Current Trends in Network Optimization

JOHN W. BANDLER, MEMBER, IEEE, AND RUDOLPH E. SEVIORA, STUDENT MEMBER, IEEE

Abstract—Some current trends in automated network design optimization which, it is believed, will play a significant role in the computer-aided design of lumped-distributed and microwave networks are reviewed and discussed. In particular, the adjoint network approach due to Director and Rohrer for evaluating the gradient vector of suitable objective functions related to network responses that are to be optimized is presented in a tutorial manner. The advantage of this method is the ease with which the required partial derivatives with respect to variable parameters, such as electrical quantities or geometrical dimensions, can be obtained using at most two network analyses. Least pth and minimax approximation in the frequency domain are considered. Networks consisting of linear time-invariant elements are treated, including the conventional lumped elements, transmission lines, RC lines, coaxial lines, rectangular waveguides, and coupled lines. To illustrate the application of the adjoint network method, an example is given concerning the optimization in the least pth sense using the Fletcher-Powell method of a transmission-line filter with frequency variable terminations previously considered by Carlin and Gupta.

I. Introduction

As the recent special issue on Computer-Oriented Microwave Practices of the IEEE Transactions on Microwave Theory and Techniques shows, microwave network optimization is widely carried out using direct search methods, i.e., iterative optimization methods which do not employ evaluation or estimation of derivatives. Murray-Lasso and Kozemchak [1], for example, used pattern search [2] to optimize the parameters of the transmission-line network shown in Fig. 1. The problem was to match the 50-ohm characteristic impedance of a transmission line to the complex input impedance of the transistor specified at a discrete set of frequencies in the band of interest. The ten parameters were the five lengths and five characteristic impedances. A problem studied by Bandler [3] was the optimization of multisection inhomogeneous rectangular waveguide impedance transformers (Fig. 2). The objective was, within certain constraints, to adjust the geometrical dimensions of the sections such that the input and output waveguides were matched over a given frequency band. In general, all waveguides had different cutoff frequencies. Responses which were optimal in the Chebyshev sense, i.e., minimax, were desired. The razor search method [4] was employed to realize them. A modified version of Rosenbrock's method [2] was used more recently by Trick and Vlach [5] to optimize the broad-band amplifier shown in Fig. 3 with, in general, complex frequency-dependent terminations. A weighted least-squares type of objective function was employed to achieve a flat power gain with a reasonable reflection coefficient in the band of interest.

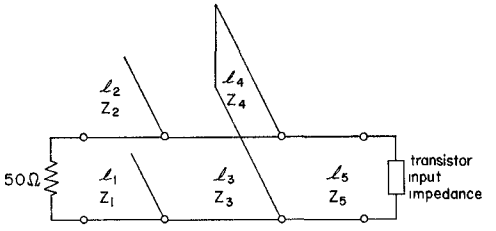

Fig. 1. Matching network optimized by Murray-Lasso and Kozemchak [1].

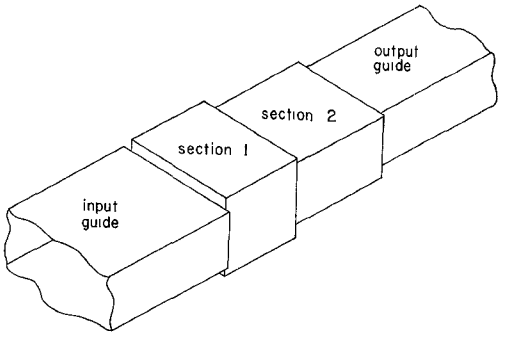

Fig. 2. Inhomogeneous rectangular waveguide impedance transformer optimized by Bandler [3].

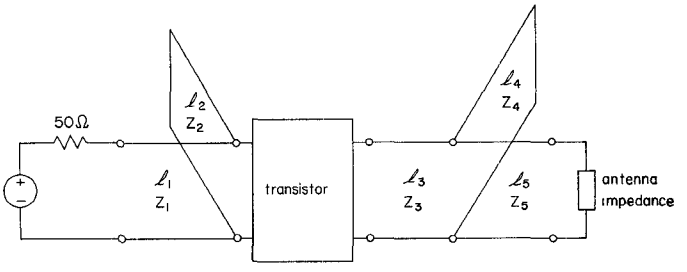

Fig. 3. Broad-band amplifier optimized by Trick and Vlach [5].

These three examples (Figs. 1 to 3) are a good indication of the state of the art in automatic optimization by computer of distributed networks in the microwave region. In the absence of a reasonably simple and efficient method of evaluating derivatives, direct search methods were probably found preferable by the authors instead of gradient methods of minimization. Consider, for example, an m-section cascaded network described

Manuscript received August 10, 1970. This work was supported by the National Research Council of Canada, under Grants A7239 and A5277. This paper is based on two papers presented at the 1970 IEEE G-MTT International Microwave Symposium, Newport Beach, Calif., May 11–14.

J. W. Bandler is with the Department of Electrical Engineering, McMaster University, Hamilton, Ont., Canada.

R. E. Seviora is with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada.

by the $ABCD$ matrix. Then

$$\begin{bmatrix} A_t & B_t \\ C_t & D_t \end{bmatrix} = \prod_{i=1}^{m} \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \tag{1}$$

so that, if some jth parameter ϕ_j appears in the nth section [6]

$$\frac{\partial}{\partial \phi_j} \begin{bmatrix} A_t & B_t \\ C_t & D_t \end{bmatrix} = \prod_{i=1}^{n-1} \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \begin{bmatrix} \frac{\partial A_n}{\partial \phi_j} & \frac{\partial B_n}{\partial \phi_j} \\ \frac{\partial C_n}{\partial \phi_j} & \frac{\partial D_n}{\partial \phi_j} \end{bmatrix} \prod_{i=n+1}^{m} \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix}. \tag{2}$$

In general, the functions involved are highly nonlinear, containing transcendental expressions. If care is not exercised to prevent reevaluation of expressions and formulas already evaluated, it may not make much difference in computing time whether analytic expressions are available for the derivatives, the derivatives are estimated numerically by differences produced by small perturbations in the parameter values, or large steps in the parameters are taken as in direct search methods.

The essence of the adjoint network method originally proposed by Director and Rohrer [7], [8] is that all required partial derivatives of the objective function may be obtained from the results of at most two complete analyses of the network regardless of the number of variable parameters and without actually perturbing them. For design of reciprocal networks on the reflection coefficient basis, for example, only one analysis yields all the information needed to compute the derivatives. The procedure is essentially an exact one, so the components could be in analytic or numerical form.

II. Tellegen's Theorem

Tellegen's theorem [9], [10], [11] is invoked to simplify the necessary derivations. Let

$$\boldsymbol{v} \triangleq \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_b \end{bmatrix} \tag{3}$$

contain all the branch voltages in a network and

$$\boldsymbol{i} \triangleq \begin{bmatrix} i_1 \\ i_2 \\ \cdot \\ \cdot \\ \cdot \\ i_b \end{bmatrix} \tag{4}$$

contain all the corresponding branch currents using associated reference directions [10].[1] Tellegen's theorem

[1] With associated reference directions, the current always enters a branch at the plus sign and leaves at the minus sign.

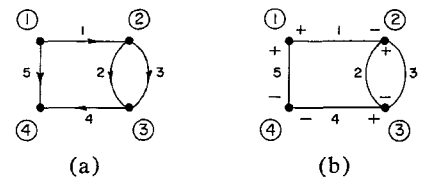

Fig. 4. Two networks having the same topology with nodes and branches correspondingly numbered.

states that if $\boldsymbol{v}$ and $\boldsymbol{i}$ satisfy Kirchhoff's voltage law (KVL) and Kirchoff's current law (KCL), respectively,

$$\boldsymbol{v}^T\boldsymbol{i} = 0. \tag{5}$$

The proof is rather straightforward [10, p. 422]. KVL requires that $\boldsymbol{v} = \boldsymbol{A}^T\boldsymbol{e}$, where $\boldsymbol{A}$ is the reduced incidence matrix of the network and $\boldsymbol{e}$ is the node-to-datum voltage vector. So

$$\boldsymbol{v}^T\boldsymbol{i} = (\boldsymbol{A}^T\boldsymbol{e})^T\boldsymbol{i} = \boldsymbol{e}^T\boldsymbol{A}\boldsymbol{i}.$$

But KCL requires that $\boldsymbol{A}\boldsymbol{i} = \boldsymbol{0}$. Therefore,

$$\boldsymbol{v}^T\boldsymbol{i} = 0.$$

As a numerical example of Tellegen's theorem consider Fig. 4, which represents two networks having the same topology. Let.

$$\boldsymbol{i} = [3 \quad -2 \quad 5 \quad 3 \quad -3]^T$$

refer, for example, to Fig. 4(a), and

$$\boldsymbol{v} = [1 \quad 2 \quad 2 \quad 3 \quad 6]^T$$

to Fig. 4(b). Then

$$\boldsymbol{v}^T\boldsymbol{i} = 3 - 4 + 10 + 9 - 18 = 0.$$

Observe that differences in elements or element values between the networks are irrelevant. Thus, $\boldsymbol{i}$ may be essentially arbitrary but subject to KCL and $\boldsymbol{v}$ arbitrary subject to KVL.

