Offprints From
Modern Filter Theory and Design
Edited by Gabor C. Temes & Sanjit K. Mitra
Published by John Wiley & Sons Inc., 1973

6

Computer-Aided Circuit Optimization

J. W. Bandler
McMaster University
Hamilton, Ontario

This chapter deals with formulations and methods which can be implemented in the ever increasing number of situations when the classical synthesis approach, whether analytic or numerical, is inappropriate. When the so-called closed-form solution is, for some reason, out of the question, the modern approach is to use efficient, iterative, automatic optimization methods to achieve a design that meets or exceeds certain requirements. Not infrequently, exact methods may be used to great advantage in providing the initial feasible design for optimization.

In order to make the mathematics tractable, usable synthesis methods are usually restricted to ideal *commensurate* networks. As soon as we have to take into account active devices, a narrow range of element values, parasitic effects, high frequency operation, nonlinearities, frequency-dependent elements, *noncommensurate* elements (e.g., mixed lumped and distributed elements, uniformly distributed transmission lines with unequal or variable lengths, etc.), elements characterized by measurement data, response constraints, and so on, classical methods of design provide, at best, only approximate answers. In some cases these answers adequately approximate the solution to the actual design problem, but in many cases they do not.

The author is not advocating numerical methods for their own sake. Generally speaking, for the same job, iterative methods require more computation time than more specialized methods which do not require iteration (if they are available). Computing time is not, however, the only criterion an engineer has to consider. For example, in deciding whether or not to devote his own time to deriving an *analytic* algorithm, as distinct from a *numerical* algorithm, he also has to ask himself how often the algorithm would be used, how well it would represent real situations, how widely applicable it would be, and last but not least, how accurate the numerical results would be. After all, as engineers, we are ultimately working toward producing meaningful numbers as the solutions to realistic design problems.

Methods for *automating* the optimal design process will be emphasized. *Ad hoc* cut-and-try techniques using a general purpose analysis program are discouraged, particularly for filter design problems with anything other than the simplest of design specifications and a handful of variable parameters. The pitfalls are the same as with automated methods, the strategy for dealing with them is inevitably less sophisticated, and in the long run it will almost certainly cost more. It is desirable that the decision making process should, as far as possible, be left to the computer.

Poor or unacceptable results in computer-aided circuit optimization (or with any design process) are felt to be most likely due to bad preparation of the problem, a lack of understanding of the hazards that can be encountered, and the wrong choice of algorithm. This chapter, therefore, attempts to show how problems within the scope of filter design may be formulated effectively as optimization problems, to explain the differences between these formulations, to indicate appropriate optimization methods, and to indicate how the results might be interpreted. Details of optimization algorithms, proofs of convergence, etc., are beyond the scope of this work. Adequate references to the original papers and relevant text books will permit the reader to investigate these for himself.

Following a section on basic concepts which are essential for an understanding of optimization theory, a formulation and description of typical objectives and objective functions is presented. Constraints and some methods of dealing with them are discussed in fair detail in the next section, including the conditions for a constrained minimum. Minimax approximation, including conditions for a minimax optimum, is then dealt with. This is followed by sections on one-dimensional search methods, direct search methods, and methods using gradient information. Least pth approximation comes naturally after a discussion of gradient methods. A fairly long section is devoted to the adjoint network method of gradient evaluation.

6.1 BASIC CONCEPTS

The problem of optimization may be stated as follows. Minimize the scalar *objective function** U where

$$U \triangleq U(\boldsymbol{\phi}) \tag{6.1}$$

subject to the *inequality constraints*

$$\mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0} \tag{6.2}$$

and *equality constraints*

$$\mathbf{h}(\boldsymbol{\phi}) = \mathbf{0}. \tag{6.3}$$

In (6.1) to (6.3), $\boldsymbol{\phi}$ is a vector of k independent variables or parameters,† thus

$$\boldsymbol{\phi} \triangleq \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_k \end{bmatrix} \tag{6.4}$$

defining a k-dimensional space. In general, we might have m inequality constraints and s equality constraints so that

$$\mathbf{g}(\boldsymbol{\phi}) \triangleq \begin{bmatrix} g_1(\boldsymbol{\phi}) \\ g_2(\boldsymbol{\phi}) \\ \vdots \\ g_m(\boldsymbol{\phi}) \end{bmatrix} \tag{6.5}$$

and

$$\mathbf{h}(\boldsymbol{\phi}) \triangleq \begin{bmatrix} h_1(\boldsymbol{\phi}) \\ h_2(\boldsymbol{\phi}) \\ \vdots \\ h_s(\boldsymbol{\phi}) \end{bmatrix} \tag{6.6}$$

The *feasible region* R is defined by all vectors $\boldsymbol{\phi}$ satisfying (6.2) and (6.3). This may be written

$$R \triangleq \{\boldsymbol{\phi} \,|\, \mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}, \mathbf{h}(\boldsymbol{\phi}) = \mathbf{0}\}. \tag{6.7}$$

* Also called *cost function*, *performance index*, or *error criterion*.

† Typically element values, residues, critical frequencies, etc.

R is said to be *closed* if, as in (6.2), equalities are allowed. If no equalities are allowed it is said to be *open*. A *proper* minimum of U located by a vector $\check{\phi}$ on the *response hypersurface* generated by $U(\phi)$ is such that

$$\check{U} \triangleq U(\check{\phi}) < U(\phi) \tag{6.8}$$

for any feasible ϕ close but not equal to $\check{\phi}$.* Since we cannot generally guarantee to find a *global minimum*, we usually have to resign ourselves to a consideration of *local minima*. Our objective then is to find a feasible $\check{\phi}$, if it indeed exists, such that

$$U(\check{\phi}) = \min_{\phi \in R} U(\phi).$$

Figure 6-1 is an illustration of the problem in two dimensions, and it contains a number of features usually encountered in optimization problems. Note that only inequality constraints, i.e., constraints of the form of (6.2), are indicated.

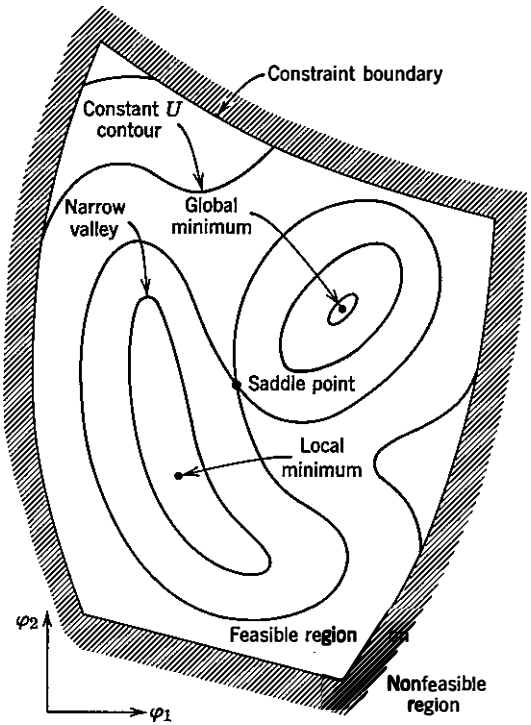

FIGURE 6-1 Some features encountered in optimization problems.

* $U(\tilde{\phi}) \leq U(\phi)$ can also define a minimum, but $\tilde{\phi}$ may then be nonunique.

Examples of *unimodal*, *multimodal*, *strictly concave*, and *strictly convex* functions of one variable are shown in Figure 6-2. A unimodal function for our purposes is one having a unique optimum in the feasible region. It may or may not be continuous with continuous derivatives. A *strictly convex*

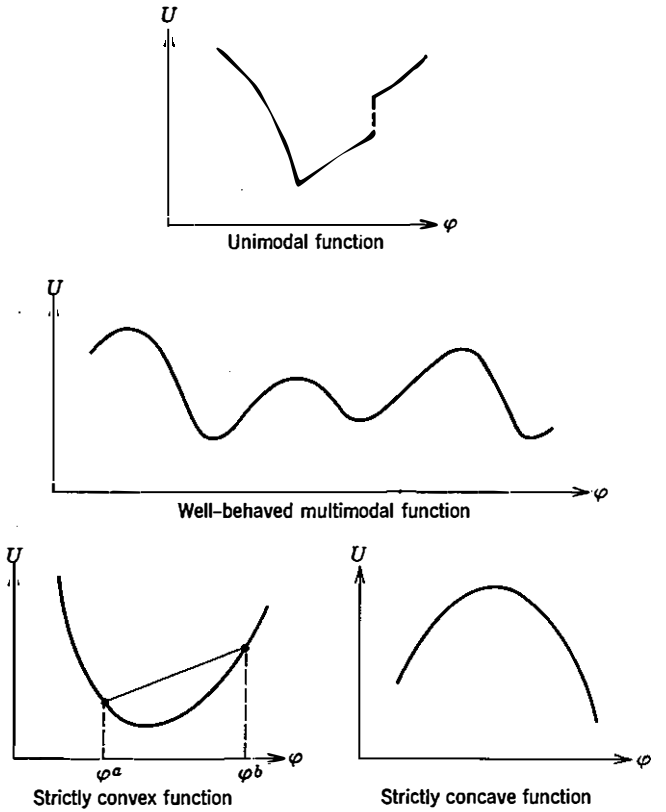

FIGURE 6-2 Functions of one variable.

function is one which can only be overestimated by a linear interpolation between two points on its surface. Thus, for $\phi^a \neq \phi^b$,

$$U(\phi^a + \lambda(\phi^b - \phi^a)) < U(\phi^a) + \lambda(U(\phi^b) - U(\phi^a))$$
$$0 < \lambda < 1 \tag{6.9}$$

for a strictly convex function. See Figure 6-3. A *strictly concave* function is one whose negative is strictly convex. Note that if we omit strictly, then we imply that equality of the function and a linear interpolation can occur, i.e., (6.9) would have to admit equalities.

