-1}^T = [y(n-1)\ y(n-2) \ldots y(n-N)]$ is the output vector.
- ♦ $X_n^T = [x(n)\ x(n-1) \ldots x(n-M)]$ is the input vector.
- ♦ $\boldsymbol{q}^T = [a_1\, a_2\, \ldots a_N\ b_0\, b_1 \ldots b_M]$ is the N+M+1 filter coefficient vector.

The error between the true output y(n) and estimate filter output is :

$$e(n) = y(n) - \hat{y}(n) = y(n) - z_n^T \boldsymbol{q} \qquad (3)$$

The filter coefficients are obtained by forming a least square solution of an over determined set of p+1 equations (note : p>M+N). The p+1 equations are written from (2) as :

$$\begin{bmatrix} \hat{y}(n) \\ \hat{y}(n-1) \\ \vdots \\ \hat{y}(n-p) \end{bmatrix} = \begin{bmatrix} Y_{n-1}^T & X_n^T \\ Y_{n-2}^T & X_{n-1}^T \\ \vdots & \vdots \\ Y_{n-p-1}^T & X_{n-p}^T \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} z_n^T \\ z_{n-1}^T \\ \vdots \\ z_{n-p}^T \end{bmatrix} \boldsymbol{q}$$

Now, the corresponding error vector is E = Y-Z θ where Z is a (p+1) × (M+N+1) matrix of input and output data samples. The least square solutions to minimize E = $e^T e$ error yields :

$$\begin{aligned} (Z^T Z)\boldsymbol{q} &= Z^T Y \\ R\boldsymbol{q} &= r \end{aligned} \qquad (4)$$

where $R = Z^T Z$ is the data autocorrelation matrix and $r = Z^T Y$ is crosscorrelation vector. From (4) we can write this matrices as :

$$R_n = Z^T Z = R_{n-1} + z_n z_n^T \qquad (5)$$

$$r_n = \sum_{i=n-p}^{n} z_i y(i) = r_{n-1} + z_n y(n) \qquad (6)$$

The filter coefficients are calculated using (4), thus, we have :

$$q_n = R_n^{-1} r_n \qquad (7)$$

Using (6), the equation (7) becomes:

$$q_n = R_n^{-1}(r_{n-1} + z_n y(n)) \qquad (8)$$

The crosscorrelation matrix is computed using (5) as follows:

$$r_{n-1} = R_{n-1}^{-1} q_{n-1} \qquad (9)$$

Then by introducing (9) in (8) we have:

$$q_n = R_n^{-1}(R_{n-1} q_{n-1} + z_n y(n)) \qquad (10)$$

And finally, we use (5) to compute the autocorrelation matrix in (10) :

$$q_n = q_{n-1} + R_n^{-1} z_n (y(n) - z_n^T q_{n-1}) \qquad (11)$$

This last equation constitutes the adaptive formula for computing the filter coefficients. To perform this algorithm we need to evaluate the inverse autocorrelation matrix used in (11). For this purpose we use the matrix inversion lemma as follows :

$$R_{n+1}^{T\ -1} = R_n^{-1} - \frac{R_n^{-1} z_{n+1} z_{n+1}^T R_n^{-1}}{1 + Z_{n+1}^T R_n^{T^{-1}} Z_{n+1}} \qquad (12)$$

This lemma requires the knowledge of the initial approximation of the inverse autocorrelation matrix noted R_0^{-1} .This matrix is initialized by :

$$R_0 = \begin{bmatrix} s_y I & 0 \\ 0 & s_x I \end{bmatrix} \qquad (13)$$

where σ_x and σ_y are the autovariance of x(t) and y(t) respectively. I is the identity matrix of appropriate dimension.

Note that the output signal is delayed from the input one due to the propagation phenomena. This delay time (Δt) must be taken in consideration in the adaptive algorithm to reduce the filter order. Hence, the algorithm would be applied to the signal defined as : $\hat{x}$(t)=x(t-Δt).

It is important to note that once the coefficients converge, no more simulation is needed. The output signal at time t (ie : y(t)) is extrapolated from the input samples x(t-i), $0 \le i \le M$ and the earliest output samples y(t-j), $0 \le j \le N$.