III. The Adjoint Network

We need to define an auxiliary network which is topologically the same as the original or given network which is to be optimized. This is called the adjoint network. Let the variables V and I refer to the original network and $\hat{V}$ and $\hat{I}$ refer to the corresponding quantities of the adjoint network. From (5)

$$\begin{aligned} \boldsymbol{V}_B{}^T\hat{\boldsymbol{I}}_B &= 0 \\ \boldsymbol{I}_B{}^T\hat{\boldsymbol{V}}_B &= 0 \end{aligned} \tag{6}$$

where subscript B implies that the associated vectors contain all corresponding complex branch voltages and currents. Perturbing elements in the original network and noting that Kirchoff's laws and hence Tellegen's theorem are applicable to the incremental changes in current and voltage, namely, $\Delta\boldsymbol{I}_B$ and $\Delta\boldsymbol{V}_B$, respectively,

$$\begin{aligned} \Delta\boldsymbol{V}_B{}^T\hat{\boldsymbol{I}}_B &= 0 \\ \Delta\boldsymbol{I}_B{}^T\hat{\boldsymbol{V}}_B &= 0 \end{aligned} \tag{7}$$

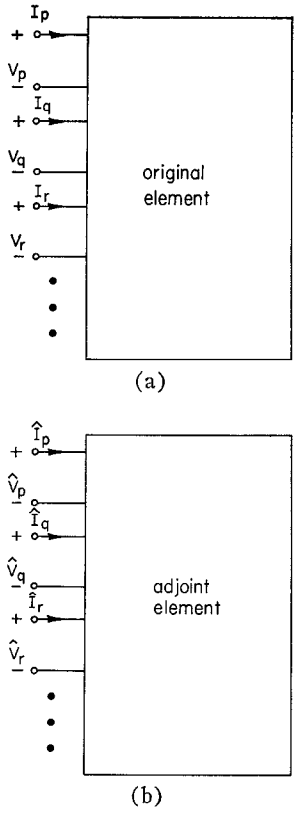

Fig. 5. (a) Multiport original element. (b) Adjoint element.

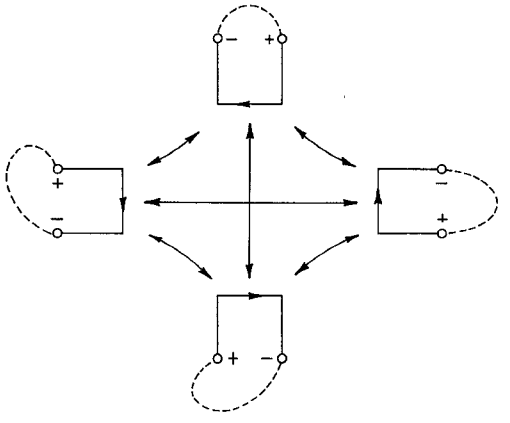

Fig. 6. Representation of multiport element or network for application of Tellegen's theorem.

so that we have the useful form

$$\Delta V_B{}^T \hat{I}_B - \Delta I_B{}^T \hat{V}_B = 0. \tag{8}$$

In general, the network to be optimized will consist of multiport elements (Fig. 5), particularly in the microwave region. To see how Tellegen's theorem may be applied, consider Fig. 6. Obviously, we can still think in terms of network graphs with branch quantities related through some appropriate matrix description. Suppose we take the hybrid matrix description

$$\begin{bmatrix} I_a \\ V_b \end{bmatrix} = \begin{bmatrix} Y & A \\ M & Z \end{bmatrix} \begin{bmatrix} V_a \\ I_b \end{bmatrix} \tag{9}$$

where

$$\begin{bmatrix} I_a \\ I_b \end{bmatrix} = \begin{bmatrix} I_p \\ I_q \\ I_r \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

$$\begin{bmatrix} V_a \\ V_b \end{bmatrix} = \begin{bmatrix} V_p \\ V_q \\ V_r \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}.$$

Perturbing the parameters of the element and neglecting higher order terms

$$\begin{bmatrix} \Delta I_a \\ \Delta V_b \end{bmatrix} = \begin{bmatrix} Y & A \\ M & Z \end{bmatrix} \begin{bmatrix} \Delta V_a \\ \Delta I_b \end{bmatrix} + \begin{bmatrix} \Delta Y & \Delta A \\ \Delta M & \Delta Z \end{bmatrix} \begin{bmatrix} V_a \\ I_b \end{bmatrix}. \tag{10}$$

Substituting (10) in (8), the terms of (8) corresponding to the element are

$$\begin{aligned} &\begin{bmatrix} \Delta V_a \\ \Delta V_b \end{bmatrix}^T \begin{bmatrix} \hat{I}_a \\ \hat{I}_b \end{bmatrix} - \begin{bmatrix} \Delta I_a \\ \Delta I_b \end{bmatrix}^T \begin{bmatrix} \hat{V}_a \\ \hat{V}_b \end{bmatrix} \\ &= \begin{bmatrix} \Delta I_a \\ \Delta V_b \end{bmatrix}^T \begin{bmatrix} -\hat{V}_a \\ \hat{I}_b \end{bmatrix} + \begin{bmatrix} \Delta V_a \\ \Delta I_b \end{bmatrix}^T \begin{bmatrix} \hat{I}_a \\ -\hat{V}_b \end{bmatrix} \\ &= \left(\begin{bmatrix} \Delta V_a \\ \Delta I_b \end{bmatrix}^T \begin{bmatrix} Y & A \\ M & Z \end{bmatrix}^T + \begin{bmatrix} V_a \\ I_b \end{bmatrix}^T \begin{bmatrix} \Delta Y & \Delta A \\ \Delta M & \Delta Z \end{bmatrix}^T \right) \\ &\quad \cdot \begin{bmatrix} -\hat{V}_a \\ \hat{I}_b \end{bmatrix} + \begin{bmatrix} \Delta V_a \\ \Delta I_b \end{bmatrix}^T \begin{bmatrix} \hat{I}_a \\ -\hat{V}_b \end{bmatrix} \end{aligned} \tag{11}$$

which can be reduced to

$$[V_a{}^T \quad I_b{}^T] \begin{bmatrix} -\Delta Y^T & \Delta M^T \\ -\Delta A^T & \phi Z^T \end{bmatrix} \begin{bmatrix} \hat{V}_a \\ \hat{I}_b \end{bmatrix} \tag{12}$$

if

$$\begin{bmatrix} \hat{I}_a \\ \hat{V}_b \end{bmatrix} \equiv \begin{bmatrix} Y^T & -M^T \\ -A^T & Z^T \end{bmatrix} \begin{bmatrix} \hat{V}_a \\ \hat{I}_b \end{bmatrix} \tag{13}$$

which defines the adjoint element. This definition causes the terms of (8) relating to the element to be expressible only in terms of the unperturbed currents and voltages associated with the original and adjoint elements and incremental changes in the elements of the matrix. Terms containing incremental changes in current and voltage have disappeared.

Table I summarizes these results and results for impedance matrix, admittance matrix, and $ABCD$ matrix descriptions. They may be derived independently or as special cases of the derivation for the hybrid matrix. Two important special cases should be noted. The first

TABLE I

Matrix Type	Original Element	Adjoint Element	Expression Yielding Sensitivity
Impedance	$V = ZI$	$\hat{V} = Z^T\hat{I}$	$I^T\Delta Z^T\hat{I}$
Admittance	$I = YV$	$\hat{I} = Y^T\hat{V}$	$-V^T\Delta Y^T\hat{V}$
Hybrid	$\begin{bmatrix} I_a \\ V_b \end{bmatrix} = \begin{bmatrix} Y & A \\ M & Z \end{bmatrix}\begin{bmatrix} V_a \\ I_b \end{bmatrix}$	$\begin{bmatrix} \hat{I}_a \\ \hat{V}_b \end{bmatrix} = \begin{bmatrix} Y^T & -M^T \\ -A^T & Z^T \end{bmatrix}\begin{bmatrix} \hat{V}_a \\ \hat{I}_b \end{bmatrix}$	$[V_a^T\ I_b^T]\begin{bmatrix} -\Delta Y^T & \Delta M^T \\ -\Delta A^T & \Delta Z^T \end{bmatrix}\begin{bmatrix} \hat{V}_a \\ \hat{I}_b \end{bmatrix}$
$ABCD$	$\begin{bmatrix} V_p \\ I_p \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} V_q \\ -I_q \end{bmatrix}$	$\begin{bmatrix} \hat{V}_p \\ \hat{I}_p \end{bmatrix} = \frac{1}{AD-BC}\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} \hat{V}_q \\ -\hat{I}_q \end{bmatrix}$	$[V_q\ I_q]\begin{bmatrix} \Delta A & -\Delta C \\ -\Delta B & \Delta D \end{bmatrix}\begin{bmatrix} \hat{I}_p \\ \hat{V}_p \end{bmatrix}$

is that adjoint of a reciprocal element is identical to the element itself. The second is that a one-port element (resistor, inductor, etc.) is accounted for by Table I.