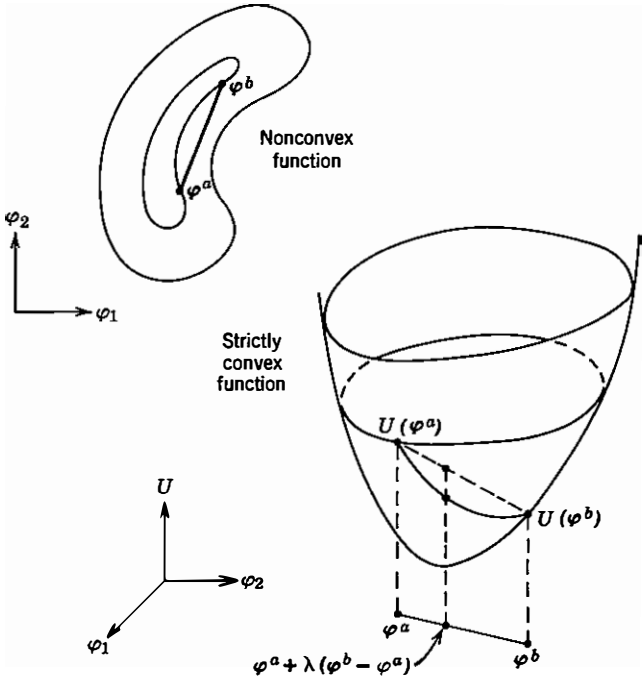

FIGURE 6-3 Illustration of convexity.

A region R is convex if for all $\phi^a, \phi^b \in R$ all points

$$\phi = \phi^a + \lambda(\phi^b - \phi^a)$$
$$0 \leq \lambda \leq 1 \tag{6.10}$$

lie in R. Illustrations of convex and nonconvex regions are given in Figure 6-4.

The first three terms of a multidimensional Taylor series expansion of $U(\phi)$ are given by

$$U(\phi + \Delta\phi) = U(\phi) + \nabla U^T \Delta\phi + \tfrac{1}{2} \Delta\phi^T \mathbf{H} \Delta\phi + \cdots \tag{6.11}$$

where the vector

$$\Delta\phi \triangleq \begin{bmatrix} \Delta\phi_1 \\ \Delta\phi_2 \\ \vdots \\ \Delta\phi_k \end{bmatrix} \tag{6.12}$$

contains k parameter increments, $\Delta\phi^T$ is the transposed (row) vector,

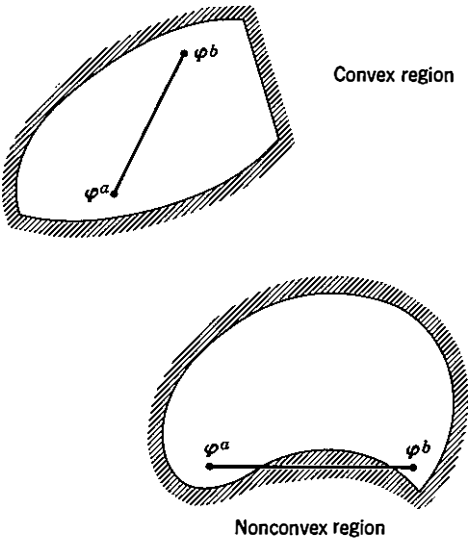

FIGURE 6-4 Convex and nonconvex regions.

$$\nabla U \triangleq \begin{bmatrix} \dfrac{\partial U}{\partial \phi_1} \\ \dfrac{\partial U}{\partial \phi_2} \\ \vdots \\ \dfrac{\partial U}{\partial \phi_k} \end{bmatrix} \tag{6.13}$$

is a vector containing the first partial derivatives of the objective function called the *gradient vector*, and

$$\mathbf{H} \triangleq \begin{bmatrix} \dfrac{\partial^2 U}{\partial \phi_1^2} & \dfrac{\partial^2 U}{\partial \phi_1\, \partial \phi_2} & \cdots & \dfrac{\partial^2 U}{\partial \phi_1\, \partial \phi_k} \\ \dfrac{\partial^2 U}{\partial \phi_2\, \partial \phi_1} & \dfrac{\partial^2 U}{\partial \phi_2^2} & \cdots & \dfrac{\partial^2 U}{\partial \phi_2\, \partial \phi_k} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 U}{\partial \phi_k\, \partial \phi_1} & \dfrac{\partial^2 U}{\partial \phi_k\, \partial \phi_2} & \cdots & \dfrac{\partial^2 U}{\partial \phi_k^2} \end{bmatrix} \tag{6.14}$$

is a symmetric $k \times k$ matrix containing the second partial derivatives and called the *Hessian* matrix. At a minimum of a continuous function with continuous first and second partial derivatives $\nabla U(\check{\boldsymbol{\phi}}) = \mathbf{0}$ and $\mathbf{H}(\check{\boldsymbol{\phi}})$ is positive semidefinite.* Invoking these conditions in (6.11) but with $\mathbf{H}$ taken as positive definite, it may be shown that (6.8) is satisfied, implying that we have a proper minimum. $U(\boldsymbol{\phi})$ is strictly convex in a region where $\mathbf{H}$ is positive definite as may be seen by relating (6.9) with (6.11).

The problem formulated in (6.1) to (6.3) is called a *mathematical programming* problem. If all the functions are linear, we have *linear programming*; if not, we have *nonlinear programming*. The term *convex programming* is often used to describe the problem defined by (6.1) and (6.2) when $U(\boldsymbol{\phi})$ is convex and $\mathbf{g}(\boldsymbol{\phi})$ is concave. Under these conditions R is convex, and $\check{U}$ is the global minimum.

In practical situations, it is usually out of the question to determine whether a specific problem falls into the domain of convex programming. Nevertheless, it seems a fair generalization to make, that the most reliable and efficient methods of optimization for practical problems are invariably those which invoke some of the nice properties of convex programming in their proofs of convergence. The better methods usually have built-in safeguards for dealing with the hazards of more general nonlinear programming problems while substantially retaining their desirable convergence features. Note that essentially unconstrained problems are regarded as special cases in the above discussion.

6.2 SOME OBJECTIVES AND OBJECTIVE FUNCTIONS

Optimization by Solving Nonlinear Equations

Classically, to find $\check{U}$ we must in general solve k nonlinear equations in k unknowns, namely

$$\nabla U = \mathbf{0}.$$

Denoting this set of equations $\mathbf{f}(\boldsymbol{\phi}) = \mathbf{0}$ where

$$\mathbf{f}(\boldsymbol{\phi}) \triangleq \begin{bmatrix} f_1(\boldsymbol{\phi}) \\ f_2(\boldsymbol{\phi}) \\ \vdots \\ f_k(\boldsymbol{\phi}) \end{bmatrix}, \tag{6.15}$$

we could define a new objective function

$$U(\boldsymbol{\phi}) = \mathbf{f}^T\mathbf{f} \tag{6.16}$$

*It should be noted that $\mathbf{H}$ might not be positive definite, even in some cases when $U(\boldsymbol{\phi})$ is strictly convex.

to be minimized. A minimum of value zero would imply that the solution to $\mathbf{f}(\phi) = \mathbf{0}$ had been found. Now, using a Taylor series expansion

$$\mathbf{f}(\phi + \Delta\phi) = \mathbf{f}(\phi) + \mathbf{J}\,\Delta\phi + \cdots \tag{6.17}$$

where

$$\mathbf{J} \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial \phi_1} & \frac{\partial f_1}{\partial \phi_2} & \cdots & \frac{\partial f_1}{\partial \phi_k} \\ \frac{\partial f_2}{\partial \phi_1} & \frac{\partial f_2}{\partial \phi_2} & \cdots & \frac{\partial f_2}{\partial \phi_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_k}{\partial \phi_1} & \frac{\partial f_k}{\partial \phi_2} & \cdots & \frac{\partial f_k}{\partial \phi_k} \end{bmatrix} \tag{6.18}$$

is a $k \times k$ *Jacobian* matrix. The well-known *Newton-Raphson* method of solution is based on the hope that, if we evaluate $\mathbf{f}$ and $\mathbf{J}$ at ϕ, then the incremental change

$$\Delta\phi = -\mathbf{J}^{-1}\mathbf{f}(\phi) \tag{6.19}$$

brings one closer to the solution. (In Section 6.8 these ideas are extended).

Quadratic Objective Function

Consider the quadratic objective function

$$U(\phi) = \tfrac{1}{2}\phi^T\mathbf{A}\phi + \mathbf{b}^T\phi + c \tag{6.20}$$

where

$\mathbf{A}$ is a $k \times k$ constant symmetric matrix,
$\mathbf{b}$ is a constant vector with k components,
c is a constant.

In this case, it is readily shown that

$$\nabla U = \mathbf{A}\phi + \mathbf{b}$$
$$\mathbf{H} = \mathbf{A}.$$

A stationary point of $U(\phi)$ can be found by solving the linear equations

$$\mathbf{A}\phi + \mathbf{b} = \mathbf{0}.$$

If $\mathbf{A}$ is nonsingular, the point is unique and can be found in a finite number of operations. The term *quadratic convergence* (Fletcher [1] prefers the term "property Q") is used to describe the convergence properties of optimization methods, which guarantee to find the minimum of a quadratic function

in a finite number of steps. Such methods can be expected to be very efficient in minimizing functions adequately representable by positive-definite quadratic forms in the vicinity of a minimum. Some of them are discussed in later sections.

Error Criteria

Most electrical network design problems can be formulated as approximation problems. Let us, therefore, introduce a weighted error or deviation between a specified function and an approximating function as

$$e(\phi, \psi) \triangleq w(\psi)[F(\phi, \psi) - S(\psi)] \tag{6.21}$$

where

$S(\psi)$ is the real or complex specified function,
$F(\phi, \psi)$ is the real or complex approximating function,
$w(\psi)$ is a weighting function,
ψ is an independent variable,
ϕ represents the adjustable parameters.

Thus $F(\phi, \psi)$ may be a network response, $S(\psi)$ may be the desired response, and ψ may be frequency or time. See Figure 6-5.

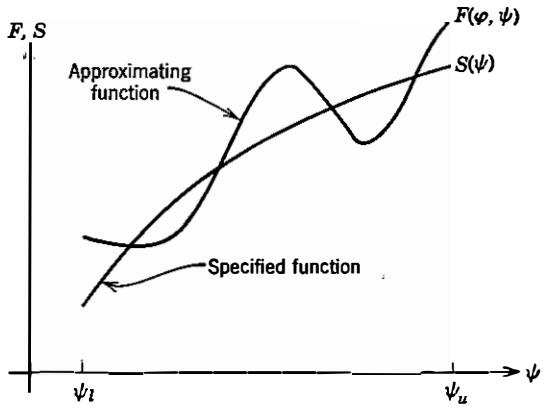

FIGURE 6-5 An approximation problem.