III. NEURAL NETWORKS

The standard MLP neural network is composed of 3 layers: Input layer, hidden layer and output layer (Fig.2). A set of (x,y) data called the training data, is generated from EM simulations. The neural model learns the (x,y) relationship from the training data.

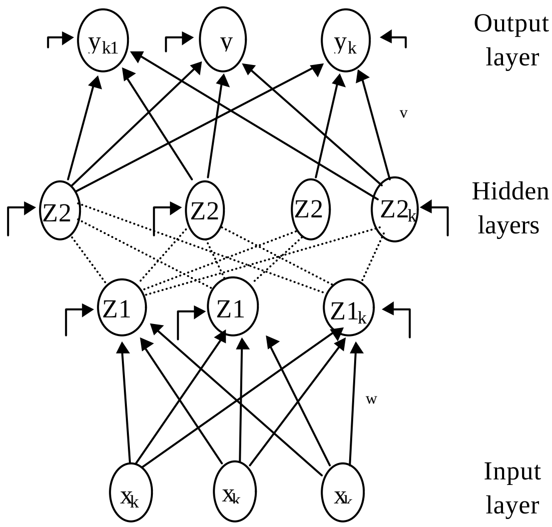


Figure.2. MLP-NN with 2 hidden layers

The training set is especially used to compute the weights and biases of the neural network. The computational algorithm called « back propagation » is then used to adjust the neural parameters[7]. Another set of data called the testing set is used to test the ability of the network for generalizing out of the training set. Training and testing sets are used to train and stop the neural learning.

The input parameters are generally the geometric dimensions of the studied system and the frequency or time. The output parameters could be the scattering parameters, the eigenvalues or eigenvectors. For our purpose, we use a conventional MLP with 2 hidden layers and an adaptive learning rate.

IV. COUPLING METHOD

As it was previously discussed, coupling the MLP-NN and the IIR filter is made in the learning set. The inputs of the neural network
are the geometric parameters and the outputs are IIR coefficients (Fig 3)

For any value inside the geometric parameters interval, the coefficients can be determined using the neural network. The filter have approximately the same behavior as the simulated response.

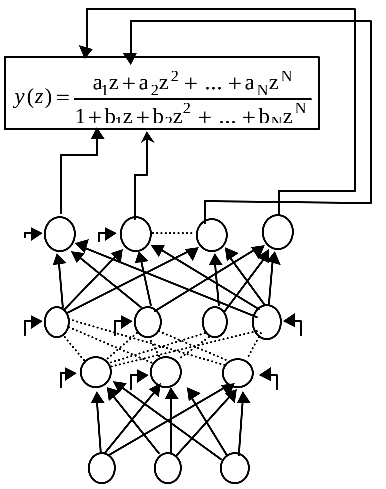


Figure. 3. Combined MLP-NN and IIR filter

V. EXPANSION OF THE METHOD TO THE FREQUENCY DOMAIN

The IIR-filter approach is still available in the spectral analysis when the filter is causal and stable. The frequency response is computed by either the Fourrier transform or the inverse z transform z = e^{jw} of the transfer function. In this case the transfer function can be written as:

$$H(w) = \frac{b_0 + b_1 e^{-jw} + ...+ b_M e^{-jMw}}{1 + a_1 e^{-jw} + a_2 e^{-2jw} + ...+ a_N e^{-jNw}} \qquad (14)$$

It is to be noted that the expression of the transfer function is close to the formulation of the Cauchy method [8] applied in modeling the scattering parameters via model reduction techniques

VI. RESULTS (TEST CASES)

VI.1. Step in width microstrip discontinuity

The first test case is a microstrip step in width discontinuity studied in the time domain using the FDTD method.