Suppose the original and adjoint networks are excited by independent sources[2] as indicated by Fig. 7. Let

$$\boldsymbol{V}_V \triangleq [V_1\ V_2 \cdots V_{n_V}]^T \tag{14}$$

be the n_V-element voltage-excitation vector,

$$\boldsymbol{I}_I \triangleq [I_{n_V+1}\ I_{n_V+2} \cdots I_{n_V+n_I}]^T \tag{15}$$

be the n_I-element current-excitation vector, so that

$$\boldsymbol{I}_V \triangleq [I_1\ I_2 \cdots I_{n_V}]^T \tag{16}$$

and

$$\boldsymbol{V}_I \triangleq [V_{n_V+1}\ V_{n_V+2} \cdots V_{n_V+n_I}]^T \tag{17}$$

respectively, are the corresponding response vectors. Thus, subscript V refers to voltage-excited ports and subscript I refers to current-excited ports. For the adjoint network, similar definitions from (14) through (17) would be distinguished by $\hat{}$.

Terms of (8) associated with the port excitations and responses are

$$\Delta\boldsymbol{V}_V^T\hat{\boldsymbol{I}}_V - \Delta\boldsymbol{I}_V^T\hat{\boldsymbol{V}}_V + \Delta\boldsymbol{V}_I^T\hat{\boldsymbol{I}}_I - \Delta\boldsymbol{I}_I^T\hat{\boldsymbol{V}}_I. \tag{18}$$

But $\Delta\boldsymbol{V}_V = \Delta\boldsymbol{I}_I = \boldsymbol{0}$ if the excitations remain constant. Expression (18), therefore, reduces to

$$-\Delta\boldsymbol{I}_V^T\hat{\boldsymbol{V}}_V + \Delta\boldsymbol{V}_I^T\hat{\boldsymbol{I}}_I. \tag{19}$$

In summary, then, (8) consists of terms of the form of (12) and similar ones as in Table I together with (19) leading, in general, to

$$\Delta\boldsymbol{I}_V^T\hat{\boldsymbol{V}}_V - \Delta\boldsymbol{V}_I^T\hat{\boldsymbol{I}}_I = \boldsymbol{G}^T\Delta\boldsymbol{\phi} \tag{20}$$

where $\boldsymbol{G}$ is a vector of sensitivities related to the adjustable network parameters contained in $\boldsymbol{\phi}$. It is seen that (20) relates changes in the port responses to changes in parameter values, which is usually what we are interested in. The form of the right-hand side of (20) is a direct consequence of the definition of the adjoint network.

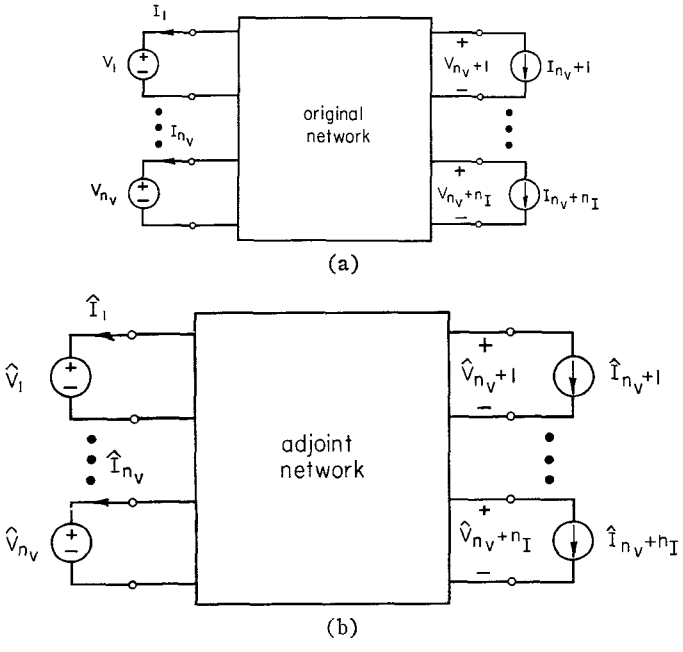

Fig. 7. (a) Excited arbitrary multiport network containing lumped and distributed elements. (b) Topologically equivalent adjoint network with corresponding port excitations.

IV. Derivation of Sensitivities

Table II presents the results of applying the formulas of Table I to a number of commonly used elements. Consider, for example, an inductor. According to the impedance formulas of Table I, the expression yielding the sensitivity is

$$I\Delta Z\hat{I} = (j\omega\, I\hat{I})\Delta L. \tag{21}$$

Taking the inductance L as the parameter, $j\omega I\hat{I}$ is the sensitivity or component of $\boldsymbol{G}$ and ΔL is the parameter increment.

Now consider a uniformly distributed line as shown in Fig. 8(a). The element is reciprocal, so that

$$\boldsymbol{Z}^T = \boldsymbol{Z} = Z\begin{bmatrix} \coth\theta & \operatorname{csch}\theta \\ \operatorname{csch}\theta & \coth\theta \end{bmatrix} \tag{22}$$

[2] Appropriate zero-valued sources are placed, for convenience, at ports which are not excited.

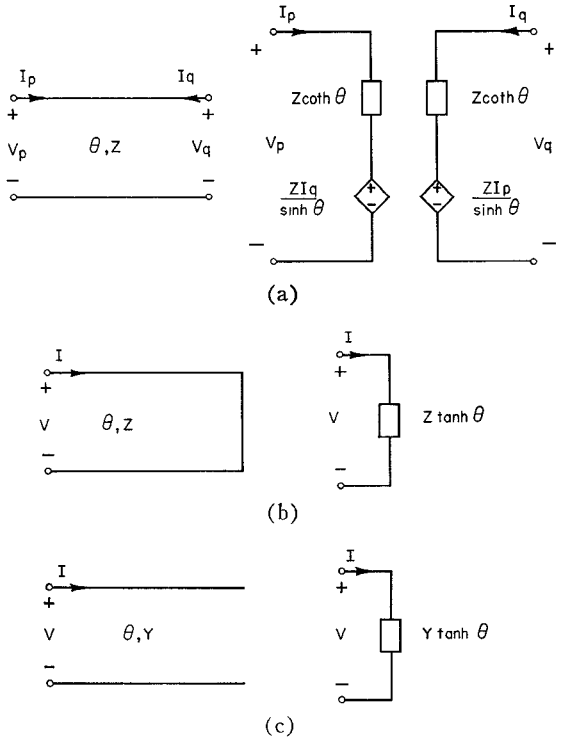

Fig. 8. Uniformly distributed elements with convenient representations. (a) Uniform line. (b) Short-circuited line. (c) Open-circuited line.

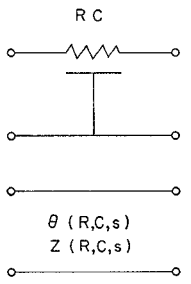

Fig. 9. Uniformly distributed RC line.

