$(-\infty, -\beta_b]$ in which β_b is a positive constant, $f(\cdot)$ is continuous on $(-\alpha_b, \infty)$, $f(\alpha)\to -\beta_b$ as $\alpha \to -\alpha_b$ from the right, $f(\alpha_1) > f(\alpha_2)$ for all α_1 and α_2 such that $\alpha_1 > \alpha_2$, $\alpha_1 > -\alpha_b$, and $\alpha_2 > -\alpha_b$, and $\alpha^{-1} f(\alpha) \to \infty$ as $\alpha \to \infty$. With the understanding that each of the functions $f_j(\cdot)$ is of the multiple-valued type just described, or of the multiple-valued type obtained by similarly modifying the function $a_j[1-\exp(-b_j x_j)]$, or of the single-valued type described two paragraphs above, the equation $F(x)+Ax=B$ can be interpreted in the following manner: $x \in E^n$ satisfies $F(x)+Ax=B$ if and only if for all j the jth component x_j of x is contained in the domain of definition of $f_j(\cdot)$, and $y+Ax=B$ for one of the values y taken on by $F(x)$. With this interpretation of $F(x)+Ax=B$, it is not a difficult matter to modify the proof given in Section II to show that there exists a solution $x \in E^n$ for each $B \in E^n$ and *any* real $n \times n$ matrix A. We leave the details of the necessary modifications of the proof to the sufficiently interested reader.

II. The Theorem and its Proof

Throughout Section II, n is an arbitrary positive integer; E^n denotes the set of all real n vectors; v_j denotes the jth component of v for all $j \in \{1, 2, \cdots, n\}$ and each $v \in E^n$; $\|v\|$ denotes $(\sum_{j=1}^{n} v_j^2)^{\frac{1}{2}}$ for all $v \in E^n$; and for all row vectors $a=(a_1, a_2, \cdots, a_n)$, a^{tr} denotes the transpose of a.

Definition

Let $\mathscr{F}$ denote the set of all mappings $F(\cdot)$ of E^n onto E^n such that $F(x)=(f_1(x_1), f_2(x_2), \cdots, f_n(x_n))^{tr}$ for all $x \in E^n$ in which for each $j=1,2,\cdots,n$, $f_j(\cdot)$ is a continuous mapping of E^1 onto itself which satisfies the following.

1) $|\alpha^{-1} f_j(\alpha)| \to \infty$ as $|\alpha| \to \infty$.
2) For some positive constant β_j, $f_j(\alpha_1) > f_j(\alpha_2)$ for all real numbers α_1 and α_2 such that $\alpha_1 > \alpha_2$, $|\alpha_1| \geq \beta_j$, and $|\alpha_2| \geq \beta_j$.

Theorem

Let $F(\cdot) \in \mathscr{F}$ and let $A(\cdot)$ be a continuous mapping of E^n into E^n such that there exist positive constants a and b with the property that $\|A(z)\| \leq a\|z\| + b$ for all $z \in E^n$. Then for each $y \in E^n$, there exists an $x \in E^n$ such that $F(x)+A(x)=y$.

Proof of the Theorem: We shall use the following lemma which is a special case of a result proved in [4]; the result is closely related to a theorem of Dubrovskii (see [4]).

Lemma

If $M(\cdot)$ is a continuous mapping of E^n into E^n such that there exist real constants $c \in (0, 1)$ and $k > 0$ with the property that $\|M(z)\| \leq c\|z\|$ for all $z \in \{z : z \in E^n, \|z\| \geq k\}$, then for each $y \in E^n$ there exists an $x \in E^n$ such that $x + M(x) = y$.

