PROCEEDINGS

EIGHTH ANNUAL ALLERTON CONFERENCE

ON CIRCUIT AND SYSTEM THEORY

A. H. Haddad
B. C. Kuo
G. Metze

Conference Co-Chairmen

Conference held
October 7, 8, 9, 1970
Allerton House
Monticello, Illinois

Sponsored by the

DEPARTMENT OF ELECTRICAL ENGINEERING
and the
COORDINATED SCIENCE LABORATORY
of the
UNIVERSITY OF ILLINOIS
Urbana, Illinois 61801

and the

CIRCUIT THEORY AND AUTOMATIC CONTROL GROUPS
OF THE
INSTITUTE OF ELECTRICAL & ELECTRONICS ENGINEERS

SENSITIVITIES IN TERMS OF WAVE VARIABLES

JOHN W. BANDLER
Dept. of Electrical Engineering
McMaster University
Hamilton, Ontario, Canada

RUDOLPH E. SEVIORA
Dept. of Electrical Engineering
University of Toronto
Toronto, Ontario, Canada

Abstract It is the purpose of this paper to show how first- and second-order sensitivities and gradients with respect to network parameters can be evaluated directly in terms of wave variables. The concept of the adjoint network is employed. The application of these ideas to computer-aided network optimization is indicated.

INTRODUCTION

There is considerable interest and activity currently in the computation of sensitivities for lumped [1], [2] and distributed [3], [4] networks using the adjoint network approach in terms of currents and voltages. These are often needed to evaluate parameter space gradients in computer-aided network optimization [1-5]. An advantage of the adjoint network approach is that at most two network analyses are required for evaluation of the gradient vector regardless of the number of parameters. In certain problems the need may arise for second-order sensitivities and methods for their evaluation have been suggested [6-8].

It is inconvenient if not impossible to work with currents and voltages for certain classes of networks. In the microwave region, for example, a wave description of networks is often preferable. There is need, then, for sensitivities or gradients of networks in terms of wave variables. It is the purpose of this paper to show how they may be obtained and to indicate practical applications.

AN IDENTITY FOR SCATTERING VARIABLES

Let us assume we have a network which is composed of (in general) one-port and multiport elements. Let the normalized incident and reflected waves at every port be denoted a and b, respectively (see Figure 1(a)). Consider now a second network of the same topology and corresponding normalizations with corresponding variables denoted α and β, respectively (see Figure 1(b)). In the ensuing discussion we will need the following identity relating the waves in both networks

$$\sum_{i\varepsilon I} (b_i\alpha_i - a_i\beta_i) = \sum_{i\varepsilon E} (b_i\alpha_i - a_i\beta_i) \qquad (1)$$

where I is an index set relating to all (interior) ports and E an index set relating to exterior ports. Since an exterior port is at the same time a port of a network element, E is a subset of I. That the above equation is true is more or less obvious in the case when the connected port pairs are normalized to the same level. However, (1) is valid even when the normalization levels are arbitrary throughout the network. This can be seen, for example, by writing first the scattering equation for an internal port junction (Figures 1(a) and (b))

This work was carried out under Grants A7239 and A5277 from the National Research Council of Canada.

$$\begin{bmatrix} a_p \\ a_q \end{bmatrix} = \underline{S}_j \begin{bmatrix} b_p \\ b_q \end{bmatrix} \tag{2a}$$

$$\begin{bmatrix} \alpha_p \\ \alpha_q \end{bmatrix} = \underline{S}_j \begin{bmatrix} \beta_p \\ \beta_q \end{bmatrix} \tag{2b}$$

where $\underline{S}_j$ is the scattering matrix of the junction, accounting for the change in normalization across the junction. Substituting (2a) and (2b) into the relevant terms of the left-hand side of (1) we have

$$\begin{bmatrix} b_p & b_q \end{bmatrix} \begin{bmatrix} \alpha_p \\ \alpha_q \end{bmatrix} - \begin{bmatrix} a_p & a_q \end{bmatrix} \begin{bmatrix} \beta_p \\ \beta_q \end{bmatrix}$$

$$= \begin{bmatrix} b_p & b_q \end{bmatrix} (\underline{S}_j - \underline{S}_j^T) \begin{bmatrix} \beta_p \\ \beta_q \end{bmatrix} . \tag{3}$$

