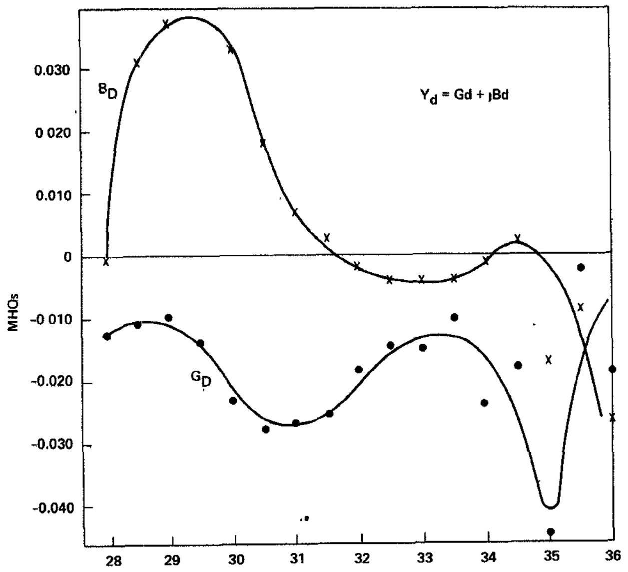

Fig. 3. Terminal diode admittance parameters for one vertical diode position.

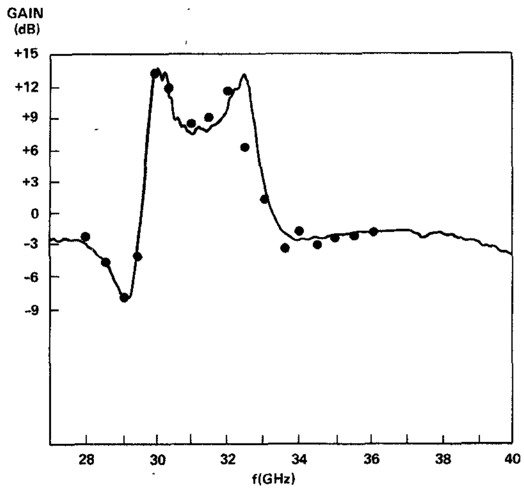

Fig. 4. Calculated versus measured gain with short 0.411 in from diode center. Terminal admittance data used in the calculations were derived from measurements taken with the short 0.142 in from center.

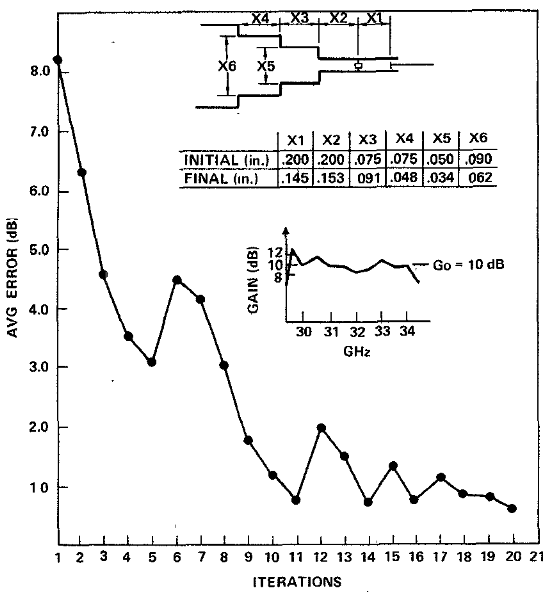

Fig. 5. Computer optimization of two-step matching transformer for 10-dB gain amplifier (30–34 GHz).