Let $\beta = \max_j \{\beta_j\}$ [see condition 2) of the Definition], and note that $f_j(\beta) > f_j(-\beta)$ for all j. For each j let $l_j(\cdot)$ denote the mapping of E^1 into E^1 defined by

$$\begin{aligned} l_j(x_j) &= (2\beta)^{-1}[f_j(\beta) - f_j(-\beta)]x_j + 2^{-1}[f_j(\beta) + f_j(-\beta)], \\ &\qquad \text{for } x_j \in [-\beta, \beta] \\ &= 0, \qquad \text{for } |x_j| > \beta \end{aligned}$$

and for each j, let $g_j(\cdot)$ denote the strictly monotone-increasing mapping of E^1 onto E^1 defined by

$$\begin{aligned} g_j(x_j) &= l_j(x_j), &&\text{for all } x_j \in [-\beta, \beta] \\ &= f_j(x_j), &&\text{for all } |x_j| > \beta. \end{aligned}$$

Then, with $G(\cdot)$ such that $G(x)=(g_1(x_1), g_2(x_2), \cdots, g_n(x_n))^{tr}$ for all $x \in E^n$, we have

$$F(x) = G(x) + \Delta(x), \qquad \text{for all } x \in E^n$$

in which

$$\begin{aligned} \Delta(x) &= (\delta_1(x_1), \delta_2(x_2), \cdots, \delta_n(x_n))^{tr} \\ \delta_j(x_j) &= f_j(x_j) - l_j(x_j), &&\text{for all } x_j \in [-\beta, \beta] \\ &= 0, &&\text{for all } |x_j| > \beta. \end{aligned}$$

We observe that $G^{-1}(\cdot)$ exists and is continuous on E^n, that $|\alpha^{-1} g_j(\alpha)| \to \infty$ as $|\alpha| \to \infty$ for all j, and that $\Delta(\cdot)$ is a continuous mapping of E^n into E^n such that for some positive constant b', $\|\Delta(x)\| \leq b'$ for all $x \in E^n$. Thus for any $y \in E^n$, the equation $F(x)+A(x)=y$ possesses a solution $x \in E^n$ if and only if there is a solution $z \in E^n$ of $z + \Delta[G^{-1}(z)] + A[G^{-1}(z)] = y$. The operator $\Delta[G^{-1}(\cdot)] + A[G^{-1}(\cdot)]$ is a continuous mapping of E^n into E^n which satisfies

$$\begin{aligned} \|\Delta[G^{-1}(x)] + A[G^{-1}(x)]\| &\leq \|\Delta[G^{-1}(x)]\| + \|A[G^{-1}(x)]\| \\ &\leq (b + b') + a\|G^{-1}(x)\| \end{aligned} \tag{3}$$

for all $x \in E^n$. Since $|\alpha^{-1} g_j(\alpha)| \to \infty$ as $|\alpha| \to \infty$ for all j, we have

$$\frac{\|G^{-1}(x)\|}{\|x\|} \to 0, \qquad \text{as } \|x\| \to \infty.$$

In particular, there exists a positive constant k such that $k \geq 4(b+b')$ and $\|G^{-1}(x)\| \leq \frac{1}{4}a^{-1}\|x\|$ for all $\|x\| \geq k$. Therefore, using (3),

$$\begin{aligned} \|\Delta[G^{-1}(x)] + A[G^{-1}(x)]\| &\leq \tfrac{1}{4}k + \tfrac{1}{4}\|x\| \\ &\leq \tfrac{1}{2}\|x\| \end{aligned}$$

for all $\|x\| \geq k$. Hence, by the Lemma, there exists a solution $z \in E^n$ of $z + \Delta[G^{-1}(z)] + A[G^{-1}(z)] = y$.

Remark

While it can be shown that the proof given above is by no means the shortest possible proof (see [5, Lemma 2.1] which can in fact be used to prove a more general result), it has the advantage that it can easily be modified to treat the important case described in Section I in which the $f_j(\cdot)$ are multiple-valued functions defined on a subset of E^1.