Since the junction (normalization step) is reciprocal (3) is equal to zero, thus the left-hand side of (1) reduces to the sum over E. For an alternative proof using voltage and current concepts see Penfield et. al. [9].

Derivations in the following sections are based on a generalized version of (1) [9], namely

$$\sum_{i\varepsilon I} ((L\ b_i)\alpha_i - (L\ a_i)\beta_i) = \sum_{i\varepsilon E} ((L\ b_i)\alpha_i - (L\ a_i)\beta_i) \tag{4}$$

where L will be first-and second-order parameter space differential operators. Proof of (4) follows similar lines to the proof of (1).

FIRST-ORDER SENSITIVITIES

Suppose a parameter ϕ in the original network is varied without affecting topology. Further, let it be contained in an n-port subnetwork characterized by a scattering matrix $\underline{S}$ such that

$$\underline{b} = \underline{S}\underline{a} \tag{5}$$

where $\underline{b}$ and $\underline{a}$ are the appropriate n-element vectors. Then

$$\frac{\partial \underline{b}}{\partial \phi} = \frac{\partial \underline{S}}{\partial \phi} \underline{a} + \underline{S} \frac{\partial \underline{a}}{\partial \phi} . \tag{6}$$

The terms relating to this subnetwork on the left-hand side of (4) become

$$\underline{a}^T (\frac{\partial \underline{S}}{\partial \phi})^T \underline{\alpha} + (\frac{\partial \underline{a}}{\partial \phi})^T \left[\underline{S}^T \underline{\alpha} - \underline{\beta} \right] \tag{7}$$

which reduces to

$$\underline{a}^T\left(\frac{\partial \underline{S}}{\partial\phi}\right)^T\underline{\alpha} \tag{8}$$

if

$$\underline{\beta} = \underline{S}^T\ \underline{\alpha}. \tag{9}$$

Equation (9) defines the corresponding subnetwork of the auxiliary or, in current parlance, the adjoint network.

If all elements of the adjoint network are defined in this way

$$\sum_{i\varepsilon E}\left(\frac{\partial b_i}{\partial\phi}\ \alpha_i - \frac{\partial a_i}{\partial\phi}\ \beta_i\right) = \underline{a}^T\left(\frac{\partial \underline{S}}{\partial\phi}\right)^T\ \underline{\alpha}\ . \tag{10}$$

Then, by suitably terminating (and exciting) the network and its adjoint, the sensitivity of the desired wave can be computed. So, for example, if we are interested in the sensitivity of the ikth term of the overall scattering matrix, the network is analyzed for $a_k=1$, $a_j=0$ for all $j\ \varepsilon\ E$, $j \neq k$ (the latter is achieved by matched terminations), and the adjoint network excitations are zero except at the ith port, where a unit source is applied. Then (10) reduces to*

$$\frac{\partial S_{ik}}{\partial\phi} = \frac{\partial b_i}{\partial\phi} = \underline{a}^T\left(\frac{\partial \underline{S}}{\partial\phi}\right)^T\ \underline{\alpha}\ . \tag{11}$$

To evaluate the sensitivities one needs to know the partial derivatives of subnetwork scattering matrices. For the majority of common design components these can usually be found after some manipulation. As it happens, quite often $\partial\underline{S}/\partial\phi$ can be expressed in terms of the original $\underline{S}$ matrix which simplifies the evaluation of the sensitivities. Thus, for example, for a piece of waveguide of characteristic impedance Z using real normalization

$$\underline{a}^T\left(\frac{\partial \underline{S}}{\partial Z}\right)^T\ \underline{\alpha} = \frac{1}{2Z}\left(\underline{a}^T\ \underline{\alpha} - \underline{b}^T\ \underline{\beta}\right)\ . \tag{12}$$