We may define a *norm*

$$\|e\|_p \triangleq \left\{\int_{\psi_l}^{\psi_u} |e(\phi, \psi)|^p \, d\psi\right\}^{1/p}, \qquad 1 \leq p \leq \infty \tag{6.22}$$

for the continuous case; and a *norm*

$$\|\mathbf{e}\|_p \triangleq \left\{\sum_i |e_i(\phi)|^p\right\}^{1/p}, \qquad \begin{matrix} i \in I \\ 1 \le p \le \infty \end{matrix} \tag{6.23}$$

for the discrete case, where

$$\mathbf{e}(\phi) \triangleq \begin{bmatrix} e_1(\phi) \\ e_2(\phi) \\ \vdots \\ e_n(\phi) \end{bmatrix}, \tag{6.24}$$

$$e_i(\phi) \triangleq e(\phi, \psi_i) = w(\psi_i)[F(\phi, \psi_i) - S(\psi_i)] \tag{6.25}$$

and

$$I \triangleq \{1, 2, \ldots, n\}. \tag{6.26}$$

Thus, I is an index set relating to discrete values of ψ on an interval $[\psi_l, \psi_u]$, which is closed and finite.

Now for well-behaved functions

$$\max_{[\psi_l, \psi_u]} |e(\phi, \psi)| = \lim_{p \to \infty} \left\{\frac{1}{\psi_u - \psi_l} \int_{\psi_l}^{\psi_u} |e(\phi, \psi)|^p \, d\psi\right\}^{1/p} \tag{6.27}$$

when $|e(\phi, \psi)|$ is defined on $[\psi_l, \psi_u]$. If $|e(\phi, \psi)|$ is continuous on a finite interval $[\psi_l, \psi_u]$, then (6.27) is certainly valid. Similarly,

$$\max_i |e_i(\phi)| = \lim_{p \to \infty} \left\{\sum_i |e_i(\phi)|^p\right\}^{1/p}, \qquad i \in I. \tag{6.28}$$

Suppose we formulate an objective function as

$$U = \int_{\psi_l}^{\psi_u} |e(\phi, \psi)|^p \, d\psi \tag{6.29}$$

for the continuous case and

$$U = \sum_{i \in I} |e_i(\phi)|^p \tag{6.30}$$

for the discrete case. The minimization of the U of (6.29) or (6.30) is called *least* p*th approximation.* A minimum for the continuous case is called a best approximation with respect to $\|e\|_p$, defined in (6.22). A minimum for the discrete case is called a best approximation with respect to $\|\mathbf{e}\|_p$, defined in (6.23). Now $\|e\|_\infty$ and $\|\mathbf{e}\|_\infty$ are called *Chebyshev* or *uniform* norms. Because of the consequences of (6.27) and (6.28), minimization with respect to $\|e\|_\infty$ or $\|\mathbf{e}\|_\infty$ is widely referred to as *minimax approximation.* Least pth approximation tends to minimax approximation as $p \to \infty$.

A word of caution concerning the weighting function $w(\psi)$ and the index p is in order. Clearly their purpose is to emphasize or deemphasize the difference between $F(\phi, \psi)$ and $S(\psi)$. Thus, an optimum with respect to one weighting function or value of p may not be an optimum with respect to another. Large errors will be emphasized by large values of p. If one knew in advance where these large errors would be, $w(\psi)$ might also be used to emphasize them. The use of $w(\psi)$ to do this is a poor approach, however, and should be discouraged.

6.3 CONSIDERATION OF CONSTRAINTS

It is rare to find any network design problem which is unconstrained. When physical considerations indicate that the optimum will lie in the interior of the feasible region, the designer is lucky and should take advantage of it. Often this will not be possible, and steps have to be taken to ensure that a realizable and practical design will be achieved. One of the great advantages of computer-aided circuit optimization is that, if the design problem has been properly formulated, a feasible design can always be achieved assuming the initial design is feasible.

Constraints in network design can take a variety of forms. They can include upper and lower bounds on parameters; they can include nonnegativity requirements on network elements. The topology, overall size, the suppression of unwanted modes of operation, considerations for parasitic effects whether reactive or lossy, and the stability of active devices can all result in constraints on parameters. Response constraints such as constraints on the phase while the amplitude is optimized can also occur.

Most network designers seem to treat constraints as an afterthought, and then complain that the optimization process gave them negative resistors, etc. Their faith in automated optimization methods is shattered as a result. The author would like to stress that a thorough consideration should be given to the constraints *before* the selection of an optimization strategy.

In this section we will look at some methods of converting constrained problems into essentially unconstrained ones. For other methods of nonlinear programming, the reader should refer elsewhere [2, 3, 4].

Transformations for Parameter Constraints

Various upper and lower bounds on the variable parameters are probably the most common kinds of constraints [5]. In Table 6-1 we show some simple parameter constraints falling into this class with appropriate transformations. It is useful to distinguish between constraints defining open and closed feasible regions. If the optimum is expected to lie away from the boundary or if it is desired to discourage the solution from getting too close, the former type might be chosen.

TABLE 6-1 Simple parameter constraints and transformations

Constraint	Transformation
$\phi_i \geq 0$	$\phi_i = \phi_i'^2$
$\phi_i > 0$	$\phi_i = \exp \phi_i'$
$\phi_i \geq \phi_{li}$	$\phi_i = \phi_{li} + \phi_i'^2$
$\phi_i > \phi_{li}$	$\phi_i = \phi_{li} + \exp \phi_i'$
$-1 \leq \phi_i \leq 1$	$\phi_i = \sin \phi_i'$
$0 \leq \phi_i \leq 1$	$\phi_i = \sin^2 \phi_i'$
$0 < \phi_i < 1$	$\phi_i = \dfrac{\exp \phi_i'}{1 + \exp \phi_i'}$
$\phi_{li} \leq \phi_i \leq \phi_{ui}$	$\phi_i = \phi_{li} + (\phi_{ui} - \phi_{li})\sin^2 \phi_i'$ $\phi_i = \frac{1}{2}(\phi_{li} + \phi_{ui}) + \frac{1}{2}(\phi_{ui} - \phi_{li})\sin \phi_i'$
$\phi_{li} < \phi_i < \phi_{ui}$	$\phi_i = \phi_{li} + (\phi_{ui} - \phi_{li}) \dfrac{\exp \phi_i'}{1 + \exp \phi_i'}$ $\phi_i = \phi_{li} + \dfrac{1}{\pi}(\phi_{ui} - \phi_{li})\cot^{-1} \phi_i'$ for $0 < \cot^{-1} \phi_i' < \pi$

More General Considerations

Parameter constraints of the form

$$\phi_{li} \leq \phi_i \leq \phi_{ui} \tag{6.31}$$

can if necessary be written as

$$\begin{aligned} \phi_i - \phi_{li} &\geq 0 \\ \phi_{ui} - \phi_i &\geq 0 \end{aligned} \tag{6.32}$$

in order to fit them into the scheme of (6.2). Frequency- or time-dependent constraints may be put into the form

$$c_j(\phi, \psi) \geq 0 \tag{6.33}$$

where j denotes some jth function, or at discrete points on the ψ-axis into the form

$$c_j(\phi, \psi_i) \geq 0 \tag{6.34}$$

where i denotes an ith sample point. The form of (6.34) is preferrable to that of (6.33), since it allows us to consider a finite rather than an infinite number of constraints.

We might eliminate a number of constraints on physical or logical grounds if, for instance,

1. $U(\phi) \to \infty$ as $g_i(\phi) \to 0$. The attenuation of a filter becomes infinite, for example, if a zero valued element short circuits the structure;
2. Some $h_i(\phi) = 0$ can be explicitly written as $\phi_j = f(\phi_1, \phi_2, \ldots, \phi_{j-1}, \phi_{j+1}, \ldots, \phi_k)$. In this case we can optimize with $k - 1$ parameters;
3. $g_i(\phi)$ is known *a priori* to be positive.

Our design problem may be so complicated that we cannot easily find an initial design to serve as a feasible starting point in the optimization process. We could try to find one by unconstrained optimization by minimizing

$$-\sum_{i=1}^{m} w_i g_i(\phi) + \sum_{j=1}^{s} h_j^2(\phi) \qquad w_i \begin{cases} = 0 & g_i(\phi) \geq 0 \\ > 0 & g_i(\phi) < 0. \end{cases} \tag{6.35}$$

If the minimum is zero we have a feasible point. Failure to converge to zero does not necessarily mean that a feasible point does not exist.

Having obtained a feasible starting point we might decide to simply reject nonfeasible points if they are obtained during optimization. Equivalently we might set $U(\phi)$ to a most unattractive value if any violation occurs. Alternatively, we could add the term

$$\sum_{i=1}^{m} w_i g_i^2(\phi) + \sum_{j=1}^{s} h_j^2(\phi) \qquad w_i \begin{cases} = 0 & g_i(\phi) \geq 0 \\ > 0 & g_i(\phi) < 0 \end{cases} \tag{6.36}$$

to the objective function. The objective function is not penalized as long as the constraints are satisfied. This procedure does not, unfortunately, always insure a strictly feasible solution.

The simple approaches just described have other disadvantages also. Discontinuities in the new function or its derivatives may be introduced. Steep walls or valleys may be formed at the boundary of the feasible region which can drastically slow down the optimization process. A method which simply rejects nonfeasible points can easily terminate at a false minimum [6].

Sequential Unconstrained Minimization Techniques

One of the best known and most highly developed of the sequential unconstrained minimization techniques (SUMT) will be briefly outlined here [7]. Consider first the problem of minimization subject to inequality constraints defined in (6.1) and (6.2). Let

$$P(\phi, r) \triangleq U(\phi) + rG(\mathbf{g}) \tag{6.37}$$

where $G(\mathbf{g})$ is continuous for $\mathbf{g} > \mathbf{0}$ and $G(\mathbf{g}) \to \infty$ for any $g_i(\boldsymbol{\phi}) \to 0$, and where $r > 0$. Two possible candidates for $G(\mathbf{g})$ immediately suggest themselves, namely

$$G(\mathbf{g}) = \sum_{i=1}^{m} \frac{1}{g_i(\boldsymbol{\phi})}, \tag{6.38}$$

and

$$G(\mathbf{g}) = -\sum_{i=1}^{m} \log g_i(\boldsymbol{\phi}). \tag{6.39}$$

Let us denote the interior of the region R of feasible points by R°, where

$$R^\circ \triangleq \{\boldsymbol{\phi} \,|\, \mathbf{g}(\boldsymbol{\phi}) > \mathbf{0}\} \tag{6.40}$$

and

$$R \triangleq \{\boldsymbol{\phi} \,|\, \mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}\}. \tag{6.41}$$

The procedure is to select a $\boldsymbol{\phi}$ and a value of r, initially $\boldsymbol{\phi}^0 \in R^\circ$ and $r_1 > 0$, respectively, and minimize the function P of (6.37). The form of this equation is such that one would expect the minimum, namely $\check{\boldsymbol{\phi}}(r_1)$, to lie in R°. Repeat the procedure for different values of r such that

$$r_1 > r_2 > \cdots r_j > 0 \qquad \text{and} \qquad \lim_{j \to \infty} r_j = 0, \tag{6.42}$$

each minimization being started at the previous minimum. The minimization of $P(\boldsymbol{\phi}, r_2)$ would be started at $\check{\boldsymbol{\phi}}(r_1)$, and so on.