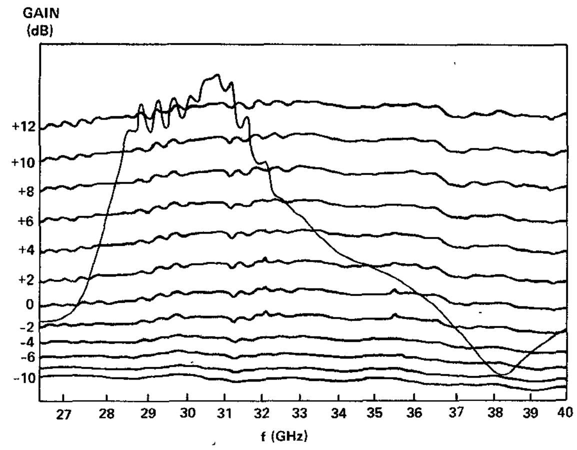

Fig. 6. Measured gain of computer optimized amplifier.

variables) was designed to optimize an amplifier for 10-dB gain over the 30–34-GHz range. Step discontinuities were accounted for in the program. Almost arbitrary initial conditions shown in Fig. 5 led to the final values and computed gain shown. The structure was fabricated and produced over 10-dB gain over a 3.5-GHz band centered 5 percent lower than calculated (Fig. 6). Other positions of the diode and short have produced gain$^{1/2}$-bandwidth products even higher than 15 GHz, indicating the potential usefulness of this type of fabrication.

REFERENCES

[1] J. G. deKoning and R. E. Goldwasser, "A wide band, two-stage *Ka*-band Gunn effect amplifier," in *1974 Millimeter Wave Techniques Conf.*, Naval Electronics Lab. Cen., vol. 1, pp. A3-1–A3-13.
J. G. deKoning *et al.*, "A four stage, 30 dB gain, 100 mW Gunn effect amplifier in *Ka*-band," in *1974 IEEE Microwave Theory and Techniques Int. Symp. Dig.*, pp. 318–319.

[2] A. Séne and F. J. Rosenbaum "A wide-band Gunn-effect CW waveguide amplifier," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 645–650, Oct. 1972.

[3] S. Baskaran and P. N. Robson, "Noise performance of InP reflection amplifiers in *Q* band," *Electron. Lett.*, vol. 8, pp. 137–138, 1972.

Optimization of Microwave Networks

CHRISTAKIS CHARALAMBOUS, MEMBER, IEEE, AND ANDREW R. CONN

***Abstract*—The application of a new algorithm for minimax optimization is investigated. Unlike most of the previously published algorithms the new algorithm uses to its advantage certain obvious properties of the minimax function, namely, that the discontinuities in the first derivatives can be characterized by projections. An N-section transmission-line transformer is used as a test problem.**

I. INTRODUCTION

The problem under consideration is to minimize $M_f(x)$ where

$$M_f(x) = \max_{1 \leq i \leq m} f_i(x)$$

$$x = [x_1 x_2 \cdots x_n]^T$$

$$[M] = \{1,2,\cdots,m\}.$$

Manuscript received February 6, 1975, revised May 5, 1975. This work was supported in part by the National Research Council of Canada under Grant A8639.

The authors are with the Department of Mathematics and Combinatorics, University of Waterloo, Waterloo, Ont., Canada.

The objective function $M_f(x)$ has discontinuous first partial derivatives at points where two or more maxima are equal, even if $f_i(x)$, $1 \leq i \leq m$ have continuous first partial derivatives. Thus we cannot directly use the well-known gradient methods to minimize $M_f(x)$.

Some of the most relevant algorithms for solving the above problem are due to Waren *et al.* [1], Osborne and Watson [2], Bandler *et al.* [3], and Bandler and Charalambous [4]–[6]. See also [7].

The first method transforms the problem into a nonlinear programming problem and solves it by well-established methods. The second and third methods use linear programming to determine the direction of search and then a linear search follows along the direction of search. The fourth method tries to smooth the function around the places where $M_f(x)$ has discontinuous first partial derivatives by using the generalized least pth objective function and by doing this we can use efficient gradient methods.

In our algorithm the direction of search at each iteration consists of two components. The first, the horizontal component which tries to keep locally the same set of the functions near active (two or more functions are considered near active if they are equal to the present maximum up to a specified tolerance) and at the same time to decrease the value of $M_f(x)$. (We do this by using projection matrices.) The second, the vertical component, attempts to satisfy the near active functions exactly by means of linearization. A linear search follows after the horizontal component has been calculated. The linear search incorporates several simple features of the algorithm, and numerical results to date suggest the resulting algorithm is very efficient. Both components are obtained by transforming the minimax problem into its equivalent nonlinear programming problem (see, for example, [1]).