Or, the arbitrariness in the normalization level can be used to advantage as demonstrated in the case of the sensitivity with respect to length of a piece of transmission line or waveguide. If the input and output waves are normalized on the impedance level equal to the characteristic impedance of the guide, then

$$\underline{a}^T\ \left(\frac{\partial \underline{S}}{\partial \ell}\right)^T\ \underline{\alpha} = -\gamma\ \underline{a}^T\underline{\beta} \tag{13}$$

where γ is the propagation constant of the waveguide.

---

*At the time of writing, it was brought to the authors' attention that some similar results have been independently derived [10].

SECOND-ORDER SENSITIVITIES

Consider two distinct non-topological parameters ϕ and ψ in the original network. Assume that they are contained in a subnetwork with scattering matrix $\underline{S}$. Applying $\partial/\partial\psi$ to (6) we get

$$\frac{\partial^2 \underline{b}}{\partial\psi\,\partial\phi} = \frac{\partial^2 \underline{S}}{\partial\psi\,\partial\phi}\,\underline{a} + \frac{\partial \underline{S}}{\partial\phi}\frac{\partial \underline{a}}{\partial\psi} + \frac{\partial \underline{S}}{\partial\psi}\frac{\partial \underline{a}}{\partial\phi} + \underline{S}\,\frac{\partial^2 \underline{a}}{\partial\psi\,\partial\phi}\,. \tag{14}$$

The terms relating to this subnetwork on the left-hand side of (4) now appropriately become

$$\underline{\alpha}^T \left(\frac{\partial^2 \underline{b}}{\partial\psi\,\partial\phi}\right) - \underline{\beta}^T \left(\frac{\partial^2 \underline{a}}{\partial\psi\,\partial\phi}\right)$$

$$= \underline{\alpha}^T \left(\frac{\partial^2 \underline{S}}{\partial\psi\,\partial\phi}\right)\underline{a} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\phi}\frac{\partial \underline{a}}{\partial\psi} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\psi}\frac{\partial \underline{a}}{\partial\phi} + \left(\underline{\alpha}^T\underline{S} - \underline{\beta}^T\right)\frac{\partial^2 \underline{a}}{\partial\psi\,\partial\phi}\,. \tag{15}$$

Using (9), we can reduce (15) to

$$\underline{\alpha}^T \left(\frac{\partial^2 \underline{S}}{\partial\psi\,\partial\phi}\right)\underline{a} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\phi}\frac{\partial \underline{a}}{\partial\psi} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\psi}\frac{\partial \underline{a}}{\partial\phi}\,. \tag{16}$$

If the whole auxiliary network is defined in accordance with (9) and if suitable terminations as described in the preceding section are applied, we obtain the counterpart of (11), namely,

$$\frac{\partial^2 S_{ik}}{\partial\psi\,\partial\phi} = \underline{\alpha}^T \left(\frac{\partial^2 \underline{S}}{\partial\psi\,\partial\phi}\right)\underline{a} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\phi}\frac{\partial \underline{a}}{\partial\psi} + \underline{\alpha}^T \frac{\partial \underline{S}}{\partial\psi}\frac{\partial \underline{a}}{\partial\phi}\,. \tag{17}$$

A special case of interest is when ψ and ϕ are identical leading to

$$\frac{\partial^2 S_{ik}}{\partial\phi^2} = \underline{\alpha}^T \left(\frac{\partial^2 \underline{S}}{\partial\phi^2}\right)\underline{a} + 2\underline{\alpha}^T \frac{\partial \underline{S}}{\partial\phi}\frac{\partial \underline{a}}{\partial\phi}\,. \tag{18}$$