I. W. Sandberg
Bell Telephone Labs., Inc.
Murray Hill, N. J. 07974

References

[1] I. W. Sandberg and A. N. Willson, Jr., "Some theorems on properties of dc equations of nonlinear networks," *Bell Syst. Tech. J.*, vol. 48, Jan. 1969, pp. 1–34.
[2] I. W. Sandberg, "Theorems on the analysis of nonlinear transistor networks," *Bell Syst. Tech. J.*, vol. 49, Jan. 1970, pp. 95–114.
[3] I. W. Sandberg and A. N. Willson, Jr., "Some network-theoretic properties of nonlinear dc transistor networks," *Bell Syst. Tech. J.*, vol. 48, May–June 1969, pp. 1293–1312.
[4] V. E. Benes and I. W. Sandberg, "Applications of a theorem of Dubrovskii to the periodic responses of nonlinear systems," *Bell Syst. Tech. J.*, vol. 43, Nov. 1964, pp. 2855–2872.
[5] R. I. Kachurovskii, "Non-linear monotone operators in Banach space," *Russ. Math. Surv.*, vol. 23, Mar.–Apr. 1968.

Conditions for a Minimax Optimum

Abstract—This paper derives and discusses necessary conditions for an optimum in nonlinear minimax approximation problems. A straightforward geometrical interpretation is presented. The results may be used to test for convergence in computer-aided network optimization, in tests for optimality in the Chebyshev sense of any given design, and to gain insight which may be helpful in developing minimax approximation algorithms.

Introduction

Of considerable importance to network designers, whether they use synthesis or optimization methods, is the optimality of their design. Conditions for optimality in the Chebyshev sense in conventional synthesis problems involving polynomials and rational functions are fairly widely appreciated. However, with the ever-increasing need for designs containing elements not conducive to the rational function approach, e.g., a mixture of lumped and distributed elements, and the inevitable application of

Manuscript received October 28, 1970; revised January 4, 1971. This work was supported by the Science and Engineering Division Research Board, McMaster University, Hamilton, Ont., Canada. This paper was presented at the 8th Annual Allerton Conference on Circuit and System Theory, Urbana, Ill., October 7–9, 1970.

automatic optimization methods involving least pth [1]–[3] and minimax [4]–[10] objectives, some means of testing for convergence to an optimum for more arbitrary problems is highly desirable.

In this paper the necessary conditions for a minimax optimum for a general nonlinear minimax approximation problem are obtained from the Kuhn–Tucker necessary conditions for a constrained optimum in nonlinear programming [11]. Just as the Kuhn–Tucker relations have a straightforward geometrical interpretation, so the conditions for a minimax optimum also have such an interpretation. This interpretation is presented and discussed. The restrictions for which the necessary conditions are sufficient are also discussed.

A comparison with theorems characterizing minimax approximations presented by Curtis and Powell [12] is made. Finally, illustrative examples demonstrate the implementation of the necessary conditions.

The Problem

Let us define errors related to the "upper" and "lower" specifications, respectively, as [8]

$$\begin{aligned} e_u(\boldsymbol{\phi}, \psi) &\triangleq w_u(\psi)(F(\boldsymbol{\phi}, \psi) - S_u(\psi)) \\ e_l(\boldsymbol{\phi}, \psi) &\triangleq w_l(\psi)(F(\boldsymbol{\phi}, \psi) - S_l(\psi)) \end{aligned} \tag{1}$$

where

- $F(\boldsymbol{\phi}, \psi)$ approximating function (network response);
- $S_u(\psi)$ upper specified function (desired response bound);
- $S_l(\psi)$ lower specified function (desired response bound);
- $w_u(\psi)$ upper weighting function;
- $w_l(\psi)$ lower weighting function;
- $\boldsymbol{\phi}$ k independent parameters;
- ψ independent variable (e.g., frequency or time).

We will make a number of assumptions which are usually fulfilled in practice. The functions $e_u(\boldsymbol{\phi}, \psi)$ and $e_l(\boldsymbol{\phi}, \psi)$ are evaluated at a finite discrete set of values of ψ taken from one or more closed intervals over each of which the functions are continuous in ψ. Therefore, let

$$\begin{aligned} e_{ui}(\boldsymbol{\phi}) &\triangleq e_u(\boldsymbol{\phi}, \psi_i), \qquad i \in I_u \\ e_{li}(\boldsymbol{\phi}) &\triangleq e_l(\boldsymbol{\phi}, \psi_i), \qquad i \in I_l \end{aligned} \tag{2}$$

where I_u and I_l are appropriate index sets.