II. THEORETICAL CONSIDERATIONS

Notation: R^n is an n-dimensional real vector space. The minimax problem can be transformed into the nonlinear programming problem: minimize z (a new independent variable) subject to

$$\phi_i(x,z) = z - f_i(x), \qquad i \in [M], \qquad z \in R^1$$

$$E(x,\epsilon) = \{i \in [M] \mid | z - f_i(x) | < \epsilon\} \text{ (active "constraints")}$$

$$I(x,\epsilon) = [M] \backslash E(x,\epsilon) \text{ (inactive "constraints")}$$

$$e = [1,0,\cdots,0]^T \text{ an } (n+1)\text{-dimensional unit vector.}$$

The Algorithm

Step 0: Label = 0.

A. To Determine the Horizontal Direction

Step 1: Set $z = M_f(x^k)$. Determine the active "constraints" at x^k. In other words determine $E(x^k,\epsilon)$.

Step 2: Determine the projection matrix. For $j = 0$ let

$$P^{(0)} = I \text{ (the } (n+1) \times (n+1) \text{ identity matrix)}$$

$$A_0 = \emptyset \text{ (the empty set).}$$

Now for an arbitrary integer $j > 0$ we shall define

$$N^{(j)} = \begin{bmatrix} \nabla\phi_{i_1}^T(x,z) \\ \cdot \\ \cdot \\ \cdot \\ \nabla\phi_{i_j}^T(x,z) \end{bmatrix}.$$

Also, $A_j = (i_1,i_2,\cdots,i_j)$ has been defined at the $j-1$ step. Let

$$P^{(j)} = I - (N^{(j)})^T[N^{(j)}(N^{(j)})^T]^{-1}N^{(j)}.$$

(*Note:* In practice, of course, we do not in fact actually compute matrix inverses but use the iterative formulas of Rosen.) $P^{(j)}$ is an $(n+1) \times (n+1)$ matrix which projects every vector orthogonal to the space spanned by the vectors $\nabla\phi_{i_1}, \nabla\phi_{i_2}, \cdots, \nabla\phi_{i_j}$. Set

$$q^{(j)} = P^{(j)}e.$$

(Note that $q^{(j)}$ is an uphill direction.)

$$i_{j+1} = \left\{\text{the } i \text{ that maximizes } \frac{(q^j)^T\nabla\phi_i(x,z)}{\| q^{(j)} \| \, \| \nabla\phi_i(x,z) \|} \mid i \in E(x,\epsilon)\backslash A_j\right\}.$$

[By $\| x \|$ we mean $(x^Tx)^{1/2}$.] If $-(q^{(j)})^T\nabla\phi_{i_{j+1}}(x,z) < 0$ then

$$A_{j+1} = A_j \cup \{i_{j+1}\}.$$

Otherwise,

$$A(x,\epsilon) = A_j$$

$$P(x,\epsilon) = p^{(j)}$$

$$q(x,\epsilon) = q^{(j)}.$$

However, if the number of elements in the index set A_j reaches $n+1$ go to Step 2a below. Otherwise, go to Step 3.

B. Check if Optimum is Reached

Step 3: If $\| q \| < \epsilon$ and label $\neq 6$ or if $\epsilon < 10^{-6}$ stop.

C. Linear Search

Step 4: Determine $\tau > 0$ such that

$$\max f_i(x^k - \tau\bar{q}), \qquad i \in [M]$$

is minimized, where $\bar{q}$ is the q of Step 2 with the first component deleted.

For details of how this is done and whether exact minimization is required, see the subalgorithm "The Linear Search Algorithm" below. Put $x = x^k - \tau\bar{q}$.

Step 5: Decision as to whether to do the vertical step or not. If the number of active constraints has not changed in three consecutive iterations, and $\| q \| < 0.1$ or if label = 6 go to Step 6. Otherwise, go to Step 1, setting $x^{k+1} = x$.