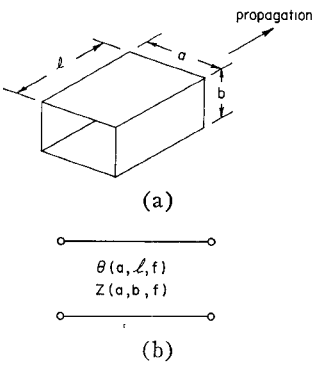

Fig. 10. Rectangular waveguide. (a) Geometrical dimensions. (b) Circuit representation.

where Z is the characteristic impedance. Using the same formula in Table I as for the inductor [12]

$$I^T\Delta Z^T\hat{I} = I^T\left(\Delta Z\begin{bmatrix}\coth\theta & \operatorname{csch}\theta\\ \operatorname{csch}\theta & \coth\theta\end{bmatrix} - \frac{Z\Delta\theta}{\sinh\theta}\begin{bmatrix}\operatorname{csch}\theta & \coth\theta\\ \coth\theta & \operatorname{csch}\theta\end{bmatrix}\right)^T\hat{I}$$

$$= \left(\frac{\Delta Z}{Z}ZI - \frac{\Delta\theta}{\sinh\theta}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}ZI\right)^T\hat{I}$$

$$= \frac{\Delta Z}{Z}V^T\hat{I} - \frac{\Delta\theta}{\sinh\theta}V^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{I}. \quad (23)$$

Corresponding expressions for the lossless transmission line of length l with $\theta = j\beta l$ and the uniform RC line (Fig. 9) with $Z = \sqrt{R/sC}$ and $\theta = \sqrt{sRC}$ are readily obtained [12] and are shown in Table II.

Consider a rectangular waveguide operating in the H_{10} mode, as shown in Fig. 10. The following model may be used if the restrictions outlined by Bandler [3] are observed:

$$Z = b\lambda_g \quad (24)$$

$$\theta = j\frac{2\pi l}{\lambda_g} = j\beta_g l \quad (25)$$

where

$$\lambda_g = \frac{\lambda}{\sqrt{1-(\lambda/2a)^2}} \quad (26)$$

where a, b, and l are the width, height, and length, respectively of the waveguide; λ_g is the guide wavelength and $\lambda = c/f$. It is readily shown, neglecting higher order terms, that

$$\Delta Z = \lambda_g\Delta b - \frac{b\lambda_g{}^3}{4a^3}\Delta a \quad (27)$$

and

$$\Delta\theta = j\beta_g\Delta l + j\frac{\beta_g l\lambda_g{}^2}{4a^3}\Delta a. \quad (28)$$

Expression (23) for the rectangular waveguide then becomes

$$-\Delta a\frac{\lambda_g{}^2}{4a^3}\left(V^T\hat{I} + \frac{\beta_g l}{\sin\beta_g l}V^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{I}\right) + \frac{\Delta b}{b}V^T\hat{I} - \frac{\beta_g\Delta l}{\sin\beta_g l}V^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{I}. \quad (29)$$

Note that the voltages and currents do not necessarily have to have any physical interpretation, their use is only in being convenient variables for analysis.

Now consider a uniform lossless coaxial line with

$$Z = \frac{1}{2\pi}\frac{Z_0}{\sqrt{\epsilon_r}}\ln\frac{d_0}{d_i} \quad (30)$$

where $Z_0 = \sqrt{\mu_0/\epsilon_0}$, ϵ_r is the relative permittivity of the medium, and d_0 and d_i are the outer and inner diameters,

TABLE II

SENSITIVITY EXPRESSIONS FOR SOME LUMPED AND UNIFORMLY DISTRIBUTED ELEMENTS

Element	Equation	Sensitivity (component of $\boldsymbol{G}$)	Increment (component of $\Delta\phi$)
Resistor	$V = RI$	$I\hat{I}$	ΔR
	$I = GV$	$-V\hat{V}$	ΔG
Inductor	$V = j\omega LI$	$j\omega I\hat{I}$	ΔL
	$I = \frac{1}{j\omega}\Gamma V$	$-\frac{1}{j\omega}V\hat{V}$	$\Delta\Gamma$
Capacitor	$V = \frac{1}{j\omega}SI$	$\frac{1}{j\omega}I\hat{I}$	ΔS
	$I = j\omega CV$	$-j\omega V\hat{V}$	ΔC
Transformer	$\begin{bmatrix} V_p \\ I_q \end{bmatrix} = \begin{bmatrix} 0 & n \\ -n & 0 \end{bmatrix}\begin{bmatrix} I_p \\ V_q \end{bmatrix}$	$V_q\hat{I}_p + I_p\hat{V}_q$	Δn
Gyrator	$\boldsymbol{V} = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}\boldsymbol{I}$	$I_p\hat{I}_q - I_q\hat{I}_p$	$\Delta\alpha$
Voltage controlled voltage source	$\begin{bmatrix} I_p \\ V_q \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mu & 0 \end{bmatrix}\begin{bmatrix} V_p \\ I_q \end{bmatrix}$	$V_p\hat{I}_q$	$\Delta\mu$
Voltage controlled current source	$\boldsymbol{I} = \begin{bmatrix} 0 & 0 \\ g_m & 0 \end{bmatrix}\boldsymbol{V}$	$-V_p\hat{V}_q$	Δg_m
Current controlled voltage source	$\boldsymbol{V} = \begin{bmatrix} 0 & 0 \\ r_m & 0 \end{bmatrix}\boldsymbol{I}$	$I_p\hat{I}_q$	Δr_m
Current controlled current source	$\begin{bmatrix} V_p \\ I_q \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}\begin{bmatrix} I_p \\ V_q \end{bmatrix}$	$-I_p\hat{V}_q$	$\Delta\beta$
Short-circuited uniformly distributed line	$V = Z\tanh\theta\, I$	$\tanh\theta\, I\hat{I}$	ΔZ
		$Z\,\mathrm{sech}^2\theta\, I\hat{I}$	$\Delta\theta$
	$I = Y\coth\theta\, V$	$-\coth\theta\, V\hat{V}$	ΔY
		$Y\,\mathrm{csch}^2\theta\, V\hat{V}$	$\Delta\theta$
Open-circuited uniformly distributed line	$V = Z\coth\theta\, I$	$\coth\theta\, I\hat{I}$	ΔZ
		$-Z\,\mathrm{csch}^2\theta\, I\hat{I}$	$\Delta\theta$
	$I = Y\tanh\theta\, V$	$-\tanh\theta\, V\hat{V}$	ΔY
		$-Y\,\mathrm{sech}^2\theta\, V\hat{V}$	$\Delta\theta$
Uniformly distributed line	$\boldsymbol{V} = Z\begin{bmatrix} \coth\theta & \mathrm{csch}\,\theta \\ \mathrm{csch}\,\theta & \coth\theta \end{bmatrix}\boldsymbol{I}$	$\frac{1}{Z}\boldsymbol{V}^T\hat{\boldsymbol{I}}$	ΔZ
		$-\frac{1}{\sinh\theta}\boldsymbol{V}^T\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\hat{\boldsymbol{I}}$	$\Delta\theta$
	$\boldsymbol{I} = Y\begin{bmatrix} \coth\theta & -\mathrm{csch}\,\theta \\ -\mathrm{csch}\,\theta & \coth\theta \end{bmatrix}\boldsymbol{V}$	$-\frac{1}{Y}\boldsymbol{I}^T\hat{\boldsymbol{V}}$	ΔY
		$-\frac{1}{\sinh\theta}\boldsymbol{I}^T\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\hat{\boldsymbol{V}}$	$\Delta\theta$
Short-circuited lossless transmission line	$V = jZ\tan\beta l\, I$	$j\tan\beta l\, I\hat{I}$	ΔZ
		$jZ\beta\sec^2\beta l\, I\hat{I}$	Δl
	$I = -jY\cot\beta l\, V$	$j\cot\beta l\, V\hat{V}$	ΔY
		$-jY\beta\csc^2\beta l\, V\hat{V}$	Δl
Open-circuited lossless transmission line	$V = -jZ\cot\beta l\, I$	$-j\cot\beta l\, I\hat{I}$	ΔZ
		$jZ\beta\csc^2\beta l\, I\hat{I}$	Δl
	$I = jY\tan\beta l\, V$	$-j\tan\beta l\, V\hat{V}$	ΔY
		$-jY\beta\sec^2\beta l\, V\hat{V}$	Δl
Lossless transmission line	$\boldsymbol{V} = -jZ\begin{bmatrix} \cot\beta l & \csc\beta l \\ \csc\beta l & \cot\beta l \end{bmatrix}\boldsymbol{I}$	$\frac{1}{Z}\boldsymbol{V}^T\hat{\boldsymbol{I}}$	ΔZ

TABLE II (*Cont.*)