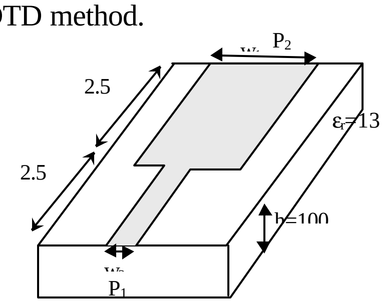


Figure 4. Microstrip step in width discontinuity

The input parameters of the neural network are the widths w_1 and w_2 as shown in Fig 4. The output parameter of the IIR filter is the voltage time domain signal between the strip and the ground at the output port P2. This structure is studied for w_1 varying from 75 µm to 100 µm and w_2 varying from 150 µm to 200 µm. The structure is excited using a Gaussian input signal at the plane P1.

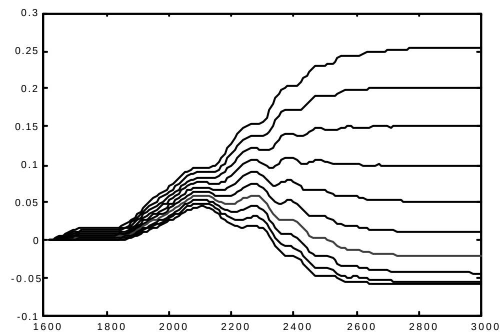


Figure 5: Convergence of the IIR filter coefficients

The step in width is designed on a GaAs (εr= 13) substrate.

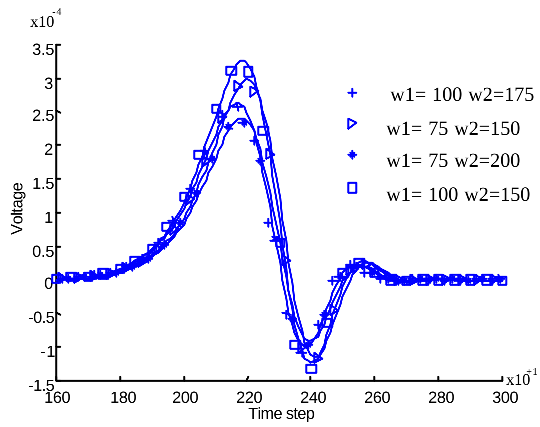


Figure 6. NN-IIR interpolated response compared to that from an EM commercial simulations for different values of w_1 and w_2.

A 30x30 IIR filter was needed to have a good approximation. Fig 5 shows the first 10^{th} filter coefficients convergence and Fig 6 shows results that are tested choosing different values for w_1 and w_2.

This test was performed for a geometry of a step in width discontinuity which was not furnished in the learning set of the neural network. Figure 6 shows the good accuracy between results using directly the FDTD EM simulator and those using NN_IIR model.

V I.2. Filter test

The purpose of this test is to model a microstrip filter (Fig 7) using the NN-IIR method. For this case, we use a 70x70 digital filter to approach the output voltage.

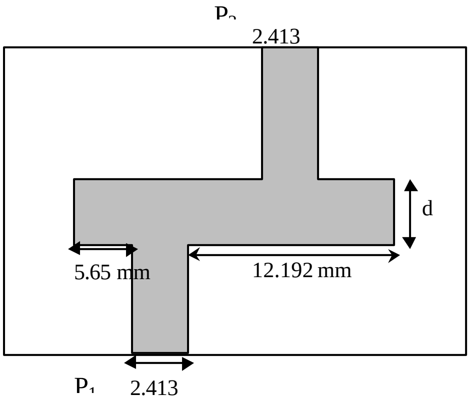


Figure 7: Microstrip filter design

This structure is studied using the FDTD method. The input parameter of the neural network is the width d (see Fig 7). The output parameter of the IIR filter is the voltage time domain signal between the strip and the ground at port P2.