If ϕ belongs to a subnetwork with a scattering matrix $\underline{S}_\phi$ and ψ belongs to a subnetwork with a scattering matrix $\underline{S}_\psi$, it may readily be shown that

$$\frac{\partial^2 S_{ik}}{\partial\psi\,\partial\phi} = \underline{\alpha}_\phi^T \frac{\partial \underline{S}_\phi}{\partial\phi}\frac{\partial \underline{a}_\phi}{\partial\psi} + \underline{\alpha}_\psi^T \frac{\partial \underline{S}_\psi}{\partial\psi}\frac{\partial \underline{a}_\psi}{\partial\phi} \tag{19}$$

where subscripts ϕ and ψ denote quantities related to the appropriate subnetworks.

COMPUTATION OF PARTIAL DERIVATIVES

The results of the previous sections can be readily exploited in the evaluation of the gradients of practically any wave-based objective function with respect to a variety of parameters including frequency. The procedure is fairly well-known and is discussed in detail in a number of references [2-6], [11].

EXAMPLES

Convenient examples to test the ideas of this paper were provided by the resistively terminated cascade of transmission lines shown in Figure 2. Two-section and three-section 10:1 quarter-wave transformers optimum over 100% bandwidth were chosen. Their responses are also shown in Figure 2. The numerical values of the parameters of the two-section design were taken from reference [12] and of the three-section design from reference [13].

Tables I and II show the components of the gradient vector of the magnitude of the reflection coefficient ρ at 0.5 GHz estimated from 1% and .01% incremental changes in the parameters compared with those obtained from the results of this paper using one network analysis. Gradient calculations at other frequencies were also made by the present method. They were used by the first author in illustrating the necessary conditions for a minimax optimum in another paper presented at this conference [14].

CONCLUSIONS

This paper has shown how first- and second-order sensitivities and gradients with respect to network parameters can be evaluated directly in terms of wave variables, without recourse to voltages and currents. The method employs the concept of the adjoint network. As a result, the same benefits in terms of ease of implementation and computational efficiency as discussed by previous authors should be enjoyed. Efficient gradient methods of minimization can thus be employed in the optimal design by computer of networks for which the wave description is more natural or preferable.

ACKNOWLEDGEMENT

The authors would like to thank Dr. M. Sablatash of the Department of Electrical Engineering, University of Toronto, for his cooperation. The authors would also like to thank A. Lee-Chan of the McMaster Data Processing and Computing Centre for checking the numerical results given in this paper.

REFERENCES

[1] 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.

[2] S.W. Director and R.A. Rohrer, "Automated network design - the frequency-domain case", IEEE Trans. Circuit Theory, vol. CT-16, pp. 330-337, August 1969.

[3] J.W. Bandler and R.E. Seviora, "Current trends in network optimization", IEEE Trans. Microwave Theory and Techniques, vol. MTT-18, December 1970.

[4] J.W. Bandler and R.E. Seviora, "Computation of sensitivities for noncommensurate networks", IEEE Trans. Circuit Theory, vol. CT-18, January 1971.

[5] R. Seviora, M. Sablatash and J.W. Bandler, "Least pth and minimax objectives for automated network design", Electronics Letters, vol. 6, pp. 14-15, January 1970.

[6] S.W. Director, "Increased efficiency of network sensitivity computations by means of LU factorization", Midwest Symposium on Circuit Theory, Austin, Texas, April 1969.

[7] P.J. Goddard and R. Spence, "Efficient method for the calculation of first- and second-order network sensitivities", Electronics Letters, vol. 5, pp. 351-352, August 1969.

[8] G.A. Richards, "Second-derivative sensitivity using the concept of the adjoint network", Electronics Letters, vol. 5, pp. 398-399, August 1969.

[9] P. Penfield, Jr., R. Spence and S. Duinker, "A generalized form of Tellegen's theorem", IEEE Trans. Circuit Theory, vol. CT-17, pp. 302-305, August 1970.