Element	Equation	Sensitivity (component of G)	Increment (component of $\Delta\phi$)
		$-\frac{\beta}{\sin\beta l}\boldsymbol{V}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{I}}$	Δl
Lossless transmission line	$\boldsymbol{I}=-jY\begin{bmatrix}\cot\beta l & -\csc\beta l\\ -\csc\beta l & \cot\beta l\end{bmatrix}\boldsymbol{V}$	$-\frac{1}{Y}\boldsymbol{I}^T\hat{\boldsymbol{V}}$	ΔY
		$-\frac{\beta}{\sin\beta l}\boldsymbol{I}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{V}}$	Δl
Rectangular waveguide operating in H_{10} mode	as for lossless transmission line with $Z=b\lambda_g$, β replaced by $\beta_g=2\pi/\lambda_g$, where $\lambda_g=\lambda/\sqrt{1-(\lambda/2a)^2}$	$-\frac{\lambda_g^2}{4a^3}\boldsymbol{V}^T\begin{bmatrix}1&\frac{\beta_g l}{\sin\beta_g l}\\ \frac{\beta_g l}{\sin\beta_g l}&1\end{bmatrix}\hat{\boldsymbol{I}}$	Δa
		$\frac{1}{b}\boldsymbol{V}^T\hat{\boldsymbol{I}}$	Δb
		$-\frac{\beta_g}{\sin\beta_g l}\boldsymbol{V}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{I}}$	Δl
Uniform RC line	as for uniformly distributed line with $Z=\sqrt{\frac{R}{sC}}$ and $\theta=\sqrt{sRC}$	$\frac{1}{2R}\boldsymbol{V}^T\begin{bmatrix}1&-\frac{\theta}{\sinh\theta}\\ -\frac{\theta}{\sinh\theta}&1\end{bmatrix}\hat{\boldsymbol{I}}$	ΔR
		$-\frac{1}{2C}\boldsymbol{V}^T\begin{bmatrix}1&\frac{\theta}{\sinh\theta}\\ \frac{\theta}{\sinh\theta}&1\end{bmatrix}\hat{\boldsymbol{I}}$	ΔC
Uniform coaxial line	as for lossless transmission line with $Z=\frac{1}{2\pi}\frac{Z_0}{\sqrt{\epsilon_r}}\ln\frac{d_0}{d_i}$ and $\beta=\beta_0\sqrt{\epsilon_r}$	$\frac{1}{d_0\ln\frac{d_0}{d_i}}\boldsymbol{V}^T\hat{\boldsymbol{I}}$	Δd_0
		$-\frac{1}{d_i\ln\frac{d_0}{d_i}}\boldsymbol{V}^T\hat{\boldsymbol{I}}$	Δd_i
		$-\frac{\beta_0\sqrt{\epsilon_r}}{\sin\beta_0\sqrt{\epsilon_r}l}\boldsymbol{V}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{I}}$	Δl
		$-\frac{1}{2\epsilon_r}\left(\boldsymbol{V}^T\hat{\boldsymbol{I}}+\frac{\beta_0\sqrt{\epsilon_r}l}{\sin\beta_0\sqrt{\epsilon_r}l}\boldsymbol{V}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{I}}\right)$	$\Delta\epsilon_r$
Coupled lines (1) capacitance matrix description	$\boldsymbol{I}=c\begin{bmatrix}C\coth\theta & -C\,\mathrm{csch}\,\theta\\ -C\,\mathrm{csch}\,\theta & C\coth\theta\end{bmatrix}\boldsymbol{V}$ where $C\triangleq\begin{bmatrix}C_{01}+C_{12} & -C_{12}\\ -C_{12} & C_{02}+C_{12}\end{bmatrix}$ (see text and Fig. 11)	$-c[V_{1p}\,V_{1q}]\begin{bmatrix}\coth\theta & -\mathrm{csch}\,\theta\\ -\mathrm{csch}\,\theta & \coth\theta\end{bmatrix}\begin{bmatrix}\hat{V}_{1p}\\ \hat{V}_{1q}\end{bmatrix}$	ΔC_{01}
		$-c[V_{2p}\,V_{2q}]\begin{bmatrix}\coth\theta & -\mathrm{csch}\,\theta\\ -\mathrm{csch}\,\theta & \coth\theta\end{bmatrix}\begin{bmatrix}\hat{V}_{2p}\\ \hat{V}_{2q}\end{bmatrix}$	ΔC_{02}
		$-c\boldsymbol{V}^T\begin{bmatrix}1'\coth\theta & -1'\,\mathrm{csch}\,\theta\\ -1'\,\mathrm{csch}\,\theta & 1'\coth\theta\end{bmatrix}\hat{\boldsymbol{V}}$	ΔC_{12}
		$-\frac{1}{\sinh\theta}\boldsymbol{I}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{V}}$ (see text for definitions of 1′, 1, and 0)	$\Delta\theta$
Coupled lines (2) even- and odd-mode description for symmetrical arrangement	$\boldsymbol{I}=-j\begin{bmatrix}\boldsymbol{Y}_M\cot\beta l & -\boldsymbol{Y}_M\csc\beta l\\ -\boldsymbol{Y}_M\csc\beta l & \boldsymbol{Y}_M\cot\beta l\end{bmatrix}$ where $\boldsymbol{Y}_M\triangleq\frac{1}{2}\begin{bmatrix}Y_e+Y_o & Y_e-Y_o\\ Y_e-Y_o & Y_e+Y_o\end{bmatrix}$ and where Y_e and Y_o are the even- and odd-mode admittances	$-\frac{1}{2Y_e}\boldsymbol{V}^T\begin{bmatrix}\hat{I}_{1p}+\hat{I}_{2p}\\ \hat{I}_{1p}+\hat{I}_{2p}\\ \hat{I}_{1q}+\hat{I}_{2q}\\ \hat{I}_{1q}+\hat{I}_{2q}\end{bmatrix}$	ΔY_e
		$-\frac{1}{2Y_o}\boldsymbol{V}^T\begin{bmatrix}\hat{I}_{1p}-\hat{I}_{2p}\\ \hat{I}_{2p}-\hat{I}_{1p}\\ \hat{I}_{1q}-\hat{I}_{2q}\\ \hat{I}_{2q}-I_{1q}\end{bmatrix}$	ΔY_o
		$-\frac{\beta}{\sin\beta l}\boldsymbol{I}^T\begin{bmatrix}0&1\\1&0\end{bmatrix}\hat{\boldsymbol{V}}$	Δl

respectively, of the line. Here,

$$\theta = j\beta_0\sqrt{\epsilon_r}\,l \tag{31}$$

where β_0 is the free-space phase constant. Thus

$$\Delta Z = Z\left[\frac{\Delta d_0}{d_0 \ln\frac{d_0}{d_i}} - \frac{\Delta d_i}{d_i \ln\frac{d_0}{d_i}} - \frac{\Delta\epsilon_r}{2\epsilon_r}\right] \tag{32}$$

and

$$\Delta\theta = j\beta_0\sqrt{\epsilon_r}\,\Delta l + \frac{j\beta_0 l\Delta\epsilon_r}{2\sqrt{\epsilon_r}}. \tag{33}$$

Expression (23) for the coaxial line becomes

$$\left[\frac{\Delta d_0}{d_0 \ln\frac{d_0}{d_i}} - \frac{\Delta d_i}{d_i \ln\frac{d_0}{d_i}}\right] \boldsymbol{V}^T\hat{\boldsymbol{I}} - \frac{\beta_0\sqrt{\epsilon_r}\,\Delta l}{\sin\beta_0\sqrt{\epsilon_r}\,l}\,\boldsymbol{V}^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{\boldsymbol{I}}$$
$$-\frac{\Delta\epsilon_r}{2\epsilon_r}\left(\boldsymbol{V}^T\hat{\boldsymbol{I}} + \frac{\beta_0\sqrt{\epsilon_r}\,l}{\sin\beta_0\sqrt{\epsilon_r}\,l}\,\boldsymbol{V}^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{\boldsymbol{I}}\right). \tag{34}$$