D. The Vertical Step

Step 6: Set $z = M_f(x)$. Determine $E(x,\epsilon)$ and hence

$$N_v = \begin{bmatrix} \nabla\phi_1^T \\ \cdot \\ \cdot \\ \cdot \\ \nabla\phi_j^T \end{bmatrix}, \quad \text{where} \quad E(x,\epsilon) = \{1,\cdots,j\}.$$

Put

$$v(x,\epsilon) = -N_v^T(N_vN_v^T)^{-1}\Phi$$

where

$$\Phi^T = (\phi_1,\cdots,\phi_j).$$

Put

$$x_{\text{temp}} = x + v \text{ (with first component missing)}$$
$$= x + \bar{v}.$$

If

$$\max_{i\in[M]} f_i(x_{\text{temp}}) < \max_{i\in[M]} f_i(x,$$

put $x^{k+1} = x_{\text{temp}}$ and go to Step 0.

Step 2a: The number of elements in A_j equals $n+1$ means one of two possibilities.

a) We have reached the neighborhood of the optimum.

b) We have constraints considered active that actually are not.

This situation is handled in two ways. First, by ensuring that we take a vertical step and second by reducing ϵ. Algorithmically, set label = 6, put $\epsilon = \epsilon/10$, and go to Step 5.

The Linear Search Algorithm

Step 1: Estimate any new function to become active. We consider all inactive constraints ϕ_i and estimate, in turn, the step size to make each ϕ_i zero. Hence we calculate

$$\tau_j = \frac{\phi_j(x^k)}{\nabla\phi_j^T(x^k)q}, \qquad j \in I(x^k,\epsilon).$$

In other words, if ϕ_i is linear we calculate τ_j so that $\phi_j((z,x^k) - \tau_j q) = 0$ and in general we linearize the ϕ's.

Step 2: Omitting unlikely values of τ_j, estimate the optimum τ by *linearizing* the minimax function.

For $j \in I(x^k,\epsilon)$ do the following. If $\tau_j < 0$ or $\tau_j > 10^4$ neglect it as inadmissible. Otherwise, calculate

$$\hat{f}_{ij} \triangleq f_i(x^k) - \tau_j \nabla f_i(x^k)^T \bar{q}, \qquad i \in [M].$$

Put

$$\hat{F}_j = \max_{i \in [M]} \hat{f}_{ij}.$$

Now, determine l such that

$$F_l = \min \hat{F}_j, \qquad j \in I(x^k,\epsilon) \backslash \{ j \mid \tau_j < 0 \quad \text{or} \quad \tau_j > 10^4 \}.$$

Put $\tau_{\text{opt}} = \tau_l$.

Step 3: Determine if τ_{opt} is acceptable. Calculate the *true* minimax value at

$$\hat{x}^k = x^k - \tau_{\text{opt}} \bar{q}.$$

If this new value is an improvement over the old value, then take $\tau = \tau_{\text{opt}}$ and $x^{k+1} = \hat{x}^k$.

Otherwise, use cubic line search on the maximum of the functions taking τ_{opt} as an upper bound.

In [8] the authors proved that the above algorithm will converge to the minimum under mild assumptions.

Comments on the Algorithm

One useful way of looking at the above algorithm is as follows.

First, our direction of search is obtained by formulating the minimax problem as a nonlinear programming problem in the standard way. Whereupon the horizontal and vertical steps are obtained analogous to Conn [9] and Conn and Pietrzykowski [10]. Secondly, however, instead of using the determined horizontal direction to minimize a penalty function (as in [9] and [10]) we proceed to do our minimization on the *min–max function directly*.

The horizontal step will give us a direction which decreases z and is orthogonal to the gradient vectors $\nabla\phi_{i_1},\nabla\phi_{i_2},\cdots,\nabla\phi_{i_j}$ ($\phi_{i_1},\phi_{i_2},\cdots,\phi_{i_j}$ is a subset of the active constraints. The remaining active constraints locally increase in the direction given by the horizontal step since this is the basis on which the projection matrix was determined).

Thus we are trying to keep the values of $\phi_{i_1},\phi_{i_2},\cdots,\phi_{i_j}$ unchanged (if all of them were exactly equal their values will be equal to zero since $z = \max_{i \in [M]} f_i$).

Now let us assume that we succeed in keeping $\phi_{i_1}\cdots\phi_{i_j}$ unchanged. Since z decreases it follows that the f_i will decrease by the same amount (from the definition of ϕ_i).