The effect of the penalty is reduced every time the parameter r is reduced, so it is reasonable to expect that, under suitable conditions

$$\lim_{j \to \infty} \check{\boldsymbol{\phi}}(r_j) = \check{\boldsymbol{\phi}}$$

since by (6.42)

$$\lim_{j \to \infty} r_j = 0$$

so that

$$\lim_{j \to \infty} U[\check{\boldsymbol{\phi}}(r_j)] = \check{U}$$

the constrained minimum. A minimum of P should always be available in R°, so any nonfeasible point that may be encountered can be rejected. This safeguard should not be overlooked, for obvious reasons.

This procedure is termed an interior point unconstrained minimization technique, and it requires an initial $\phi^0 \in R^\circ$. If one is not available, the following approach may be adopted. Let

$$S \triangleq \{s \mid g_s(\phi) \leq 0, \quad s \in \{1, 2, \ldots, m\}\}$$
$$T \triangleq \{t \mid g_t(\phi) > 0, \quad t \in \{1, 2, \ldots, m\}\}.$$

Now define a

$$P(\phi, r) = -\sum_{s \in S} g_s(\phi) + r \sum_{t \in T} G_t(g_t(\phi)) \tag{6.43}$$

to be minimized for a sequence of r values satisfying (6.42). The implications of (6.43) are that any satisfied constraints are prevented from becoming violated while an attempt to satisfy the rest is being made. As soon as any constraint is satisfied the corresponding index is transferred from S to T, and the procedure repeated. When S becomes empty we have obtained a $\phi^0 \in R^\circ$ and the solution process of the problem can commence.

To prove convergence one must invoke the requirements for convex programming (See Section 6.1). In practice, however, the conditions may be difficult to verify even if they hold. Nevertheless, the method should work successfully on a wide variety of practical problems for which convergence is not readily proved. Bad initial choices of r and ϕ will slow down convergence. Too large a value of r_1 may render the first few minima of P to be relatively independent of U, whereas too small a value may render the penalty ineffective except near the boundary where elongated valleys with steep sides are produced. Because of this and the fact that a sequence of unconstrained problems has to be solved, efficient gradient methods are generally required.

A reduction factor of 10 for the values of r is probably as good as any once the process has started. The arbitrariness of this can be somewhat alleviated by using the SUMT method without the r parameters [7, 8].

To include equality constraints the term

$$\frac{1}{r^{1/2}} \sum_{j=1}^{s} h_j^2(\phi) \tag{6.44}$$

can be added to the right hand side of (6.37). Clearly, as $r \to 0$, $\mathbf{h}(\boldsymbol{\phi})$ must approach $\mathbf{0}$ or a minimum will not be reached.

The reader is referred to a number of selected references which discuss or extend SUMT [7, 8, 9, 10]. A lucid discussion is found in Chapter 5 of Kowalik and Osborne [8].

Conditions for a Constrained Minimum

Necessary conditions which a stationary point $\boldsymbol{\phi}^\circ$ must satisfy in the problem of minimizing $U(\boldsymbol{\phi})$ subject to $\mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}$ can be formulated. Assume $U(\boldsymbol{\phi})$ and $\mathbf{g}(\boldsymbol{\phi})$ to be differentiable in the neighborhood of a feasible stationary point $\boldsymbol{\phi}^\circ$, then

$$\nabla U(\boldsymbol{\phi}^\circ) = \sum_{i=1}^{m} u_i \nabla g_i(\boldsymbol{\phi}^\circ) \tag{6.45}$$

and

$$\mathbf{u}^T \mathbf{g}(\boldsymbol{\phi}^\circ) = 0 \tag{6.46}$$

where

$$\mathbf{u} \triangleq \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \geq \mathbf{0}.$$

These necessary conditions can be interpreted as follows: $\nabla U(\boldsymbol{\phi}^\circ)$ is a nonnegative linear combination of the gradients $\nabla g_i(\boldsymbol{\phi}^\circ)$ of those constraints which are active* at $\boldsymbol{\phi}^\circ$.

Under the conditions of convex programming, i.e., if $U(\boldsymbol{\phi})$ is convex, $\mathbf{g}(\boldsymbol{\phi})$ is concave, and R° is nonempty, the conditions become sufficient for $\boldsymbol{\phi}^\circ$ to be $\check{\boldsymbol{\phi}}$, the constrained minimum. The relations (6.45) and (6.46) are called the Kuhn-Tucker relations [11]. An interpretation is sketched in Figure 6-6. Note that if we have been using a reliable optimization method, and if the relations are satisfied, we can be reasonably sure that a local minimum has been attained even if the convexity requirements are not met.

For a detailed treatment of the Kuhn-Tucker relations, including their derivation and a discussion of the constraint qualification which must also hold, the reader is referred to an appropriate book such as Zangwill [4].

* A constraint $g_i(\boldsymbol{\phi}) \geq 0$ is active at $\boldsymbol{\phi}^\circ$ if $g_i(\boldsymbol{\phi}^\circ) = 0$.

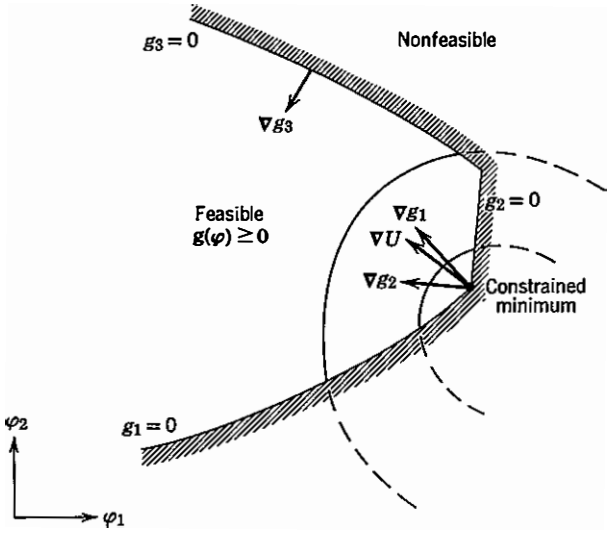

FIGURE 6-6 Sufficient conditions for a constrained minimum, $u_1 > 0$, $u_2 > 0$, $u_3 = 0$.

6.4 MINIMAX APPROXIMATION

Classically, minimax approximation (See Section 6.2 for definitions) has implied the selection of the coefficients of a suitable polynomial or rational function so that it fits some desired specification (usually continuous on a closed interval) in an optimal equal-ripple manner. The Remez method and its generalizations are notable examples of iterative processes for obtaining best approximations using polynomials and rational functions.

The number of practical problems in filter design which can be solved by the classical approach is certainly diminishing in comparison with those that need solving, notwithstanding progress in transformations in the frequency variable, Richard's transformation for transmission-line networks, and so on. This section will, therefore, emphasize less specialized methods applicable to a wider range of practical design problems. Recent references are available which discuss in detail methods well-suited to polynomials and rational functions in the context of filter design [12, 13, 14, 15].

Formulation in Terms of Inequality Constraints

Figure 6-7 illustrates a typical filter design problem. We would like to find the (constrained) parameters of a suitable network so that certain passband and stopband specifications are met or exceeded. Assuming the approximat-

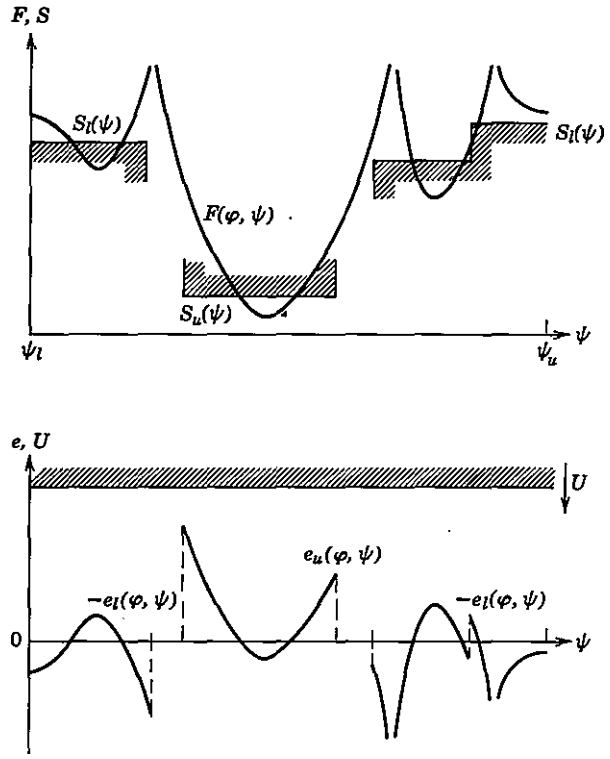

FIGURE 6-7 Typical filter design problem (the specifications are violated).

ing function and the specifications are real, let the error functions e_u and e_l be given by

$$\begin{aligned} e_u(\phi, \psi) &\triangleq w_u(\psi)[F(\phi, \psi) - S_u(\psi)] \\ e_l(\phi, \psi) &\triangleq w_l(\psi)[F(\phi, \psi) - S_l(\psi)] \end{aligned} \tag{6.47}$$

so that

$$\begin{aligned} e_{ui}(\phi) &\triangleq e_u(\phi, \psi_i), \qquad i \in I_u \\ e_{li}(\phi) &\triangleq e_l(\phi, \psi_i), \qquad i \in I_l. \end{aligned} \tag{6.48}$$

This is simply a generalization of (6.21), (6.25), and (6.26), where the symbols have the same meaning. In the present case of (6.47) and (6.48), the subscript u refers to the upper or *passband* specification, the l to the lower or *stopband* specification.

Since approximation in the time and other domains can also be formulated in these terms, ψ is used rather than frequency. Furthermore, the index sets I_u and I_l are not necessarily disjoint.

The optimization problem can now be specified as: minimize the quantity U subject to

$$\begin{aligned} U &\geq e_{ui}(\phi), \qquad i \in I_u \\ U &\geq -e_{li}(\phi), \qquad i \in I_l \end{aligned} \tag{6.49}$$

and also to all other constraints, such as on ϕ. Observe that U is an *additional* independent variable. As shown in Figure 6-7 it may be visualized as a level or ceiling which is forced down on the deviations e_u and $-e_l$.