The discrete approximation problem can be written in the form

$$\text{minimize } U = \phi_{k+1} \tag{3}$$

subject to

$$\begin{aligned} \phi_{k+1} - e_{ui}(\boldsymbol{\phi}) \geq 0, &\qquad i \in I_u \\ \phi_{k+1} + e_{li}(\boldsymbol{\phi}) \geq 0, &\qquad i \in I_l. \end{aligned} \tag{4}$$

This is equivalent to the problem of finding

$$\min_{\boldsymbol{\phi}} \{\max_i [e_{ui}(\boldsymbol{\phi}), -e_{li}(\boldsymbol{\phi})]\}. \tag{5}$$

It is assumed that a minimum exists in a closed and bounded region of points $\boldsymbol{\phi}$ and that $e_{ui}(\boldsymbol{\phi})$ and $e_{li}(\boldsymbol{\phi})$ are continuous for all i with continuous partial derivatives at least in the neighborhood of the minimum. It is also assumed that (4) satisfy the constraint qualification.

The Conditions

Theorem 1

At an optimum point $\boldsymbol{\phi}^o$ for the minimax approximation problem

$$\sum_{i \in J_u} u_{ui} \nabla e_{ui}(\boldsymbol{\phi}^o) = \sum_{i \in J_l} u_{li} \nabla e_{li}(\boldsymbol{\phi}^o)$$

$$\sum_{i \in J_u} u_{ui} + \sum_{i \in J_l} u_{li} = 1$$

$$\begin{aligned} u_{ui} \geq 0, &\qquad i \in J_u \\ u_{li} \geq 0, &\qquad i \in J_l \end{aligned}$$

where

$$\nabla \triangleq \left[\frac{\partial}{\partial \phi_1} \frac{\partial}{\partial \phi_2} \cdots \frac{\partial}{\partial \phi_k}\right]^T$$

and where $e_{ui}(\boldsymbol{\phi}^o)$ for $i \in J_u$ and $-e_{li}(\boldsymbol{\phi}^o)$ for $i \in J_l$ are the equal maxima. The special case $S_{ui} = S_{li} = S_i$, $w_{ui} = w_{li} = w_i$, $I_u = I_l = I$ is also accommodated.

Proof: Invoking the Kuhn–Tucker relations, we may write at $\boldsymbol{\phi} = \boldsymbol{\phi}^o$

$$\begin{bmatrix} \nabla U \\ \dfrac{\partial U}{\partial \phi_{k+1}} \end{bmatrix} = \sum_{i \in I_u} u_{ui} \begin{bmatrix} \nabla \\ \dfrac{\partial}{\partial \phi_{k+1}} \end{bmatrix} (\phi_{k+1} - e_{ui}(\boldsymbol{\phi})) + \sum_{i \in I_l} u_{li} \begin{bmatrix} \nabla \\ \dfrac{\partial}{\partial \phi_{k+1}} \end{bmatrix} (\phi_{k+1} + e_{li}(\boldsymbol{\phi})) \tag{6}$$

$$\begin{aligned} u_{ui}(\phi_{k+1} - e_{ui}(\boldsymbol{\phi})) = 0, &\qquad i \in I_u \\ u_{li}(\phi_{k+1} + e_{li}(\boldsymbol{\phi})) = 0, &\qquad i \in I_l \end{aligned} \tag{7}$$

where

$$\begin{aligned} u_{ui} \geq 0, &\qquad i \in I_u \\ u_{li} \geq 0, &\qquad i \in I_l. \end{aligned} \tag{8}$$

At a minimum at least one of the constraints (4) must equal zero or U could be further reduced without violation of the constraints. Let J_u be an index set corresponding to possible n_u constraints related to the upper specifications which are zero and J_l be an index set corresponding to possible n_l constraints related to the lower specifications which are zero. Then

$$\begin{aligned} u_{ui} &\begin{cases} \geq 0, & i \in J_u \\ = 0, & i \notin J_u \end{cases} \\ u_{li} &\begin{cases} \geq 0, & i \in J_l \\ = 0, & i \notin J_l. \end{cases} \end{aligned} \tag{9}$$