In the case of linear functions the above is exactly true. In the general case it is approximately true in that we are linearizing locally.

This shows that the horizontal direction tries to follow the path of discontinuous derivatives. The way we build up the projection matrix guarantees that the gradient vectors $\nabla\phi_{i_1},\nabla\phi_{i_2},\cdots,\nabla\phi_{i_j}$ are *linearly independent* and therefore the matrix $N^{(j)}(N^{(j)})^T$ is nonsingular and its inverse exists (this was implemented by Conn and Pietrzykowski [10]). This is an important difference between the way we handle the projection matrix and the way Rosen does [11].

The vertical step tries to make the near active functions equal, and by doing this an effort is made to get *exactly* on the line of the discontinuous derivatives, which is very desirable when we are close to the optimum point.

In practice we do not use an exact cubic linear search but merely ask for sufficient improvement in the minimax value.

If in Step 4 of the main algorithm we decided to do the vertical step then we dispense with the estimation of τ_{opt} as above and merely do the cubic search. The motivation for this is as follows. The estimates for τ_i are based on the surmise that some new function will become active, whereas the vertical step is based on the assumption that this will not be the case.

III. EXAMPLES

The algorithm has been applied to a wide range of problems. We will illustrate here its performance on the design of a two-section and a three-section 100-percent relative bandwidth 10:1 transmission-line transformer problem [3], [4], [6]. Specifically, we want to

$$\min_{x} \max_{\psi \in [\psi_l,\psi_u]} |\rho(x,\psi)|$$

where ρ is the reflection coefficient, ψ is the frequency in gigahertz, $\psi_l = 0.5$ GHz, $\psi_u = 1.5$ GHz. As usual we will use a finite number of sample points of ψ, and so we obtain

$$f_i(x) = |\rho_i(x)| = |\rho(x,\psi_i)|.$$

All the numerical results were obtained by using double precision IBM 360/75 computer.

Two-Section Transformer

For the two-section transformer we used 11 uniformly spaced frequency points of ψ. The section lengths l_1 and l_2 and the characteristic impedances z_1 and z_2 may be considered as variables.

At first we kept $l_1 = l_2 = l_q$ and used z_1 and z_2 as variables. l_q is the quarter-wavelength at center frequency. Starting from the point $z_1 = 1$, $z_2 = 3$, the new algorithm generated the results shown in Table I. It is important to note that the algorithm took 9 iterations and *only* 16 function evaluations to produce very accurate results. This means that the linear search algorithm works extremely well in this case. In Fig. 1 we show the path taken by the algorithm, from which it can be seen that the algorithm follows ridges ($a(b)$ denotes the point reached at iteration a after b function evaluations).

Table II shows the results from four different starting points.

Table III shows the results by varying the lengths and the impedances from two different starting points.

Three-Section Transformer

For the three-section transformer we used the following sample points of ψ in gigahertz

$$\{0.5,0.6,0.7,0.77,0.9,1.0,1.1,1.23,1.4,1.5\}.$$

Table IV shows the results from two different starting points. It should be noted that, from the second starting point, we do not

TABLE I
RESULTS FOR THE TWO-SECTION 10:1 QUARTER-WAVE TRANSFORMER OVER 100-PERCENT BANDWIDTH

$\ell_1 = \ell_2 = \ell_q =$ Starting point $z_1 = 1.0$, $z_2 = 3.0$				
Number of Iterations	Number of Function Evaluations [a]	x_1	x_2	$M_f(x)$
0	1	1	3	
1	2	1.49152	3.00439	0.54402
2	3	1.62678	2.99891	0.49863
3	4	1.94797	3.94177	0.44792
4	5	1.98645	3.93111	0.43749
5	7	2.14988	4.31621	0.43134
6	8	2.15828	4.31264	0.42933
7	10	2.20976	4.44209	0.42882
8	13	2.22696	4.45590	0.42877
9	14	2.22777	4.45550	0.42858
9	16	2.23330	4.46661	0.42857

[a] The number of function evaluations are cumulative, i.e., iteration seven, for example, took two function evaluations.