At a minimum at least one constraint in (6.49) must be an equality. Otherwise U can be lowered without violation. Further if

1. $\check{U} < 0$, the minimum amount by which the network response exceeds the specifications is maximized;
2. $\check{U} > 0$, the maximum amount by which the network response violates the specifications is minimized.

For loss or phase equalization, or time-domain approximation, for example, we might have only one specification, namely, $S(\psi)$. To treat these special cases we simply drop the subscripts u and l in (6.47) to (6.49) and the objective is equivalent to minimizing

$$U = \max_{i \in I} |e_i(\phi)|. \tag{6.50}$$

The weighting functions in (6.47) serve the following purpose. If one is much larger than the other, it emphasizes the deviation associated with it at the expense of the rest of the response if the specifications are violated. When the specifications are satisfied (we can now set the weighting function effectively to infinity if required), effort is switched to the rest of the response.

Methods for Minimax Approximation

An approach successfully implemented in optimal filter design [16] and reviewed by Waren, Lasdon, and Suchman [17] is to use sequential unconstrained minimization. We could, for example, define

$$\begin{aligned} P(\phi, U, r) = U + r\Bigg\{ & \sum_{i \in I_u} \frac{w_{ui}}{U - e_{ui}(\phi)} \\ & + \sum_{i \in I_l} \frac{w_{li}}{U + e_{li}(\phi)} \\ & + \text{other terms} \Bigg\} \end{aligned} \tag{6.51}$$

where the other terms might include parameter constraints. Note that in (6.51) the elements of $\mathbf{g}(\phi)$ include

$$\frac{1}{w_{ui}}[U - e_{ui}(\phi)] \geq 0, \qquad i \in I_u$$

$$\frac{1}{w_{li}}[U + e_{li}(\phi)] \geq 0, \qquad i \in I_l. \tag{6.52}$$

Further, it should be remembered that U is an independent variable. The appropriate formulations described in Section 6.3 are thus applicable to minimax approximation.

Ishizaki and Watanabe [18] have described a method in many respects similar to the more recent one by Osborne and Watson [19], which applies linear programming iteratively to achieve a best approximation in the minimax sense. Let us concern ourselves with the objective suggested by (6.50). This should not, however, be taken to imply that the method is less general than the one already outlined.

Linearizing $e_i(\phi)$, which is taken as real, at some point ϕ^j the problem becomes one of minimizing U subject to

$$\begin{aligned} &\frac{1}{w_i}[U - e_i(\phi^j) - \nabla e_i^T(\phi^j)\,\Delta\phi^j] \geq 0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \ldots, n > k \\ &\frac{1}{w_i}[U + e_i(\phi^j) + \nabla e_i^T(\phi^j)\,\Delta\phi^j] \geq 0 \end{aligned} \tag{6.53}$$

and other (linearized) constraints. Noting that the variables for linear programming should all be nonnegative, and imposing a rather practical constraint that the elements of ϕ should not change sign we have the linear program in $\mathbf{x} \triangleq [x_1\, x_2 \cdots x_{k+1}]^T$ such as to

$$\text{minimize } U = x_{k+1}$$

subject to

$$\pm\{e_i(\phi^j) + \nabla e_i^T(\phi^j) \begin{bmatrix} \phi_1^j x_1 - \phi_1^j \\ \phi_2^j x_2 - \phi_2^j \\ \vdots \\ \phi_k^j x_k - \phi_k^j \end{bmatrix}\} \leq x_{k+1}, \; i = 1, 2, \ldots, n > k \tag{6.54}$$

$$\mathbf{x} \geq \mathbf{0}$$

where

$$x_i \triangleq \frac{\Delta\phi_i^j}{\phi_i^j} + 1, \qquad i = 1, 2, \ldots, k.$$

The solution produces a direction given by $\Delta\phi^j$. Next we find α^j such that $\max_i |e_i(\phi^j + \alpha^j \Delta\phi^j)|$ is a minimum, set $\phi^{j+1} = \phi^j + \alpha^j \Delta\phi^j$ and repeat the process. For conditions for convergence the reader is referred to the original papers [18, 19]. Other linearized constraints can also be considered [14, 15]. Clearly such an approach is directly applicable to linear functions such as polynomials, for which $k + 1$ equal extrema results at the optimum.

Bandler, Srinivasan, and Charalambous [20] have described a descent type of algorithm for minimax approximation which also employs linear programming. Basically, the algorithm attempts to find a locally optimal downhill direction for the problem of minimizing U, where

$$U = \max_{i \in I} f_i(\phi), \tag{6.55}$$

where the $f_i(\phi)$ are real nonlinear differentiable functions generally. Linearizing $f_i(\phi)$ and letting

$$J \triangleq \{i \,|\, f_i(\phi) = \max_i f_i(\phi), \qquad i \in I\} \tag{6.56}$$

we can obtain, at some feasible point ϕ^j, the first-order changes

$$\Delta f_i(\phi^j) = \nabla f_i^T(\phi^j)\, \Delta\phi^j, \qquad i \in J. \tag{6.57}$$

In order for $\Delta\phi^j$ to define a descent direction for $\max_{i \in I} f_i(\phi)$ we must have

$$\nabla f_i^T(\phi^j)\, \Delta\phi^j < 0, \qquad i \in J.$$

Consider

$$\Delta\phi^j = -\sum_{i \in J} \alpha_i^j \nabla f_i(\phi^j) \tag{6.58}$$

$$\sum_{i \in J} \alpha_i^j = 1 \tag{6.59}$$

$$\alpha_i^j \geq 0, \qquad i \in J, \tag{6.60}$$

which suggests the linear program:

$$\text{maximize } \alpha_{r+1}^j \geq 0 \tag{6.61}$$

subject to

$$-\nabla f_i^T(\phi^j) \sum_{i \in J} \alpha_i^j \nabla f_i(\phi^j) \leq -\alpha_{r+1}^j, \qquad i \in J \tag{6.62}$$

plus (6.59) and (6.60), where it is assumed that J has r elements.

Observe that J should be nonempty, and that if J has only one element, we obtain the steepest descent direction for the corresponding maximum of the $f_i(\phi)$. The solution to the linear program provides $\Delta\phi^j$. We then find γ^j corresponding to the minimum value of $\max_{i \in I} f_i(\phi^j + \gamma^j \Delta\phi^j)$. ϕ^{j+1} is set to $\phi^j + \gamma^j \Delta\phi^j$ and the procedure is repeated.

In practice, we will not have a set of $f_i(\phi)$ identically equal to the maximum value. An appropriate tolerance must, therefore, be introduced into (6.56) and a more suitable selection procedure for the elements of J formulated. For further details the original paper should be consulted [20]. It can be proved that the algorithm will, if correctly implemented, converge to the minimax solution.

Example 6-1. Figure 6-8 shows an example of minimax approximation [21]. The objective was to find

$$\check{U} = \min_{\varphi} \left\{ \max_{[f_l,\, f_u]} |\rho(\phi, f)| \right\}$$

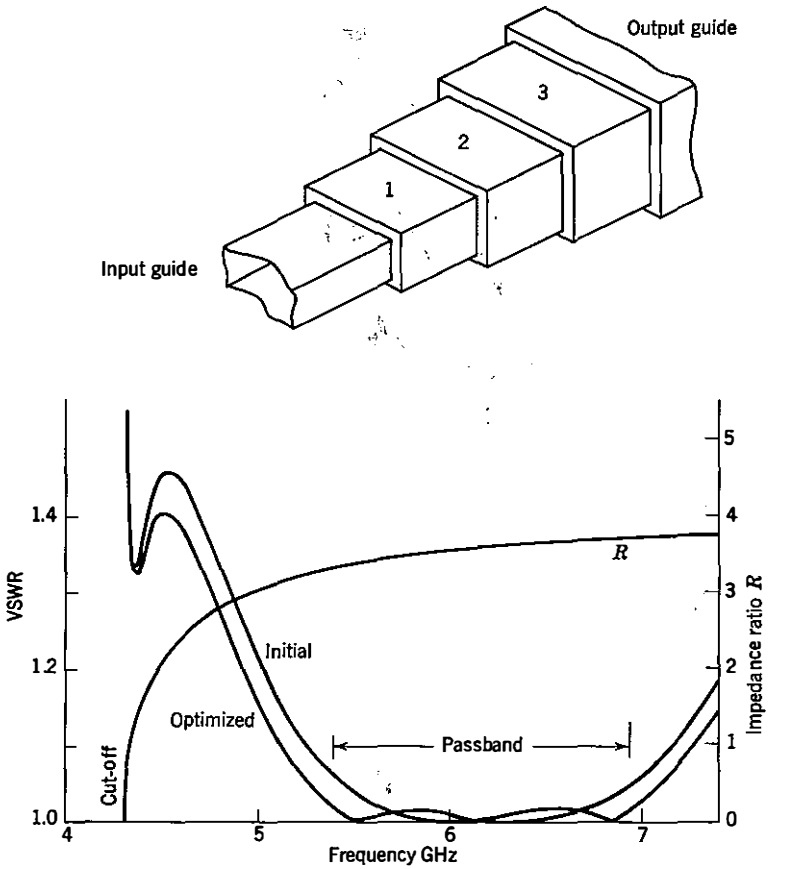

FIGURE 6-8 Example of constrained minimax approximation.

for the 3-section inhomogeneous rectangular waveguide impedance transformer, where ρ is the reflection coefficient, and f is frequency in GHz. The parameters ϕ to be varied were the actual geometrical dimensions of the sections. The lower and upper band edges were $f_l = 5.4$ GHz and $f_u = 6.95$ GHz. It should be noted that (1) both input and output waveguides had different cut-off frequencies so that an exact synthesis was not possible, (2) severe constraints were placed on the parameters for a variety of physical reasons, (3) discontinuity susceptances could be taken directly into account, and (4) the razor search method [22] (See Section 6.6) was employed. The reader is referred to Bandler [21] for further details of this type of problem and for some other numerical results.