Note also that

$$\begin{aligned} \nabla U &= \nabla \phi_{k+1} = \mathbf{0} \\ \frac{\partial U}{\partial \phi_{k+1}} &= 1 \\ \frac{\partial e_{ui}(\boldsymbol{\phi})}{\partial \phi_{k+1}} &= \frac{\partial e_{li}(\boldsymbol{\phi})}{\partial \phi_{k+1}} = 0 \end{aligned} \tag{10}$$

everywhere. The necessary conditions now become

$$\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \sum_{i \in J_u} u_{ui} \begin{bmatrix} -\nabla e_{ui}(\boldsymbol{\phi}^o) \\ 1 \end{bmatrix} + \sum_{i \in J_l} u_{li} \begin{bmatrix} \nabla e_{li}(\boldsymbol{\phi}^o) \\ 1 \end{bmatrix} \tag{11}$$

from which the theorem follows.

Geometrically, the necessary conditions have a straightforward interpretation. To see this, let the problem be written in the form

$$\text{minimize } U = \phi_{k+1} \tag{12}$$

subject to

$$\phi_{k+1} \geq f_i(\boldsymbol{\phi}), \qquad i = 1, 2, \cdots, m. \tag{13}$$

The necessary conditions for an optimum are

$$\sum_{i=1}^{m_0} u_i \nabla f_i(\boldsymbol{\phi}^o) = \mathbf{0} \tag{14}$$

$$\sum_{i=1}^{m_0} u_i = 1 \tag{15}$$

$$u_i \geq 0, \qquad i = 1, 2, \cdots, m_0 \tag{16}$$

where it is assumed, for convenience, that the first m_0 constraints are equalities or that the first m_0 of f_i are equal maxima.

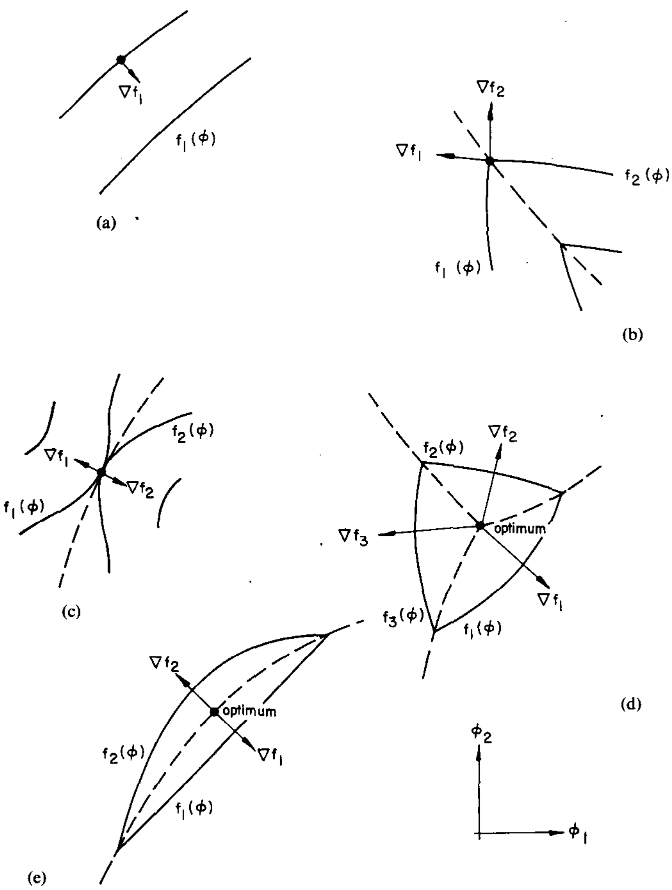

Fig. 1. Geometrical interpretation. (a) and (b): Necessary conditions not satisfied. (c)–(e): Necessary conditions satisfied.