Finally, consider the admittance matrix formulas of Table I applied to the pair of coupled lines above a ground plane [13], [14] shown in Fig. 11. The admittance matrix description is

$$\begin{bmatrix}I_{1p}\\ I_{2p}\\ I_{1q}\\ I_{2q}\end{bmatrix} = c\begin{bmatrix}\boldsymbol{C}\coth\theta & -\boldsymbol{C}\operatorname{csch}\theta\\ -\boldsymbol{C}\operatorname{csch}\theta & \boldsymbol{C}\coth\theta\end{bmatrix}\begin{bmatrix}V_{1p}\\ V_{2p}\\ V_{1q}\\ V_{2q}\end{bmatrix} \tag{35}$$

where subscript p denotes the two ports formed between each conductor and the ground plane at one end and q the corresponding ports at the other end; subscript 1 refers to one conductor and 2 to the other. The matrix $\boldsymbol{C}$ is given by

$$\boldsymbol{C} \triangleq \begin{bmatrix}C_{01}+C_{12} & -C_{12}\\ -C_{12} & C_{02}+C_{12}\end{bmatrix} \tag{36}$$

the elements of which are defined in Fig. 11. Treating C_{01}, C_{02}, C_{12}, and θ as variables we have

$$-\boldsymbol{V}^T\Delta\boldsymbol{Y}^T\hat{\boldsymbol{V}}$$
$$= -c[V_{1p}V_{1q}]\begin{bmatrix}\coth\theta & -\operatorname{csch}\theta\\ -\operatorname{csch}\theta & \coth\theta\end{bmatrix}\begin{bmatrix}\hat{V}_{1p}\\ \hat{V}_{1q}\end{bmatrix}\Delta C_{01}$$
$$-c[V_{2p}V_{2q}]\begin{bmatrix}\coth\theta & -\operatorname{csch}\theta\\ -\operatorname{csch}\theta & \coth\theta\end{bmatrix}\begin{bmatrix}\hat{V}_{2p}\\ \hat{V}_{2q}\end{bmatrix}\Delta C_{02}$$
$$-c\boldsymbol{V}^T\begin{bmatrix}\mathbf{1}'\coth\theta & -\mathbf{1}'\operatorname{csch}\theta\\ -\mathbf{1}'\operatorname{csch}\theta & \mathbf{1}'\coth\theta\end{bmatrix}\hat{\boldsymbol{V}}\Delta C_{12}$$
$$-\frac{c}{\sinh\theta}\boldsymbol{V}^T\begin{bmatrix}-\boldsymbol{C}\operatorname{csch}\theta & \boldsymbol{C}\coth\theta\\ \boldsymbol{C}\coth\theta & -\boldsymbol{C}\operatorname{csch}\theta\end{bmatrix}\hat{\boldsymbol{V}}\Delta\theta \tag{37}$$

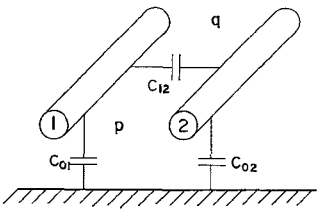

Fig. 11. Coupled lines above a ground plane with static capacitances per unit length.

where

$$\mathbf{1}' \triangleq \begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}. \tag{38}$$

The last term may be rewritten as

$$-\frac{1}{\sinh\theta}\boldsymbol{I}^T\begin{bmatrix}\mathbf{0} & \mathbf{1}\\ \mathbf{1} & \mathbf{0}\end{bmatrix}\hat{\boldsymbol{V}}\Delta\theta \tag{39}$$

where

$$\mathbf{0} \triangleq \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \tag{40}$$

and

$$\mathbf{1} \triangleq \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}. \tag{41}$$

These results are summarized in Table II, along with expressions based on the approach using even- and odd-mode characteristic admittances.

V. Gradient Computations

There are a number of ways in which the adjoint network method can be used effectively in gradient computations.

Consider Figs. 12 and 13. Fig. 12(a) depicts the situation when insertion loss or gain is to be optimized. Here we are interested at some frequency in the partial derivatives of I_L with respect to the parameters and hence ∇I_L. Fig. 13(a) is appropriate for design on the reflection coefficient basis. In this case we are interested at some frequency in ∇I_g. Suppose the adjoint networks are excited as shown in Figs. 12(b) and 13(b). Then, for Fig. 12, (20) can be reduced to

$$\Delta I_L\hat{V}_L = \boldsymbol{G}^T\Delta\boldsymbol{\phi}. \tag{42}$$

Dividing by $\hat{V}_L$ we have

$$\Delta I_L = \nabla I_L{}^T\Delta\boldsymbol{\phi} = \left[\frac{1}{\hat{V}_L}\boldsymbol{G}^T\right]\Delta\boldsymbol{\phi}$$

from which

$$\nabla I_L = \frac{1}{\hat{V}_L}\boldsymbol{G}. \tag{43}$$

For Fig. 13, (20) can be reduced to

$$\Delta I_g\hat{V}_g = \boldsymbol{G}^T\Delta\boldsymbol{\phi}. \tag{44}$$

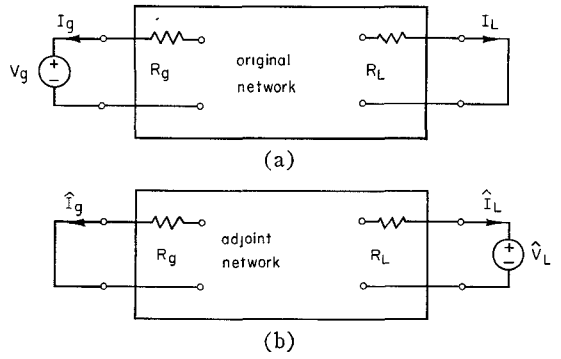

Fig. 12. Special case of Fig. 7 for insertion loss design.

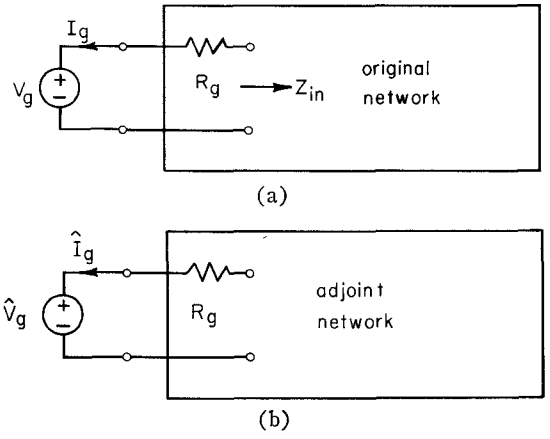

Fig. 13. Special case of Fig. 7 for reflection coefficient design.

Noting that (44) has the same form as (42) we get

$$\nabla I_g = \frac{1}{\hat{V}_g} G. \tag{45}$$

Observe that ∇I_L in (43) and ∇I_g in (45) are evaluated from the currents and voltages present in the unperturbed original and adjoint networks. At most, two network analyses using *any* suitable method will, therefore, yield the information required for the evaluation. Of course, if desired, analytic expressions for the partial derivatives could also be found by this approach.[3] It is interesting to note that for design of reciprocal networks on the reflection coefficient basis we are at liberty to set $\hat{V}_g \equiv V_g$ and use the results of just *one* analysis at each frequency.

To relate ∇I_L or ∇I_g to the gradient vector of suitable least pth or minimax objective functions [2]–[4], [6], [15]–[18] is a straightforward process [12]. In anticipation of the numerical example (Section VI), we will first consider discrete least pth approximation using the reflection coefficient. Let

$$U = \sum_{\Omega_d} \frac{1}{p} \mid \rho(j\omega_d) \mid^p, \qquad \omega_d \in \Omega_d \tag{46}$$

[3] It is debatable, however, whether any computational advantage would, in general, be gained by deriving analytic expressions.

where ρ is the reflection coefficient between R_g and the one-port network, Ω_d is a set of discrete frequencies ω_d, and p is any positive integer. Suppose it is required to minimize U. In this case we are trying to approximate zero reflection coefficient in a least pth sense. For large p we would expect a nearly equal-ripple response to correspond to the minimum of U [2].

$$\rho(j\omega_d) = \frac{Z_{in}(j\omega_d) - R_g}{Z_{in}(j\omega_d) + R_g} = 1 - \frac{2R_g}{Z_{in}(j\omega_d) + R_g}$$
$$= 1 + \frac{2R_g I_g(j\omega_d)}{V_g(j\omega_d)} \tag{47}$$

so that

$$\nabla U = \sum_{\Omega_d} \mathrm{Re}\left\{ \mid \rho(j\omega_d) \mid^{p-2} \rho^*(j\omega_d) \nabla \rho(j\omega_d) \right\}$$
$$= \sum_{\Omega_d} \mathrm{Re}\left\{ \frac{2R_g}{V_g(j\omega_d)} \mid \rho(j\omega_d) \mid^{p-2} \rho^*(j\omega_d) \nabla I_g(j\omega_d) \right\}. \tag{48}$$

If, instead of minimizing (46), the problem is to minimize a nonnegative independent variable U, subject to

$$U \geq g(\omega_d) = \tfrac{1}{2} \mid \rho(j\omega_d) \mid^2, \qquad \omega_d \in \Omega_d \tag{49}$$

then we have minimax approximation [2], for which

$$\nabla g(\omega_d) = \mathrm{Re}\left\{ \frac{2R_g}{V_g(j\omega_d)} \rho^*(j\omega_d) \nabla I_g(j\omega_d) \right\}. \tag{50}$$