Example 6-2. Let us consider in a little more detail, the optimization of a seven-section cascaded transmission-line filter of the type shown in Figure 6-9. It is terminated at each end by

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

where f is frequency in GHz and $f_c = 2.077$ GHz. The frequency variation of the terminations is thus like that of rectangular waveguides operating in the H_{10} mode with cut-off frequency 2.077 GHz. This interesting problem was

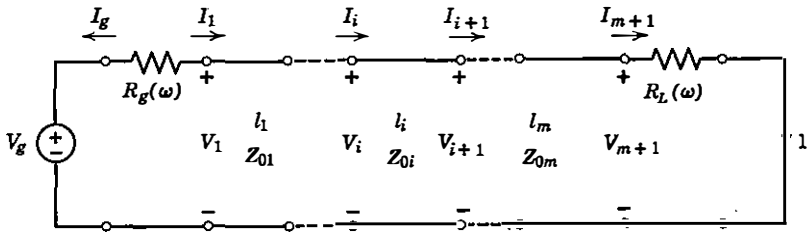

FIGURE 6-9 Cascaded transmission-line filter between frequency-variable resistors.

previously considered by Carlin and Gupta [23]. All section lengths were kept fixed at 1.5 cm so that the maximum stopband insertion loss would occur at about 5 GHz. The passband 2.16 to 3 GHz was selected, for which a maximum of 0.4 dB loss was specified. The solution obtained by the method of Carlin and Gupta was used as the initial design as shown by Figure 6-10.

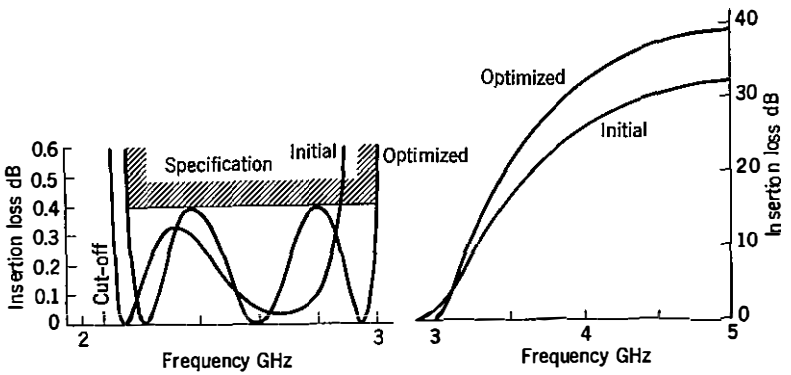

FIGURE 6-10 Comparison between the initial and optimized responses of the filter of Figure 6-9.

As optimized by Bandler and Lee-Chan [24], the problem was to minimize $\max_i f_i(\phi)$ where

$$f_i(\phi) = \begin{cases} \frac{1}{2}[|\rho_i(\phi)|^2 - r^2] & \text{in the passband} \\ \frac{1}{2}[1 - |\rho_i(\phi)|^2] & \text{in the stopband} \end{cases}$$

$$\phi = \begin{bmatrix} Z_{01} \\ Z_{02} \\ \vdots \\ Z_{07} \end{bmatrix}$$

r is the reflection coefficient magnitude corresponding to an insertion loss of 0.4 dB, and $\rho_i(\phi)$ is the reflection coefficient of the filter at the ith frequency point. In particular, 22 uniformly spaced frequencies were selected from the passband and a single frequency, namely, 5 GHz for the stopband. The appropriately optimized response is shown in Figure 6-10. These results have also been reproduced by the method of Bandler, Srinivasan, and Charalambous [20], using

$$\phi = \begin{bmatrix} Z_{01} \\ Z_{02} \\ Z_{03} \\ Z_{04} \end{bmatrix}$$

and letting $Z_{05} = Z_{03}$, $Z_{06} = Z_{02}$, $Z_{07} = Z_{01}$.

The method used to analyze the filter at each frequency is suggested in Figure 6-9. A load current of 1 amp was assumed and a simple $ABCD$ matrix analysis was carried out to find all the other voltage and current variables shown (V_g will, of course, be generally complex and frequency dependent in

this case). The appropriate partial derivatives were obtained from *one* such analysis per frequency point, using the adjoint network method (Section 6.9).

Conditions for a Minimax Optimum

To derive some insight into the necessary conditions which a stationary point ϕ° must satisfy in a minimax approximation problem [25], let us reduce it to the form

$$\text{minimize } U = \phi_{k+1} \tag{6.63}$$

subject to constraints of the form

$$\phi_{k+1} \geq f_i(\phi), \qquad i = 1, 2, \ldots, m. \tag{6.64}$$

Rewriting the constraints as

$$g_i(\phi) \triangleq \phi_{k+1} - f_i(\phi) \geq 0, \qquad i = 1, 2, \ldots, m, \tag{6.65}$$

allows us to apply the Kuhn-Tucker relations (Section 6.3). Assuming U and the $f_i(\phi)$ to be differentiable in the neighborhood of ϕ°, we have at $\phi = \phi^\circ$

$$\begin{bmatrix} \nabla U \\ \dfrac{\partial U}{\partial \phi_{k+1}} \end{bmatrix} = \sum_{i=1}^{m} u_i \begin{bmatrix} \nabla \\ \dfrac{\partial}{\partial \phi_{k+1}} \end{bmatrix} (\phi_{k+1} - f_i(\phi))$$

$$\mathbf{u}^T \mathbf{g} = 0, \tag{6.66}$$

where $\mathbf{u}$ is defined by (6.46). But

$$\begin{aligned} \nabla U &= \nabla \phi_{k+1} = \mathbf{0} \\ \frac{\partial U}{\partial \phi_{k+1}} &= 1 \\ \frac{\partial f_i(\phi)}{\partial \phi_{k+1}} &= 0 \end{aligned} \tag{6.67}$$

everywhere. Furthermore, at least one constraint must be an equality. For convenience, assume the first m_0 constraints are equalities. Then

$$\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \sum_{i=1}^{m_0} u_i \begin{bmatrix} -\nabla f_i(\phi^\circ) \\ 1 \end{bmatrix} \tag{6.68}$$

since

$$u_i = 0, \qquad i = m_0 + 1, m_0 + 2, \ldots, m.$$

Alternatively, the necessary conditions may be written as

$$\sum_{i=1}^{m_0} u_i \nabla f_i(\phi^\circ) = \mathbf{0}$$

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

$$u_i \geq 0, \qquad i = 1, 2, \ldots, m_0. \tag{6.69}$$

An interpretation of these relations is sketched in Figure 6-11. Under the conditions of convex programming, the $f_i(\phi)$ would have to be convex, and the conditions become sufficient for ϕ° to be $\tilde{\phi}$, the minimax optimum. Often m_0 will be equal to $k + 1$, but this is not a general requirement. The reader should observe the correspondence between (6.58) to (6.60) for $\Delta\phi^j = \mathbf{0}$ with (6.69). More insight into these relations, in particular as they relate to filter problems, should be gained by referring to Bandler [25].

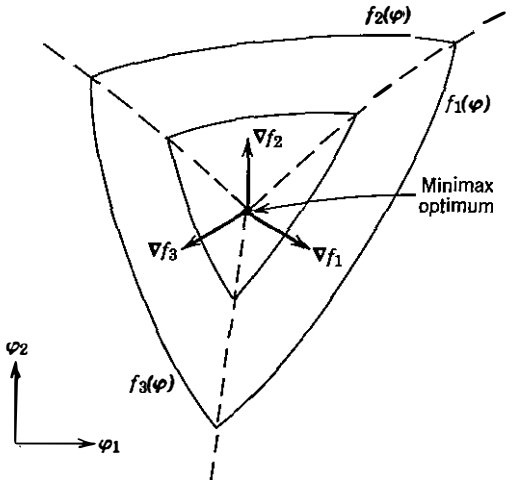

FIGURE 6-11 Sufficient conditions for a minimax optimum, $u_1 > 0$, $u_2 > 0$, $u_3 > 0$.

6.5 ONE-DIMENSIONAL SEARCH METHODS OF MINIMIZATION

Three main possible reasons spring to mind for investigating the optima of functions of one variable. The obvious one is that this might be the problem we are given. The second is that the multidimensional method we are using may call for a one-dimensional search for a minimum in some feasible downhill direction.* The third is that we may be dealing with an approximation problem for which the extrema of the error function are required during an optimization process.

*That is, in a feasible direction for which U is decreasing.

Powerful methods are available for functions known to be unimodal on an interval. We can broadly distinguish two classes, first the *elimination* methods which chop away subintervals not containing the optimum in an efficient manner with no assumptions except unimodality; second the *approximation* or *interpolation* methods which assume the function is smooth and well-represented by a low-order polynomial near the optimum.

Without loss of generality and to simplify discussions we will assume we have a function U of a single variable ϕ.

Elimination Methods

At the start of the jth iteration of a search for a minimum of a unimodal function suppose we have an *interval of uncertainty* I^j where, referring to Figure 6-12,

$$I^j \triangleq u - l \tag{6.70}$$

with $\phi_u^j = u$, $\phi_l^j = l$. Further, we have two interior points $\phi_a^j = a$ and $\phi_b^j = b$ at which we have evaluated the objective function. Let $U(a)$ and $U(b)$ be denoted U_a and U_b, respectively. Note that we take

$$l < a < b < u. \tag{6.71}$$

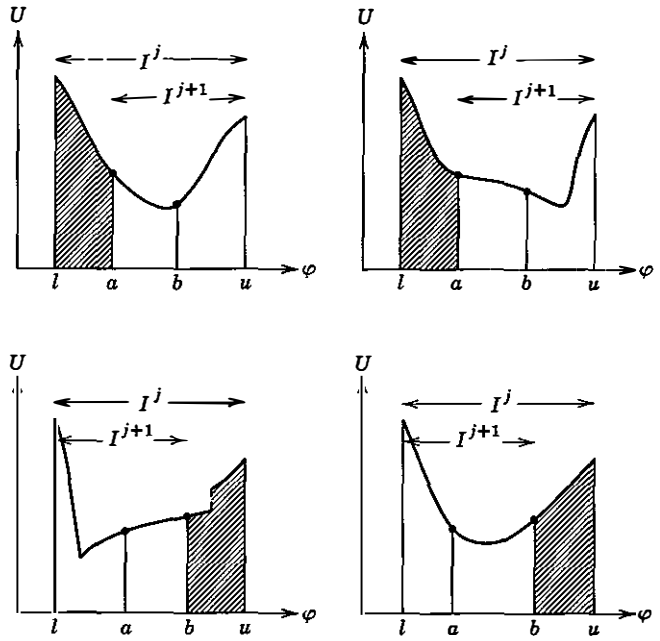

FIGURE 6-12 Reduction of interval of uncertainty.

Two conclusions can be drawn:

1. If $U_a > U_b$, the minimum lies in $[a, u]$ and $I^{j+1} = u - a$.
2. If $U_a < U_b$, the minimum lies in $[l, b]$ and $I^{j+1} = b - l$.

The difference between two well-known and efficient methods, the Fibonacci search and the Golden Section search, is in how these interior points are located. Let us discuss the slightly less efficient but simpler Golden Section search method. The reader is referred elsewhere for more detailed accounts of the various methods [6, 8, 26].