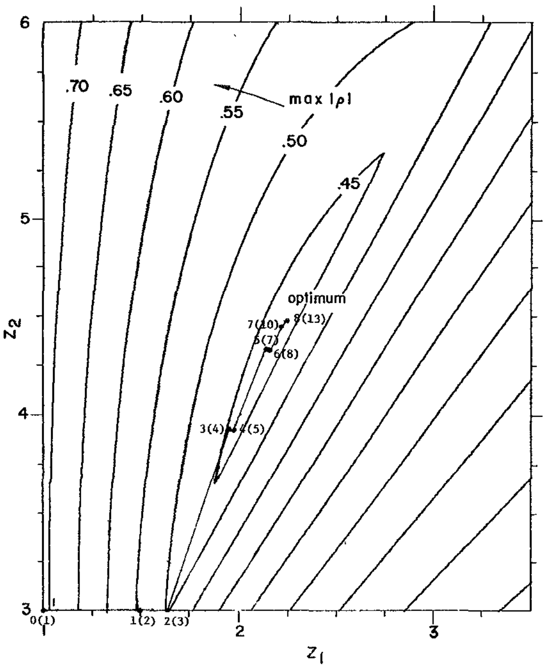

Fig. 1. Example illustrating the path taken by the new algorithm.

TABLE II

OPTIMIZATION OF A TWO-SECTION QUARTER-WAVE TRANSFORMER OVER 100-PERCENT BANDWIDTH ($l_1 = l_2 = l_q$)

Starting Point		Function Evaluations[a]
z_1	z_2	
1.0	3.0	16
1.0	6.0	14
3.5	6.0	13
3.5	3.0	14

[a] Number of function evaluations to reach a reflection coefficient of 0.42857.

TABLE III

OPTIMIZATION OF A TWO-SECTION 10:1 TRANSFORMER OVER A 100-PERCENT BANDWIDTH WITH VARIABLE LENGTHS AND IMPEDANCES

Starting Point				Function Evaluations	
ℓ_1/ℓ_q	ℓ_2/ℓ_q	z_1	z_2	A	B
1.0	1.0	1.0	3	10	16
1.2	0.8	3.5	3.0	41	75

A: Function evaluations required to bring the reflection coefficient within 0.01 percent of its optimum value.

B: Function evaluations required to bring the reflection coefficient to 0.42857.

appear to be converging rapidly to the solution. However, at the point obtained in 80 function evaluations the "active" functions are all within 10^{-8} of each other and the directional derivative for z in the corresponding nonlinear programming problem is of order 10^{-5}. Furthermore, the active functions are correctly identified. Consequently, it is reasonable to assume that the answer obtained is acceptable. Should full accuracy be desired it can be obtained, but, understandably, convergence is slow. In fact the answer is obtained to 7 significant figures after 200 function evaluations.

TABLE IV

OPTIMIZATION OF A THREE-SECTION TRANSFORMER OVER A 100-PERCENT BANDWIDTH WITH VARIABLE LENGTHS AND IMPEDANCES

Starting Point						Function Evaluations	
ℓ_1/ℓ_q	ℓ_2/ℓ_q	ℓ_3/ℓ_q	z_1	z_2	z_3	A	B
1.0	1.0	1.0	1.0	3.16228	10.0	38(0.19735)	67(0.19729)
0.8	1.2	0.8	1.5	3.0	6.0	79(0.19731)	162(0.19729)

In comparing our numerical results with the previous published results [3], [4], [6], it can be seen that the new algorithm is at least as efficient as the existing algorithms, and our experience suggests it is better.

IV. CONCLUSIONS

A new algorithm for computer-aided minimax optimization has been presented. From the limited numerical results presented here, along with other experience (see [8]) we have, it seems that the algorithm is efficient. It is important to note that if the functions are linear then the minimax function will be piecewise linear and the present algorithm will take into consideration this fact. The authors are presently investigating this and preliminary results are encouraging. In fact the algorithm [12] will terminate in a finite number of steps in this case, and appears superior to any previous linear L_∞ algorithm. Also, we found that our linear search works very well, and to the authors' knowledge it is the first time a special linear search was used to minimize minimax functions.