Fig. 1 shows examples for $k=2$. Observe that the conditions are not satisfied in Fig. 1(a) and (b). They are satisfied in Fig. 1(c), (d), and (e). The points indicated in Fig. 1(d) and (e) are clearly optimum points.

Theorem 2

If the relations in Theorem 1 are satisfied at a point $\boldsymbol{\phi}^o$ and the $e_{ui}(\boldsymbol{\phi})$ for $i \in J_u$ are convex and the $e_{li}(\boldsymbol{\phi})$ for $i \in J_l$ are concave, then $\boldsymbol{\phi}^o$ is optimal.

The proof for the above sufficiency theorem would follow similar lines to those required for proving the sufficiency of the Kuhn–Tucker relations.

Now let

$$\boldsymbol{D} \triangleq [\boldsymbol{d}_1\ \boldsymbol{d}_2 \cdots \boldsymbol{d}_r] \tag{17}$$

and

$$\boldsymbol{\rho} \triangleq [\rho_1\ \rho_2 \cdots \rho_r]^T \tag{18}$$

where

$$r = n_u + n_l \tag{19}$$

and where each $\boldsymbol{d}_j$ or ρ_j for $j=1, 2, \cdots, n_u$ corresponds to a $\nabla e_{ui}(\boldsymbol{\phi}^o)$ or u_{ui}, respectively, for $i \in J_u$ and each $\boldsymbol{d}_j$ or ρ_j for $j=n_u+1, n_u+2, \cdots, r$ corresponds to a $\nabla e_{li}(\boldsymbol{\phi}^o)$ or $-u_{li}$, respectively, for $i \in J_l$. Then at an optimum, from Theorem 1

$$\boldsymbol{D\rho} = \boldsymbol{0} \tag{20}$$

is satisfied for at least one $\rho_j \neq 0$.

The two theorems of Curtis and Powell [12] concerning minimax approximations follow immediately. Adapting them to the present notation, the first one would state that the rank of $\boldsymbol{D}$ is less than r; the second that, for $r=k+1$, the signs of the maximum deviations are either all the same as or all opposite to the signs of the corresponding elements of $\boldsymbol{\rho}$. Thus, upper errors correspond to the positive ρ_j and lower errors to negative ρ_j. The present Theorem I appears essentially to contain both the theorems of Curtis and Powell without emphasizing the special case $r=k+1$.

TABLE I
OPTIMUM TWO-SECTION 10- TO 1-Ω QUARTER-WAVE TRANSFORMER WITH 100-PERCENT BANDWIDTH

Maximum of Reflection Coefficient Magnitude		0.4286	
Frequencies (GHz) at which Maximum Occurs	0.5	1.0	1.5
Parameter Values		Gradient Vectors	
$l_1=7.49482$ cm	−0.07333	0.00000	0.21998
$Z_1=2.2361$ Ω	−0.18254	0.36506	−0.18253
$l_2=7.49482$ cm	−0.07333	0.00000	0.21998
$Z_2=4.4721$ Ω	0.09127	−0.18253	0.09127

TABLE II
OPTIMUM THREE-SECTION 10- TO 1-Ω QUARTER-WAVE TRANSFORMER WITH 100-PERCENT BANDWIDTH

Maximum of Reflection Coefficient Magnitude		0.1973		
Frequencies (GHz) at which Maximum Occurs	0.5	0.770	1.230	1.5
Parameter Values		Gradient Vectors		
$l_1=7.49482$ cm	−0.04377	−0.03227	0.05155	0.13131
$Z_1=1.63471$ Ω	−0.43556	0.38111	0.38111	−0.43555
$l_2=7.49482$ cm	−0.09129	0.06453	−0.10310	0.27387
$Z_2=3.16228$ Ω	0.00000	0.00000	0.00000	0.00000
$l_3=7.49482$ cm	−0.04377	−0.03227	0.05155	0.13131
$Z_3=6.11729$ Ω	0.11639	−0.10184	−0.10184	0.11639

EXAMPLES

Curtis and Powell [12] gave two examples to illustrate the application of their theorems. The necessary conditions derived here confirm their results.