Whatever the outcome of comparing U_a and U_b, we want

$$I^{j+1} = u - a = b - l, \tag{6.72}$$

which is achieved by symmetrical placement of a and b on $[l, u]$. We want to minimize I^{j+1} and use one of the points in our new interval again which leads to

$$I^{j+2} = u - b = a - l. \tag{6.73}$$

Combining (6.70) to (6.73)

$$I^j = I^{j+1} + I^{j+2}. \tag{6.74}$$

To reduce the interval of uncertainty by a constant factor τ at each iteration:

$$\frac{I^j}{I^{j+1}} = \frac{I^{j+1}}{I^{j+2}} = \tau. \tag{6.75}$$

Equations (6.74) and (6.75) lead to

$$\tau^2 = \tau + 1, \tag{6.76}$$

the solution of relevance being $\tau = 1/2(1 + \sqrt{5}) \cong 1.618034$. The division of a line according to (6.74) and (6.75) is called the Golden Section of a line.

At the jth iteration of this scheme

$$\begin{aligned} \phi_a^j &= \frac{1}{\tau^2} I^j + \phi_l^j \\ \phi_b^j &= \frac{1}{\tau} I^j + \phi_l^j \end{aligned} \qquad j = 1, 2, 3, \ldots. \tag{6.77}$$

Note that each iteration except the first involves only one function evaluation due to symmetry. Depending on the outcome of the jth iteration, the appropriate quantities are set for the $(j + 1)$th iteration and the procedure repeated. After n function evaluations

$$\frac{I^1}{I^n} = \tau^{n-1}. \tag{6.78}$$

For a desired accuracy of σ, n should be chosen such that

$$\tau^{n-2} < \frac{\phi_u^1 - \phi_l^1}{\sigma} \leq \tau^{n-1}. \tag{6.79}$$

It is readily shown that Golden Section provides an interval of uncertainty only about 17% greater than Fibonacci search for large n. The latter method also has the disadvantage that the number of function evaluations needs to be fixed in advance.

It is also possible to construct a scheme described by Temes [15] whereby the initial interval of uncertainty does not have to be fixed in advance. This scheme has been used with the method of Bandler, Srinivasan and Charalambous [20] (Section 6.4).

Interpolation Methods

There are several interpolation methods, including quadratic and cubic, which are available [8, 26, 27, 28]. A rather straightforward method suggested by Davies, Swann, and Campey [8, 26] will be described here. The method does not require a unimodal interval containing the minimum to be known in advance, but the unimodality restriction should hold.

Evaluate $U^i \triangleq U(\phi^0 + \alpha^i s)$ for

$$\alpha^0 = 0$$

$$\alpha^i = \sum_{j=1}^{i} 2^{j-1}\,\delta, \qquad i = 1, 2, \ldots \tag{6.80}$$

where s determines the negative gradient direction, i.e.,

$$s \triangleq \left. \frac{-\dfrac{\partial U}{\partial \phi}}{\left|\dfrac{\partial U}{\partial \phi}\right|} \right|_{\phi=\phi^0} \tag{6.81}$$

and $\delta > 0$, e.g., 1% of ϕ^0, is a convenient increment. Thus α^i is a positive step in the direction of decreasing U. When, for some i,

$$U(\alpha^i) > U(\alpha^{i-1}), \tag{6.82}$$

evaluate U^{i+1} at

$$\alpha^{i+1} = \alpha^{i-1} + (\alpha^{i-1} - \alpha^{i-2}). \tag{6.83}$$

It should be clear that we now have four uniformly spaced points on the α axis, namely, α^{i-2}, α^{i-1}, α^{i+1}, and α^i in order of increasing α. Note that $i \geq 2$.

$$\begin{aligned} &\text{If} \quad U(\alpha^{i+1}) < U(\alpha^{i-1}), \quad \text{let} \quad a = \alpha^{i-1},\ b = \alpha^{i+1},\ c = \alpha^i. \\ &\text{If} \quad U(\alpha^{i+1}) > U(\alpha^{i-1}), \quad \text{let} \quad a = \alpha^{i-2},\ b = \alpha^{i-1},\ c = \alpha^{i+1}. \end{aligned} \tag{6.84}$$

It is easily shown that the minimum of a quadratic fitted at a, b, and c is at

$$\alpha_{\min} = b + \frac{(b-a)(U_a - U_c)}{2(U_a - 2U_b + U_c)}. \tag{6.85}$$

Evaluation of U at $\alpha_{\min}$ gives the estimate of the minimum and completes one stage of the method. A new stage with reduced δ can be started at b or $\alpha_{\min}$, whichever corresponds to a smaller U.

6.6 DIRECT SEARCH METHODS OF MINIMIZATION

Direct search methods as interpreted by this author are methods which do not depend explicitly on evaluation or estimation of the gradient vector of the objective function. Such methods have enjoyed fairly wide use in network optimization [6, 21, 22, 29]. To what extent they will remain competitive, however, in the light of currently available methods of evaluating derivatives (See Section 6.9), remains to be seen.

One of the simplest methods is the one-at-a-time method. As Figure 6-13 shows, this process basically consists of letting one parameter vary until no improvement is obtained, and then another one, and so on. Progress is fairly slow on valleys not oriented in the direction of any coordinate axis.

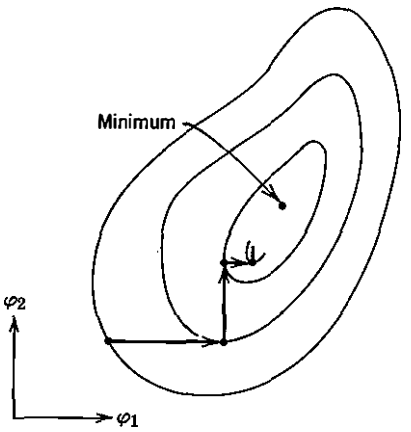

FIGURE 6-13 One-at-a-time search.

Obviously we need to consider more efficient methods. Two widely used methods will be reviewed, namely the pattern search method of Hooke and Jeeves [30] and the simplex method of Nelder and Mead [31]. Other well-known methods are Rosenbrock's method [32], the Powell-Zangwill method [28, 33], and the method of Davies, Swann, and Campey [26]. These methods are discussed in some of the general references [1, 6, 8, 26, 34].

Pattern Search

An advantage the *pattern search* method has over the one-at-a-time method is that it attempts to detect the presence of a valley and align a direction of search along it. The tactics employed by pattern search will be explained by means of the example shown in Figure 6-14.

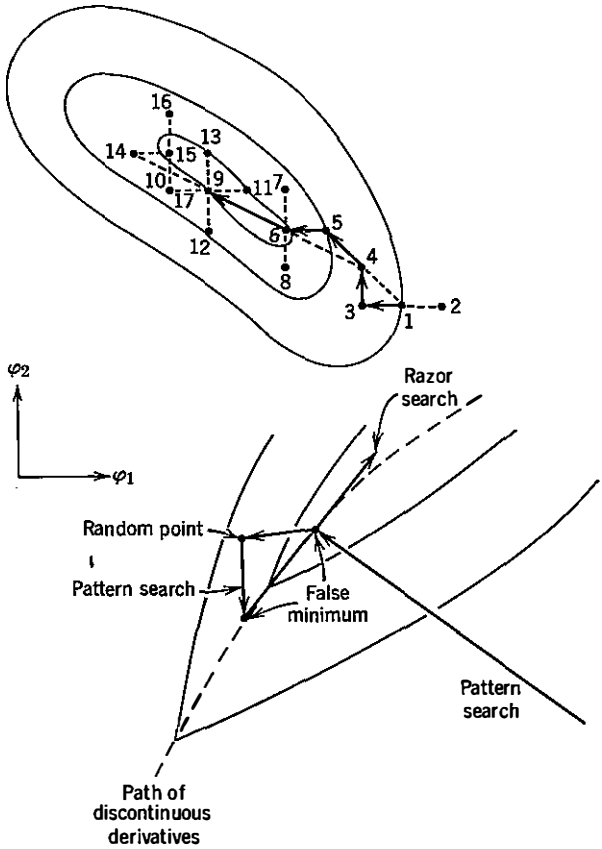

FIGURE 6-14 Following valleys by pattern search and razor search.

The first *base point* $\mathbf{b}^1$ is taken as the starting point $\boldsymbol{\phi}^1$. A series of *exploratory moves* from $\boldsymbol{\phi}^1$ is initiated to find the second base point. In the example, ϕ_1 is incremented leading us to $\boldsymbol{\phi}^2$. Now $U^2 > U^1$ so $\boldsymbol{\phi}^2$ is rejected, and ϕ_1 is incremented in the opposite direction to $\boldsymbol{\phi}^3$. Exploration with ϕ_1 is over. $U^3 < U^1$ so $\boldsymbol{\phi}^3$ is retained and exploration with ϕ_2 begins. $U^4 < U^3$ so $\boldsymbol{\phi}^4$ is retained in place of $\boldsymbol{\phi}^3$. The first set of exploratory moves is complete, and so $\boldsymbol{\phi}^4$ becomes the second base point $\mathbf{b}^2$. In the expectation that our

success would be repeated we make a *pattern move* to $\phi^5 = 2\mathbf{b}^2 - \mathbf{b}^1$, which is in the direction $\mathbf{b}^2 - \mathbf{b}^1$. By another set of exploratory moves we try to find the most promising point in the vicinity of ϕ^5. Here, this point is ϕ^6 which becomes the third base point $\mathbf{b}^3$, since $U^6 < U^4$. The search continues with a pattern move in the direction $\mathbf{b}^3 - \mathbf{b}^2$ to ϕ^9.

The pattern direction is destroyed when a pattern move followed by exploration fails, as around ϕ^{14}. The strategy is to return to the previous base point. If the exploratory moves around the base point fail, as around ϕ^9, the parameter increments are reduced and the procedure is restarted at that point. The search may be terminated either when the parameter increments fall below prescribed levels or the number of function evaluations or running time have reached upper limits.

The *razor search* method of Bandler and Macdonald [22] is a development of pattern search suited to direct optimization in the minimax sense without using derivatives. The name was suggested by the fact that "razor sharp" valleys are, in general, generated by an attempt to minimize functions of the form of (6.50). Paths of discontinuous derivatives are found along the bottom of such valleys, as indicated in Figures 6-11 and 6-14.