Finally, let us address ourselves to the approximation problem considered by Director and Rohrer [8], generalizing it to least p [19], [20]. Equation (20) is readily rearranged to give

$$G = \sum_{i=1}^{n_V} \hat{V}_i \nabla I_i - \sum_{i=n_V+1}^{n_V+n_I} \hat{I}_i \nabla V_i \tag{51}$$

since, neglecting higher order terms,

$$\Delta I_i = \nabla I_i{}^T \Delta\phi$$
$$\Delta V_i = \nabla V_i{}^T \Delta\phi.$$

Given, for example, the objective function

$$U = \sum_{i=1}^{n_V+n_I} \int_\Omega \frac{1}{p} \mid e_i(\phi, j\omega) \mid^p d\omega \tag{52}$$

where

$$e_i(\phi, j\omega) \triangleq w_i(\omega)(F_i(\phi, j\omega) - S_i(j\omega)) \tag{53}$$

where

$$F_i(\phi, j\omega) \triangleq \begin{cases} I_i(\phi, j\omega) & i = 1, 2, \cdots, n_V \\ V_i(\phi, j\omega) & i = n_V + 1, \cdots, n_V + n_I \end{cases} \tag{54}$$

and Ω defines the frequency range of interest. Here, $S_i(j\omega)$ is a desired complex port response with $w_i(\omega)$ a nonnegative real weighting function. In this case

$$\nabla U = \sum_{i=1}^{n_V+n_I} \int_\Omega \mathrm{Re}\left\{ \mid e_i(\phi, j\omega) \mid^{p-2} w_i(\omega) \right.$$
$$\left. \cdot e_i^*(\phi, j\omega) \nabla F_i(\phi, j\omega) \right\} d\omega. \tag{55}$$

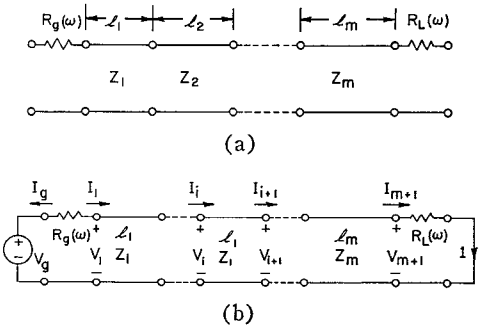

Fig. 14. Cascaded transmission lines terminated in frequency variable resistances.

By comparing (51), (54), and (55) it is seen that by arranging for the adjoint network voltage and current excitations to be given by

$$|e_i(\phi, j\omega)|^{p-2} w_i(\omega) e_i^*(\phi, j\omega) = \begin{cases} \hat{V}_i(j\omega) & i = 1, 2, \cdots, n_V \\ -\hat{I}_i(j\omega) & i = n_V + 1, \cdots, n_V + n_I \end{cases} \tag{56}$$

we obtain

$$\nabla U = \int_\Omega \mathrm{Re}\,\{\boldsymbol{G}\}\, d\omega. \tag{57}$$

If there is no excitation at a particular port the appropriate source is obviously set to zero. If the response at a particular port is not to be controlled the corresponding adjoint excitation should be zero. Elements or parameters not to be varied during optimization do not, of course, contribute to ϕ or $\boldsymbol{G}$.

VI. Example

Carlin and Gupta [21] recently considered the optimal design of filters with lumped–distributed elements or frequency-variable terminations. Although any of their design examples are amenable to computer-oriented optimization techniques, let us discuss the design of the symmetrical seven-section cascaded transmission-line filter shown in Fig. 14(a).

The terminating impedances are real but frequency dependent, specifically

$$R_g(\omega) = R_L(\omega) = 377/\sqrt{1 - (f_c/f)^2}$$

where f is the frequency in GHz and

$$f_c = 2.077 \text{ GHz}.$$

Thus, the terminating impedances can be thought of as rectangular waveguides operating in the H_{10} mode with cutoff frequency 2.077 GHz. Carlin and Gupta required a passband insertion loss of less than 0.4 dB over 2.16 to 3 GHz and an edge to the useful band of 5 GHz. They constrained all section lengths to be 1.5 cm so that each section would be quarter wave at 5 GHz and causing the maximum insertion loss to occur at that frequency. The response of their design is shown in Figs.

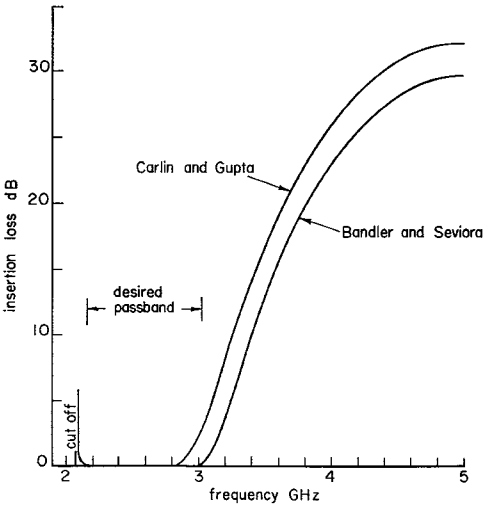

Fig. 15. The response of the seven-section filter whose configuration is shown in Fig. 14. The authors' response was optimized for minimum passband insertion loss.

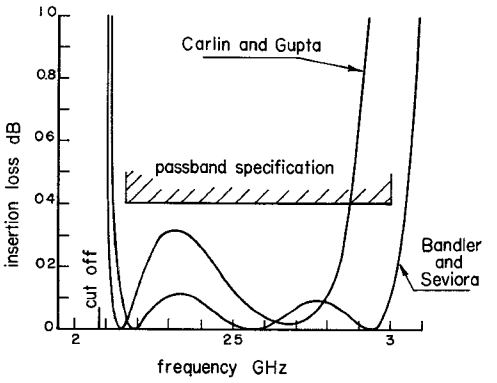

Fig. 16. Details of the passband insertion loss of the seven-section filter.

15 and 16 and the values of characteristic impedance in Table III.

A question of interest to the present authors is this: how small can the passband insertion loss be made under the constraints of the problem? (Note that the question is trivial if the terminating impedances were frequency independent or if the section lengths were freely variable.)

The least pth objective function of (46) was set up using 51 uniformly spaced points over the range 2.16 to 3 GHz and with $p = 10$. Optimization was carried out by the Fletcher–Powell method [22], the required derivatives being evaluated from the results of *one* analysis of the network of Fig. 14(b). To apply the adjoint network method a simple $ABCD$ matrix analysis algorithm employing the approach indicated in Fig. 14(b) was written. Instead of fixing V_g, it was found more convenient to assume that $I_L = 1$ and to calculate the required currents and voltages including V_g. The appropriate formulas from Table II were used (not forgetting to reverse the currents at the junctions when necessary). The design parameter values of Carlin and Gupta were used as starting values.

TABLE III
COMPARISON OF PARAMETER VALUES FOR THE SEVEN-SECTION FILTER

Characteristic Impedances (normalized)	Carlin and Gupta [21]	Bandler and Seviora
Z_1	1476.5	1469.5
Z_2	733.6	763.2
Z_3	1963.6	1945.1
Z_4	461.8	558.7
Z_5	1963.6	1945.1
Z_6	733.6	763.2
Z_7	1476.5	1469.5

The resulting response is plotted in Figs. 15 and 16 and the final parameter values are given in Table III. Observe the almost equal-ripple behavior of the response with a maximum insertion loss over the passband of about 0.1 dB. It would appear then that under the design constraints imposed by Carlin and Gupta, a much lower maximum passband insertion loss is probably not achieveable. This was verified more recently by applying a minimax approximation algorithm to the same problem. A substantially equal-ripple response was obtained with a maximum insertion loss of 0.086 dB. (The algorithm uses the general philosophy behind the razor search method [4] but relies on gradient information generated by the adjoint network method.)

The reader should note that our design is not optimal in the filtering sense required by Carlin and Gupta; to achieve this one would want to maximize the stopband insertion loss subject to a passband insertion loss less than or equal to 0.4 dB. Allowing the section lengths to vary might also improve the response somewhat.

VII. Discussion

A nonexistent lumped element may be thought of as an appropriate zero-valued element connected between two nodes. Since the gradients depend only on voltages between nodes and currents through branches, they may be evaluated with respect to such nonexistent elements. If an increase in element value is indicated, the element can be grown from a short circuit or open circuit, as appropriate. Thus, changes in topology can be accommodated by this means. The adjoint network method does not seem, however, to provide any clear advantage over other methods as an aid to choosing the best topology except possibly in computation time. A direct search method, for example, can also investigate changes with respect to zero-valued elements.

As the authors have found [12], it is not, in general, obvious what kind of element should be grown, whether lumped or distributed, when distributed elements are also allowed. A knowledge only of the currents and voltages is not really sufficient. Furthermore, a variety of physical, economic, and other practical constraints on circuit configuration will also affect the choice. How would one decide, for example, whether a short-circuited transmission line should be grown rather than a lumped inductor?