The necessary conditions will be applied to verify the optimality of two optimum quarter-wave transmission-line transformers. The examples were previously studied by Bandler and Macdonald [7], [9]. The necessary gradient vectors for both examples are derived by Bandler and Seviora elsewhere [13]. Tables I and II summarize the problems and list the partial derivatives.

By inspection, multipliers in the ratio

$$3:2:1$$

for Table I, and by solving three equations in three unknowns

$$3.000:2.812:1.760:1.000$$

for Table II can be found (within rounding errors) for the gradient vectors to satisfy the necessary conditions for optimality.

It is also interesting to relate the gradient vectors in Table I to the various contour diagrams given by Bandler and Macdonald [7]. If only Z_1 and Z_2 are assumed variable, then the design may be optimal over the band 0.5 to 1.0 GHz as well as 0.5 to 1.5 GHz. If only l_1 and l_2 are assumed variable, the design may be optimal over the band 0.5 to 1.5 GHz but not from 0.5 to 1.0 GHz. These and other conclusions can be quickly reached by applying Theorem 1 as easily as from physical considerations. The three-section design (Table II) can be investigated in a similar manner, if desired.

Conclusions

Observe that no assumptions on the number of equal extrema as compared with the number of parameters has been made. Further, note that if (by optimization or otherwise) we have an almost equal-ripple solution, the conditions might still be valid if it is assumed that the weighting functions are readjusted to equalize the extrema. Consider, for example, Fig. 1(d). Thus the progress of least pth approximation [1]–[3] could also be checked.

Very wide application of these ideas to approximation problems in network design is envisaged, particularly when nonclassical procedures have to be employed. The capability of evaluating gradient vectors associated with network responses is required. Using the adjoint network method [3], [13], [14], this can be readily done for a wide class of networks.

Acknowledgment

The author wishes to thank Dr. E. Della Torre, Department of Electrical Engineering, McMaster University, and R. E. Seviora, Department of Electrical Engineering, University of Toronto, for numerous and useful discussions on the subject matter of this paper; and A. Lee-Chan, McMaster Data Processing and Computing Centre, for programming assistance.

John W. Bandler
Dep. Elec. Eng.
McMaster Univ.
Hamilton, Ont., Canada

References

[1] G. C. Temes and D. Y. F. Zai, "Least pth approximation," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-16, May 1969, pp. 235–237.
[2] 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.
[3] J. W. Bandler and R. E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, Dec. 1970, pp. 1159–1170.
[4] A. D. Waren, L. S. Lasdon, and D. F. Suchman, "Optimization in engineering design," *Proc. IEEE*, vol. 55, Nov. 1967, pp. 1885–1897.
[5] Y. Ishizaki and H. Watanabe, "An iterative Chebyshev approximation method for network design," *IEEE Trans. Circuit Theory*, vol. CT-15, Dec. 1968, pp. 326–336.
[6] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the nonlinear case," *Comput. J.*, vol. 12, Feb. 1969, pp. 63–68.
[7] J. W. Bandler and P. A. Macdonald, "Cascaded noncommensurate transmission-line networks as optimization problems," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-16, Aug. 1969, pp. 391–394.
[8] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, Aug. 1969, pp. 533–552.
[9] J. W. Bandler and P. A. Macdonald, "Optimization of microwave networks by razor search," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, Aug. 1969, pp. 552–562.
[10] J. W. Bandler, "Computer optimization of inhomogeneous waveguide transformers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, Aug. 1969, pp. 563–571.
[11] H. W. Kuhn and A. W. Tucker, "Non-linear programming," in *Proc. 2nd Symp. on Math. Statistics and Probability*. Berkeley, Calif.: University of California Press, 1951, pp. 481–493.
[12] A. R. Curtis and M. J. D. Powell, "Necessary conditions for a minimax approximation," *Comput. J.*, vol. 8, Jan. 1966, pp. 358–361.
[13] J. W. Bandler and R. E. Seviora, "Sensitivities in terms of wave variables," in *Proc. 8th Annu. Allerton Conf. Circuit and System Theory* (Urbana, Ill., Oct. 1970), pp. 379–387.
[14] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, Aug. 1969, pp. 318–323.