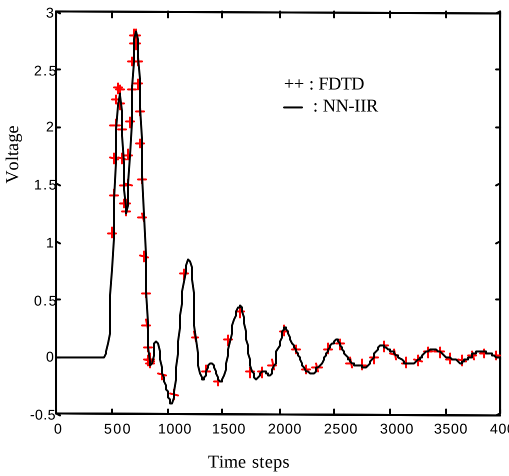


Figure 8. NN-IIR interpolated response compared to that from an EM commercial simulations for d= 5.08 mm

This structure is studied for d varying from 2.45mm to 10 mm. The structure is excited using a gaussian input signal at the port P1. The structure is designed on a Teflon (εr= 2.2) substrate. Figure 8 shows the good accuracy between the FDTD EM simulator results and those using the NN_IIR model for d=5.08 mm This latter results were not furnished while establishing the NN).

VI.3. Chamfered bend test (Frequency Domain).

An NN-IIR model is developed for modeling a microstrip chamfered bend structure like that shown in Fig 9. The test was performed in the frequency range [1 GHz-15 GHz] and for the geometrical parameter, d, varying between 571 µm and 1261µm.

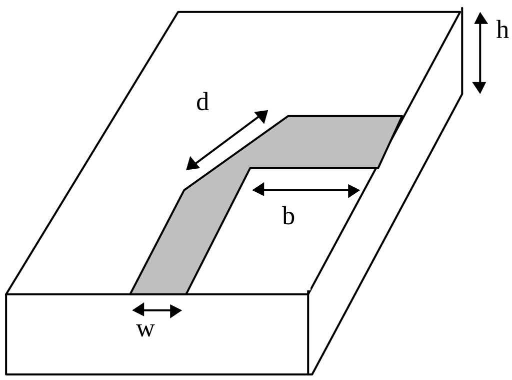


Figure.9. Microstrip chamfered bend structure with h= 635µm, b=500µm and er =9.6

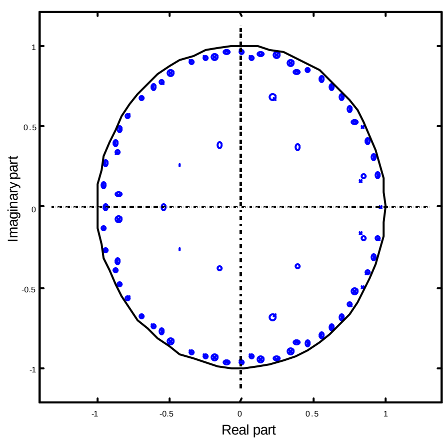


Figure.10. Poles and zeros position in complex plane with d=575

The IIR filter is built for 4 structures corresponding to different values of the geometrical parameter (Fig.9). The studied response is taken to be the S_{11} parameter. A filter order of M=N=70 was needed to fit accurately the system response. However, it is interesting to note that a large set of poles and zeros cancels each other as shown in (Fig.10). Hence, we select only those which are different. It is important to note that conjugate zeros correspond to a minimum pic and conjugate poles to maximum one.
The poles and zeros are then interpolated by a neural network which gives the poles and zeros variation in the z-plane (Fig 11 and 12).

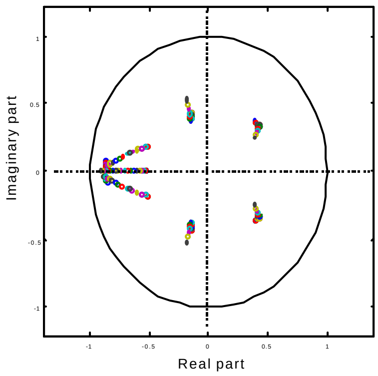


Figure.11. Poles positions in z-plane for d varying between 575µm to 261µm

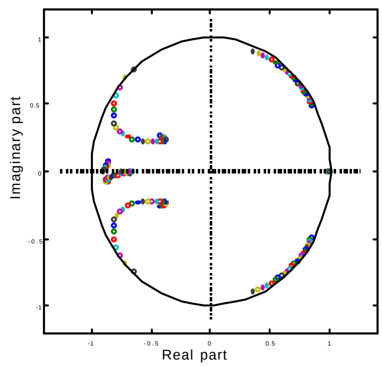


Figure.12. Zeros positions in z-plane for d varying between 575 µm to 1261 µm

For any value inside the considered “d” interval, the corresponding poles and zeros are computed using the neural network and hence, the frequency response of the output parameter can be easily determined. Fig 13 gives a comparison between results for the output response for different values of the geometrical parameter d, once calculated by a rigorous analysis using a frequency domain electromagnetic commercial simulator and another time using our NN-IIR model. The excellent agreement of the results demonstrates the validity of our proposed method