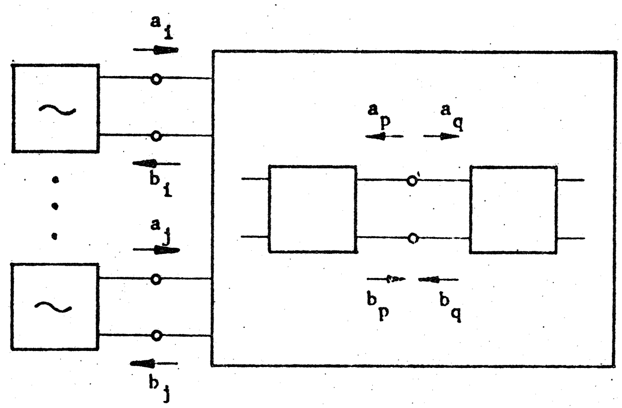
[10] V.A. Monaco and P. Tiberio, "On linear network scattering matrix sensitivity", Alta Frequenza, vol. 39, pp. 193-195, February 1970.

[11] G.C. Temes, "Exact computation of group delay and its sensitivities using adjoint-network concept", Electronics Letters, vol. 6, pp. 483-485, July 1970.

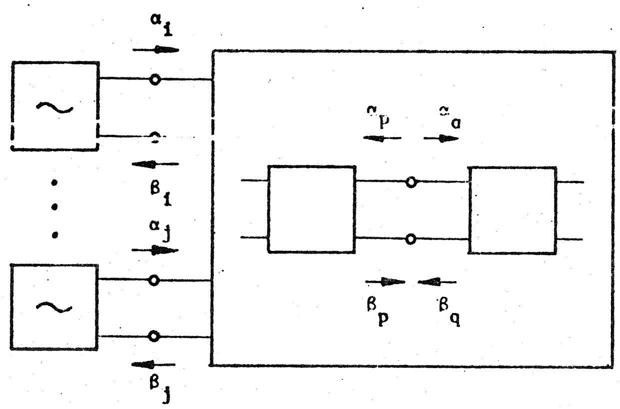
[12] J.W. Bandler and P.A. Macdonald, "Cascaded noncommensurate transmission-line networks as optimization problems", IEEE Trans. Circuit Theory, vol. CT-16, pp. 391-394, August 1969.

[13] J.W. Bandler and P.A. Macdonald, "Optimization of microwave networks by razor search", IEEE Trans. Microwave Theory and Techniques, vol. MTT-17, pp. 552-562, August 1969.

[14] J.W. Bandler, "Conditions for a minimax optimum", 8th Annual Allerton Conference on Circuit and System Theory, Urbana, Illinois, October 1970.



(a)



(b)

Figure 1. Two networks of the same topology and with corresponding port normalizations.