Many circuit designers claim to have had success in computer-aided network design using direct search methods, so why should they adopt a gradient method? Well, if it is steepest descent they are thinking of, they are better off using the direct search methods. The Fletcher–Powell method [22], on the other hand, is, at the time of writing, still most widely acknowledged as the most powerful unconstrained minimization method available. Factors affecting the choice of an optimization method undoubtedly include familiarity with a particular program, the presence of constraints, the type of approximation required (whether least pth or minimax), the number of variables, and the available computation system [23]. As far as approximation methods are concerned, algorithms which should benefit considerably from the adjoint network method of evaluating derivatives are the minimax approximation methods of Lasdon and Waren [6], [15], Ishizaki and Watanabe [16], Osborne and Watson [24], and the least pth approximation method of Temes and Zai [17], [18].

Extensions of the adjoint network method to second-order network sensitivities have been presented [25]–[27]. The results may be used with those optimization methods, such as the Newton method [2], which require second derivatives. However, since gradient methods involving only first derivatives are generally considered superior, it seems unlikely that widespread application of these results will be seen in the very near future.

VIII. Conclusions

The ease of implementation of the adjoint network method of evaluating partial derivatives and the immediate savings in computation time for the computer-aided design of circuits make it very attractive. A great deal of the uncertainty and inefficiency inherent in the numerical estimation of partial derivatives can be eliminated.

It is believed that this approach will find very wide application. Another very recent report in this area and of interest to microwave engineers is available [28]. There seems little doubt, from the circuit designer's point of view at any rate, that the introduction of the adjoint network method by Director and Rohrer is a turning point in computer-aided design.

IX. Acknowledgment

The authors would like to thank Dr. M. Sablatash of the Department of Electrical Engineering, University of Toronto, Ont., Canada, for his stimulation of this work, A. Lee-Chan of the McMaster Data Processing and Computing Centre, McMaster University for his programming assistance, and W. J. Butler, of General Electric Company, Schenectady, N. Y., for useful discussions. R. E. Seviora received financial assistance under a Mary H. Beatty Fellowship from the University of Toronto.

References

[1] M. A. Murray-Lasso and E. B. Kozemchak, "Microwave circuit design by digital computer," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 514–526, August 1969.
[2] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 533–552, August 1969.
[3] ——, "Computer optimization of inhomogeneous waveguide transformers." *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 563–571, August 1969.
[4] J. W. Bandler and P. A. Macdonald, "Optimization of microwave networks by razor search," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 552–562, August 1969.
[5] T. N. Trick and J. Vlach, "Computer-aided design of broad-band amplifiers with complex loads," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 541–547, September 1970.
[6] L. S. Lasdon and A. D. Waren, "Optimal design of filters with bounded, lossy elements," *IEEE Trans. Circuit Theory*, vol. CT-13, pp. 175–187, June 1966.
[7] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 318–323, August 1969.
[8] ——, "Automated network design—the frequency-domain case," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 330–337, August 1969.
[9] B. D. H. Tellegen, "A general network theorem, with applications," *Philips Res. Reps.*, vol. 7, pp. 259–269, August 1952.
[10] C. A. Desoer and E. S. Kuh, *Basic Circuit Theory*. New York: McGraw-Hill, 1969, ch. 9.
[11] P. Penfield, Jr., R. Spence, and S. Duinker, *Tellegen's Theorem and Electrical Networks*. Cambridge, Mass.: MIT Press, 1970.
[12] J. W. Bandler and R. E. Seviora, "Computation of sensitivities for noncommensurate networks" (to be published).
[13] E. M. T. Jones and J. T. Bolljahn, "Coupled-strip-transmission-line filters and directional couplers," *IRE Trans. Microwave Theory Tech.*, vol. MTT-4, pp. 75–81, April 1956.
[14] R. J. Wenzel, "Theoretical and practical applications of capacitance matrix transformations to TEM network design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-14, pp. 635–647, December 1966.
[15] A. D. Waren, L. S. Lasdon, and D. F. Suchman, "Optimization in engineering design," *Proc. IEEE*, vol. 55, pp. 1885–1897, November 1967.
[16] Y. Ishizaki and H. Watanabe, "An iterative Chebyshev approximation method for network design," *IEEE Trans. Circuit Theory*, vol. CT-15, pp. 326–336, December 1968.
[17] G. C. Temes and D. Y. F. Zai, "Least *p*th approximation," *IEEE Trans. Circuit Theory* (Correspondence), vol. CT-16, pp. 235–237, May 1969.
[18] G. C. Temes, "Optimization methods in circuit design," in *Computer Oriented Circuit Design*, F. F. Kuo and W. G. Magnuson, Jr., Eds. Englewood Cliffs, N. J.: Prentice-Hall, 1969.
[19] S. W. Director, "Network design by mathematical optimization," *WESCON*, San Francisco, Calif., August 1969.
[20] R. Seviora, M. Sablatash, and J. W. Bandler, "Least *p*th and minimax objectives for automated network design," *Electron. Lett.*, vol. 6, pp. 14–15, January 8, 1970.
[21] H. J. Carlin and O. P. Gupta, "Computer design of filters with lumped-distributed elements or frequency variable terminations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 598–604, August 1969.
[22] R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," *Comput. J.*, vol. 6, pp. 163–168, June 1963.
[23] J. W. Bandler, "Automatic optimization of engineering designs —possibilities and pitfalls," *NEREM Rec.*, pp. 26–27, November 1969.
[24] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the nonlinear case," *Comput. J.*, vol. 12, pp. 63–68, February 1969.
[25] S. W. Director, "Increased efficiency of network sensitivity computations by means of LU factorization," presented at the Midwest Symposium on Circuit Theory, Austin, Tex., April 1969.
[26] P. J. Goddard and R. Spence, "Efficient method for the calculation of first- and second-order network sensitivities," *Electron. Lett*, vol. 5, pp. 351–352, August 7, 1969.
[27] G. A. Richards, "Second-derivative sensitivity using the concept of the adjoint network," *Electron. Lett.*, vol. 5, pp. 398–399, August 21, 1969.
[28] M. E. Mokari-Bolhassan and T. N. Trick, "Design of microwave integrated amplifiers," presented at the Midwest Symposium on Circuit Theory, Minneapolis, Minn., May 1970.

Correspondence

Ferrite Microstrip Phase Shifters with Transverse and Longitudinal Magnetization

***Abstract*—Phase shifts of opposing sign are produced in a linear section of microstrip by transverse and longitudinal magnetization of the ferrite substrate. Nonreciprocal phase shift is also produced by the transverse magnetization. Theoretical calculations of phase shift that account for both the diamagnetic effects and the tensor properties of the ferrite permeability agree well with properly constructed experimental measurements. These measurements use closed magnetic circuits to remove the nonuniform demagnetization effects.**

A lightweight reciprocal phase shifter has been constructed that utilizes both transverse and longitudinal magnetization at low drive power with closed magnetic circuits to obtain a high figure of merit.

Introduction

Experimentally observed properties of ferrite microstrip phase shifters have been described in a number of published papers [1], [2]. Approximate theoretical analyses have also been presented that give a partial explanation of some of the properties observed. Refinements in both the theories presented and the experimental measurements performed are required before reasonable correlation can be obtained between them. Some of these refinements are described in this correspondence. The discussion is limited to ferrite phase shifters that are composed of linear sections of microstrip. The closely coupled meander-line configurations are not considered.

Manuscript received April 13, 1970; revised July 17, 1970. This paper was presented at the 1970 International Microwave Symposium, Newport Beach, Calif., May 11–14.

Longitudinal and Transverse Magnetization

Previous investigations have been concentrated on the effects of longitudinal magnetization (parallel to the microstrip) on the phase-shifting properties of linear sections of microstrip. The effect of a transverse magnetization in the plane of the substrate on the phase velocity of the propagating fields has been assumed to be nil. As shown by a comparison of Figs. 1 and 2, the application of an external transverse magnetic field to the ferrite substrate generates a phase shift of roughly the same size as that introduced by longitudinal fields. The phase shift induced by transverse magnetization has both a reciprocal and a nonreciprocal component. The composite phase shift induced by transverse magnetization is in the direction opposite to the phase shift generated by the longitudinal magnetization. Some nonreciprocal phase shift has been previously reported [3] for widely spaced meander lines, but the effect was attributed to coupling between the lines. The presence of the nonreciprocal phase shift in a single linear section of microstrip is evidence of the fact that the