VII. CONCLUSION

A new method has been proposed for precise and rapid EM modeling of complex microwave structures using efficient optimization and parameterization techniques in the time and frequency domains. It combines the powerful learning possibilities of neural networks and those of the accurate fitting of IIR filters in time domain. This method fits the desired response by determining the corresponding transfer function between

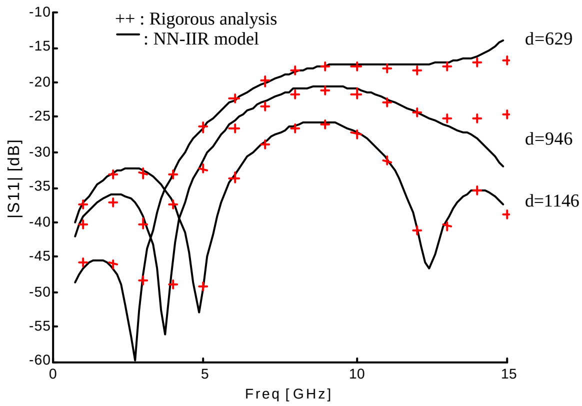


Figure 13. NN-IIR interpolated response compared to that from an EM commercial simulations for different values of d

the input excitation and the output response of the studied structure.

The extension of this technique to applications in the frequency domain is shown to be still precise. The characterization of the transfer function by its poles and zeros decreases the number of samples that are required to be furnished to the neural network to learn the behavior of the response. It is shown that interpolating of poles and zeros in the frequency domain is more precise because it is found that the number of dominant or significant poles and zeros is smaller.

Optimizing using this technique become an easy task as the maximums and the minimums of the IIR filter transfer function correspond to the required poles and zeros.

This technique is distinguished from other known ones by its capacity of interpolating functions (filter transfer function) and not only points.

This work is significant due to the growing demand of having accurate means of parameterization and optimization in existing electromagnetic simulations.

VIII. References

[1] K. Guillouard,et al , « A new global finite element analysis of microwave circuits including lumped elements », IEEE MTT, vol.44, no. 12, pp.2587-2594.

[2] Y. Bégassat, L.A. Ramboz, M.F. Wong, V. Fouad Hanna, J. Citerne, «Modeling of an X-band planar balanced mixer with a mixed SPICE-FDTD program», Microwave and Optical Technology letters, Oct. 1996, vol. 13, no.2, pp. 99-102.

[3] P. Watson and K. C. Gupta. Design and optimization of CPW circuits using EM-ANN models for CPW components. IEEE Transactions on microwave theory and techniques, VOL. 45, No. 12, December 1997. pp 2515-2523.

[4] S. Veluswami, M. Nakhla, Q.J. Zhang. The Application of Neural Network to EM-Based Simulation and Optimization of Interconnects in High-Speed VLSI circuits. MTT (1997), vol 45, N°45, pp. 712-722.

[5] H. Zaabab , Q. J. Zhang, M. S. Nakhla. A Neural Network Modeling Approach to Circuit Optimization and Statistical Design. MTT. (1995). Vol 43, N°6, pp 1349-1358.

[6] W. Kumpel, I. Wolff Digital Signal Processing of Time Domain Filed Simulation Results Using The System Identification Method. IEEE MTT-S Digest, pp 793-796. 1996.

[7] S. Haykin . Neural Networks a Comprehensive Foundation. (1993). IEEE press.

[8] R. S. Adve,et al, Aplication of the Cauchy Method for Extrapolating/Interpolation Narrow Band System Responses , Transactions Microwave Theory and Techniques, Vol 45, No 5, pp 837-845, May 1997.