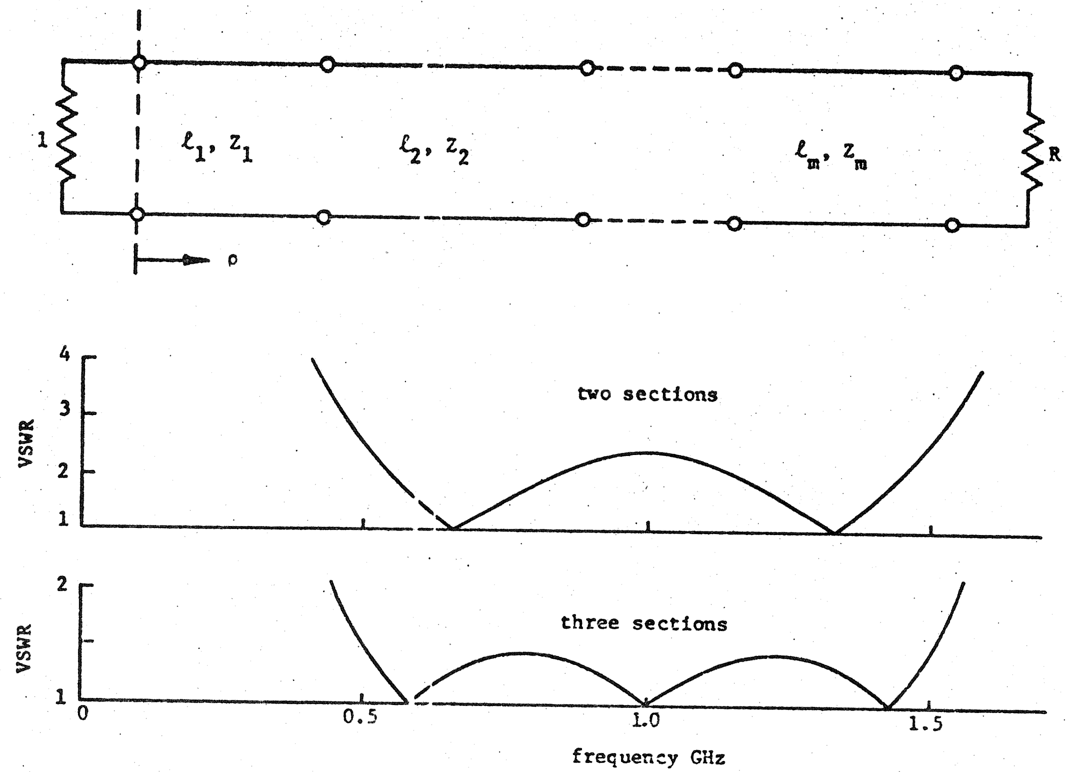


Figure 2. m-section resistively terminated cascade of transmission lines. Also shown are the responses of 100% bandwidth optimum two- and three-section quarter-wave transformers for R=10.

TABLE I

COMPARISON OF GRADIENT COMPONENTS OF THE REFLECTION COEFFICIENT OF THE 2-SECTION TRANSFORMER AT 0.5 GHz ($|\rho| = 0.4286$)

Parameter values	Gradient components		
	1% increment	.01% increment	adjoint network
$\ell_1 = 7.49482$ cm	-7.4397×10^{-2}	-7.3337×10^{-2}	-7.3326×10^{-2}
$Z_1 = 2.2361\ \Omega$	-1.8250×10^{-1}	-1.8254×10^{-1}	-1.8254×10^{-1}
$\ell_2 = 7.49482$ cm	-7.3745×10^{-2}	-7.3330×10^{-2}	-7.3326×10^{-2}
$Z_2 = 4.4721\ \Omega$	9.0050×10^{-2}	9.1260×10^{-2}	9.1272×10^{-2}

TABLE II

COMPARISON OF GRADIENT COMPONENTS OF THE REFLECTION COEFFICIENT OF THE 3-SECTION TRANSFORMER AT 0.5 GHz ($|\rho| = 0.1973$)

Parameter values	Gradient components		
	1% increment	.01% increment	adjoint network
$\ell_1 = 7.49482$ cm	-4.4498×10^{-2}	-4.3777×10^{-2}	-4.3770×10^{-2}
$Z_1 = 1.63471\ \Omega$	-4.3461×10^{-1}	-4.3555×10^{-1}	-4.3556×10^{-1}
$\ell_2 = 7.49482$ cm	-9.1695×10^{-2}	-9.1294×10^{-2}	-9.1289×10^{-2}
$Z_2 = 3.16228\ \Omega$	-6.7×10^{-4}	-6.5×10^{-6}	4.0×10^{-7}
$\ell_3 = 7.49482$ cm	-4.3545×10^{-2}	-4.3767×10^{-2}	-4.3770×10^{-2}
$Z_3 = 6.11729\ \Omega$	1.1543×10^{-1}	1.1638×10^{-1}	1.1639×10^{-1}