It should be pointed out that our horizontal component is the same as the direction of search at each iteration of the algorithm proposed by Bandler *et al.* [3], but in their paper they have not considered the vertical step which means that their algorithm will only reach the optimum within a supplied tolerance and final convergence in their case is very difficult. Furthermore, they take the linear programming approach whereas we use orthogonal projections. The connections between these two approaches are well known. It is suggested that optimization problems of a similar nature, for example, standard nonlinear programming problems with an objective function that is continuous but has discontinuous derivatives, might be solved by an analogous method.

The mathematical background with proofs of convergence, and so on, is rather involved and lengthy. It is available elsewhere [8].

REFERENCES

[1] A. D. Waren, L. S. Lasdon, and D. F. Suchman, "Optimization in engineering design," *Proc. IEEE*, vol. 55, pp. 1885–1897, Nov. 1967.

[2] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the non-linear case," *Comput. J.*, vol. 12, pp. 63–68, 1969.

[3] J. W. Bandler, T. V. Srinivasan, and C. Charalambous, "Minimax optimization of networks by grazor search," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 596–604, Sept. 1972.

[4] J. W. Bandler and C. Charalambous, "Practical least *p*th optimization of networks," *IEEE Trans. Microwave Theory Tech. (1972 Symp. Issue)*, vol. MTT-20, pp. 834–840, Dec. 1972.

[5] C. Charalambous and J. W. Bandler, "Nonlinear minimax optimization as a sequence of least *p*th optimization with finite values of *p*," Faculty Eng., McMaster Univ. Hamilton, Ont., Canada. Rep. SOC-3, 1973.

[6] ——, "New algorithms for network optimization," *IEEE Trans. Microwave Theory Tech. (1973 Symp. Issue)*, vol. MTT-21, pp. 815–818, Dec. 1973.

[7] C. Charalambous, "A unified review of optimization", *IEEE Trans. Microwave Theory Tech. (Special Issue on Computer-Oriented Microwave Practices)*, vol. MTT-22, pp. 289–300, Mar. 1974.

[8] C. Charalambous and A. R. Conn, "An efficient method to solve the minimax problem directly," Dep. Combinatorics and Optimization, Univ. Waterloo, Waterloo, Ont., Canada, Rep. CORR-74-19.

[9] A. R. Conn, "Constrained optimization using nondifferentiable penalty function," *SIAM J. Num. Anal.*, vol. 10, pp. 760–784, 1973.

[10] A. R. Conn and T. Pietrzykowski, "A penalty function method converging directly to a constrained optimum," Dep. Combinatorics and Optimization, Univ. Waterloo, Waterloo, Ont. Canada, Rep. 73-11, 1973.
[11] J. B. Rosen, "The gradient projection method for nonlinear programming, Part 1: linear constraints," *J. Soc. Ind. Appl. Math.*, vol. 8, 1960,
[12] C. Charalambous and A. R. Conn, "A new algorithm for discrete L_∞ linear approximation," to be published.

Minimax Optimization by Algorithms Employing Modified Lagrangians

OLOV EINARSSON

Abstract—Two general, modified Langrangian algorithms related to recent developments in nonlinear programming are presented. The methods give accurate results and are easy to program. An N-section transmission-line transformer is used as a test problem for minimax (equal ripple) optimization and the methods are compared to existing algorithms for network optimization.

I. INTRODUCTION

There exists a large class of optimization problems of engineering interest where some finite-dimensional functional is minimized (or maximized) subject to an equal ripple condition. The purpose of this short paper is to draw attention to the existence of two effective, recent algorithms which can be applied with advantage to this type of problem. While not new, the methods do not appear to have been applied to microwave problems before. The methods proposed are quite general and the choice of a transmission-line transformer problem as an example is only dictated by its use as a test problem in previous works on minimax optimization [1]–[3].