On the Formal Theory of Nonuniform Transmission Lines

Abstract—Some results are given in the formal theory of nonuniform transmission lines whose constitutive parameters are bounded measurable functions of the geometric abscissa x and holomorphic functions of the complex variable p in some region D of the complex plane.

Manuscript received July 7, 1970; revised January 26, 1971. This work has been partially sponsored by the Consiglio Nazionale delle Ricerche, Rome, Italy.

Introduction

The formal theory of generalized nonuniform transmission lines has been investigated by Wohlers [1], [2], mostly under the assumption that the line parameters (i.e., the per-unit-length longitudinal impedance and transverse admittance) can be represented each as the product of a function of the complex variable p, holomorphic in some region of the p plane, and of a piecewise continuous function of the geometric abscissa x.

In this correspondence, a formal theory is developed based on a more general mathematical model, for which the line parameters are simply assumed to be functions of both p and x, holomorphic with respect to p in some region D of the complex plane and measurable with respect to x in the interval $]0, l[$ (where l is the line length).

The main result is an existence and uniqueness theorem, ensuring the analyticity of the chain matrix in the same region D where the line parameters are holomorphic (or in particular the entireness of the chain matrix as a consequence of the latter being entire). Under the further assumption of passivity, the existence of the scattering and immittance representations is proven. Some results are given for nonreciprocal and/or active structures and generalized lossless lines.

The Existence and Uniqueness Theorem

The formal theory of generalized nonuniform lines can be based upon the existence and uniqueness theorem for the following abstract Cauchy problem:

$$du/dx = Au + f \tag{1a}$$

$$u(0) = u_0 \tag{1b}$$

where

$$u(x) = \begin{bmatrix} V(x, p) \\ I(x, p) \end{bmatrix}; \qquad f(x) = \begin{bmatrix} E(x, p) \\ J(x, p) \end{bmatrix} \tag{2}$$

$$A(x) = -\begin{bmatrix} 0 & z(x, p) \\ y(x, p) & 0 \end{bmatrix}. \tag{3}$$

The meaning of the symbols should be self-explanatory.

To set up a convenient mathematical framework for solving problem (1), we introduce the following abstract spaces.

B: the Banach space of the ordered pairs of functions

$$f(p) = \begin{bmatrix} f_1(p) \\ f_2(p) \end{bmatrix} \tag{4}$$

which are holomorphic in some region D of the complex p plane and continuous in its closure $\bar{D}$, endowed with the norm

$$\|f\|_B = \max_{p\in\bar{D}} |f_1(p)| + \max_{p\in\bar{D}} |f_2(p)|. \tag{5}$$

$L^q(0, l; B)$: the Banach space of the functions u: $]0, l[\to B$ whose qth power is summable in $]0, l[$, endowed with the norm

$$\|u\|_{L^q(0,l;B)} = \left(\int_0^l \|u(x)\|_B^q \, dx \right)^{1/q}. \tag{6}$$

$W^{kq}(0, l; B)$: the Banach space of the functions u: $]0, l[\to B$, such that $u \in L^q(0, l; B)$ and $(d^m u/dx^m) \in L^q(0, l; B)$ $(m = 0, 1, 2, \cdots, k)$, with the norm

$$\|u\|_{W^{kq}(0,l;B)} = \left(\sum_{m=0}^{k} \int_0^l \left\| \frac{d^m u}{dx^m} \right\|_B^q dx \right)^{1/q}. \tag{7}$$

$L^\infty(0, l; B)$: the Banach space of the functions u: $]0, l[\to B$, measurable and bounded, with the norm

$$\|u\|_{L^\infty(0,l;B)} = \operatorname*{ess\,sup}_{0\le x\le l} \|u(x)\|_B. \tag{8}$$