Consider the following minimax problem. Find the vector x which minimizes the real-valued function $f(x)$; i.e., find

$$\min_{x \in R^n} f(x) \tag{1}$$

where $f(x)$ is defined as

$$f(x) \triangleq \max_{\nu \in I} \tfrac{1}{2} \mid \rho(x,\nu) \mid^2 \tag{2}$$

and ρ is the reflection coefficient of the N-section lossless transmission-line transformer shown in Fig. 1. In (2) the frequency ν, normalized to some suitable frequency ν_0, is varied either over a closed interval

$$I = [\nu_1,\nu_M] \tag{3a}$$

or over a finite set

$$I = \{\nu_i\}_1{}^M. \tag{3b}$$

The components of the n-dimensional vector x in (1) are the (real) characteristic impedances and the lengths of the transmission-line sections. In one version the lengths of the sections are kept constant and equal to $\lambda_0/4$ where $\lambda_0 = c/\nu_0$. The corresponding x vector is

Manuscript received November 4, 1974; revised May 5, 1975.
The author was on leave from the University of Lund, at the Division of Radio and Electrical Engineering, National Research Council of Canada, Ottawa, Ont., Canada. He is now back at the College of Engineering, University of Lund, Lund, Sweden.

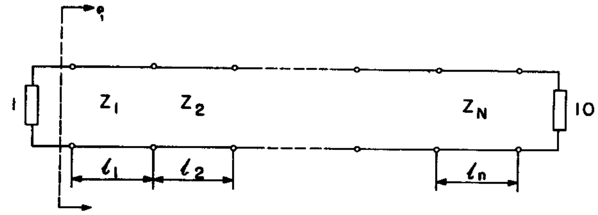

Fig. 1. 100-percent relative bandwidth 10:1 transmission-line transformer.

$$x = (Z_1,Z_2,\cdots,Z_N), \qquad n = N$$
$$l_1 = l_2 = \cdots = l_N = \lambda_0/4. \tag{4a}$$

Alternatively, the lengths of the transmission-line sections may also be varied, resulting in an x vector

$$x = (Z_1,Z_2,\cdots,Z_N,l_1,l_2,\cdots,l_N), \qquad n = 2N. \tag{4b}$$

The solution of the transmission-line transformer problem is known in terms of Chebyshev polynomials. The optimal lengths all turn out to be equal to $\lambda_0/4$ and the optimal impedances can be determined from the polynomial expressing the insertion loss function [4]. However, the methods of this investigation do not rely on this special polynomial structure of the problem.

II. DISCUSSION OF METHODS

It is readily seen that the unconstrained minimization problem given by (1), (2), and (3b) is equivalent to the following problem:

$$\min_{x \in R^n} f_1(x) \tag{5}$$

subject to the $M - 1$ nonlinear constraints

$$f_i(x) - f_1(x) \leq 0, \qquad i = 2,3,\cdots,M \tag{6}$$

where we have defined

$$f_i(x) \triangleq \tfrac{1}{2} \mid \rho(x,\nu_i) \mid^2 \tag{7}$$

and where we have used the fact that $\mid \rho(x,\nu) \mid$ takes its maximum value at the left end point ν_1 of the frequency range.

One well-established way of handling a nonlinear constrained optimization problem is to introduce a Lagrange multiplier for each constraint and construct a Lagrangian which will be stationary at the solution point. However, in the treatment of nonconvex problems, the usefulness of the Lagrangian is limited by the fact that a stationary point may not necessarily correspond to a minimum. Hestenes [5] and Powell [6] independently discovered that this drawback could be overcome by augmenting the Lagrangian with a quadratic penalty function term making the problem locally convex in a neighborhood of the solution point. A large number of different modified Lagrangians (also called "exact penalty functions" or "augmented Lagrangians") have been proposed both for problems with equality and inequality constraints [7], [8]. The modified Lagrangians employed in this short paper are given explicitly by (A1) and (A4) of Appendix A.

In most cases the Lagrange multipliers related to an optimization problem are unknown and have to be calculated during the minimization procedure. One way of doing this is to employ the saddle-point property of the modified Lagrangian $L(x,\mu)$ in the product space $[x,\mu]$ where μ is the multiplier vector, and solve the dual problem, i.e., to find

$$\max_{\mu \in R^m} \min_{x \in R^n} L(x,\mu) \tag{8}$$

by iterating alternately in x and μ space. The updating of the multipliers can be done in different ways. A very simple rule is used in the Hestenes–Powell method. For the routine VF01A [9] advantage is taken of the fact that if the unconstrained minimizations are