An investigation of the behavior of pattern search in the optimization of cascaded noncommensurate transmission lines acting as impedance transformers between resistive terminations was carried out [29]. It was observed that pattern search failed only when a sharp valley whose contours lay entirely within a quadrant of the coordinate axes was encountered. In that case no improvement was possible by searching parallel to these axes.

The razor search method makes a random move from a point where pattern search fails (assuming a false minimum) and uses pattern search to return to the path of discontinuous derivatives. (See Figure 6-14.) When pattern search fails again, an attempt is made to establish a pattern in the apparent downhill direction and resume with pattern search. The results shown in Figure 6-8 were produced by the razor search method [21].

An observation worth making here is that manual network optimization in the minimax sense, using an interactive system and employing, say, the one-at-a-time method, can easily terminate at a false minimum. A false minimum in the present context is a point representing a possibly equal-ripple response but which is not a local optimum in the minimax sense.

The Simplex Method

In simplex methods of nonlinear optimization, the objective function is evaluated at the $k + 1$ vertices of a *simplex* in k-dimensional space. In two dimensions, for example, we would have a triangle, for three dimensions a tetrahedron. An attempt is then made to replace the point with the greatest objective function value by another point.

A method having very desirable valley-following properties is the one due to Nelder and Mead [31]. The basic move is to *reflect* the point having the greatest function value in the centroid of the simplex formed by the remaining points. If the reflected point results in a function value lower than the current lowest, an *expansion* is attempted. Otherwise the point is retained if it results in a function value lower than the second highest. *Contraction* is attempted if reflection fails. Finally, *shrinking* of the simplex about the vertex corresponding to the lowest function value occurs following an unsuccessful attempt at contraction. Some of these moves are illustrated in Figure 6-15.

An example of the simplex strategy is shown in Figure 6-16. Observe that ϕ^4, ϕ^6, ϕ^8, ϕ^9, and ϕ^{10} have resulted from reflection; ϕ^5 from expansion; and ϕ^7 and ϕ^{11} from contraction. The reader should follow the strategy through carefully to ensure his understanding of it. Its desirable valley-following properties result from its ability to align elongated simplexes in the

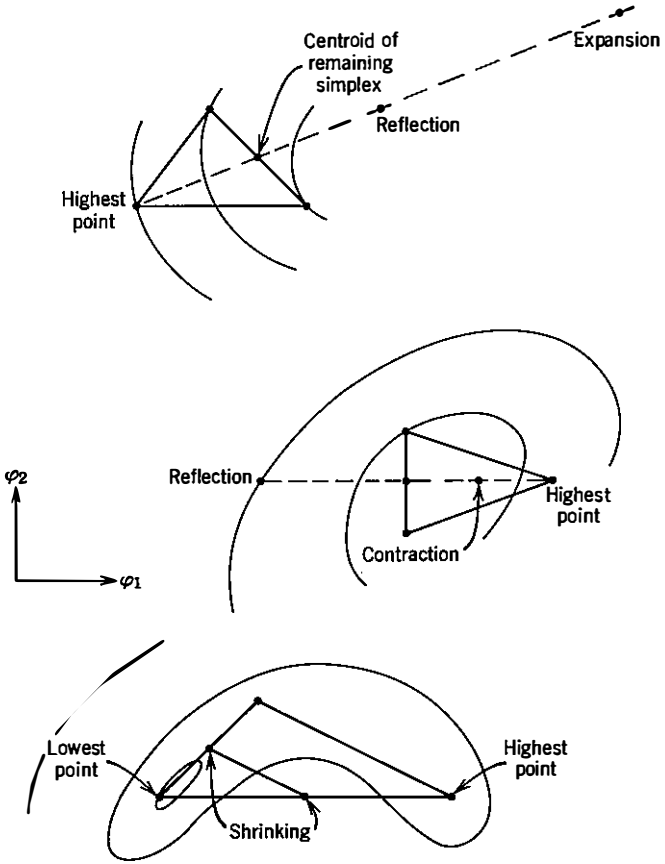

FIGURE 6-15 Examples of moves made by the simplex method.

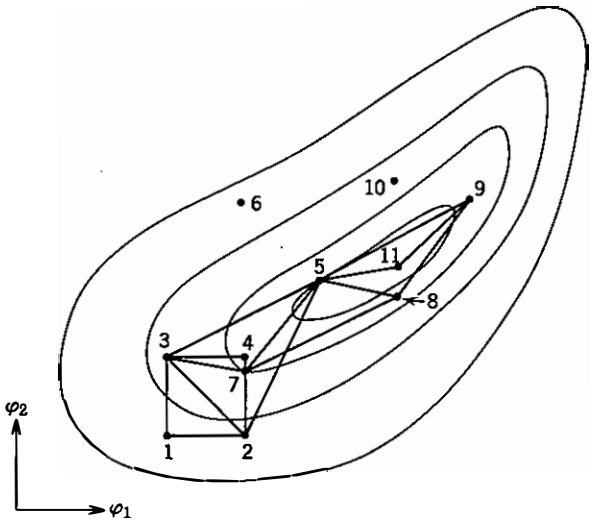

FIGURE 6-16 Optimization by the simplex method.

direction of the valleys. In particular, repeated success, for example, if a long straight valley is being followed, tends to increase the size of the moves, whereas repeated failure, for example, if a bend in the valley is encountered, tends to cause a decrease in the size of the moves.

It has been claimed to the author on a number of occasions that, unlike some other direct search methods, the simplex method can be successfully employed for minimax approximation. In the author's experience the simplex method is no less infallible than pattern search, for example. The principal fallacy in the argument is the assumption that, if the method requires no derivative information, it can necessarily handle problems with discontinuous derivatives.

6.7 GRADIENT METHODS OF MINIMIZATION

We turn our attention now to a class of minimization methods which require derivatives. By and large the most efficient algorithms currently available rely on evaluation of the gradient vector [1, 5, 6, 8, 14, 15, 26, 27, 35].

Steepest Descent

At the jth iteration of most gradient methods, we proceed to

$$\boldsymbol{\phi}^{j+1} = \boldsymbol{\phi}^j + \alpha^j \mathbf{s}^j \tag{6.86}$$

where $\mathbf{s}^j$ is (hopefully) a downhill direction of search and $\alpha^j > 0$ is a scale factor chosen to minimize $U(\boldsymbol{\phi}^j + \alpha^j \mathbf{s}^j)$. One-dimensional minimization methods suitable for this purpose were discussed in Section 6.5.

The most obvious choice for $\mathbf{s}^j$ is the *steepest descent* direction at $\boldsymbol{\phi}^j$, defined as follows. Referring back to (6.11), we note that a first-order change in the objective function is given by

$$\Delta U = \nabla U^T \Delta\boldsymbol{\phi}. \tag{6.87}$$

If $\Delta\boldsymbol{\phi} = \alpha\mathbf{s}$, where $\alpha > 0$ is fixed and $\|\mathbf{s}\| = 1^*$, then it is easy to show that the $\mathbf{s}$ minimizing ΔU is

$$\mathbf{s} = -\frac{\nabla U}{\|\nabla U\|} \tag{6.88}$$

The $\mathbf{s}$ in (6.88) is the negative of the *normalized gradient* vector. Although $-\nabla U/\|\nabla U\|$ provides the greatest local change, success of the steepest descent method is highly dependent on scaling. As Figure 6-17 shows, the

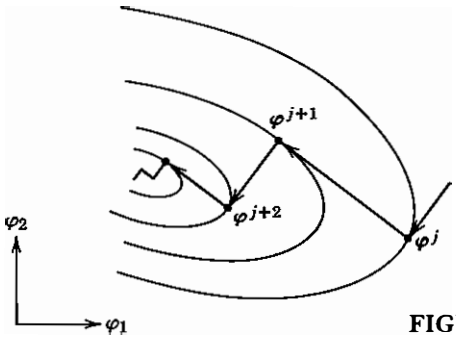

FIGURE 6-17 A steepest-descent strategy.

first few iterations may give good reduction in U, but subsequently the method usually deteriorates rapidly into oscillations, and progress becomes very slow.

The Newton Method

This method was already mentioned in Section 6.2 in the context of solution of nonlinear equations. Differentiating the Taylor series (6.11)

$$\nabla U(\boldsymbol{\phi} + \Delta\boldsymbol{\phi}) = \nabla U(\boldsymbol{\phi}) + \mathbf{H}\,\Delta\boldsymbol{\phi} + \cdots. \tag{6.89}$$

For $\boldsymbol{\phi} + \Delta\boldsymbol{\phi}$ to be the minimizing point $\bar{\boldsymbol{\phi}}$, $\nabla U(\boldsymbol{\phi} + \Delta\boldsymbol{\phi})$ should be $\mathbf{0}$ so that, neglecting higher-order terms,

$$\Delta\boldsymbol{\phi} = -\mathbf{H}^{-1}\,\nabla U. \tag{6.90}$$

* The expression $\|\cdot\|$ is the *Euclidean* norm. It has the form of (6.23) with $p = 2$.

This incremental change takes us to the minimum in only one iteration if we are dealing with a quadratic function (Section 6.2). It is instructive to compare (6.90) with (6.19).

When U is not quadratic, we could try the iterative scheme

$$\boldsymbol{\phi}^{j+1} = \boldsymbol{\phi}^j - \mathbf{H}^{-1} \nabla U^j \tag{6.91}$$

where $\mathbf{H}^{-1}$ is the inverse of the Hessian matrix at the jth iteration. This scheme has, however, several disadvantages. $\mathbf{H}$ must be positive definite otherwise divergence could occur. In particular, $-\mathbf{H}^{-1} \nabla U^j$ might not point downhill. To counteract these possibilities, the modification

$$\boldsymbol{\phi}^{j+1} = \boldsymbol{\phi}^j - \alpha^j \mathbf{H}^{-1} \nabla U^j \tag{6.92}$$

can be employed where α^j is chosen to minimize U^{j+1}. This might also be ineffective: α^j may have to be negative; $\mathbf{H}$ may be locally singular. Finally, the computation of $\mathbf{H}$ and its inverse are time consuming.

Conjugate Directions

Certain gradient methods which exploit the properties of *conjugate directions* associated with quadratic functions and do not explicitly evaluate $\mathbf{H}$ or its inverse are highly effective. Before discussing them let us define conjugate directions.

The directions $\mathbf{u}_i$ and $\mathbf{u}_j$ are said to be conjugate with respect to a positive definite matrix $\mathbf{A}$ if

$$\mathbf{u}_i^T \mathbf{A} \mathbf{u}_j = 0, \qquad i \neq j. \tag{6.93}$$

In Figure 6-18, a two-dimensional interpretation of conjugate directions is given. Methods which generate such directions will minimize a quadratic

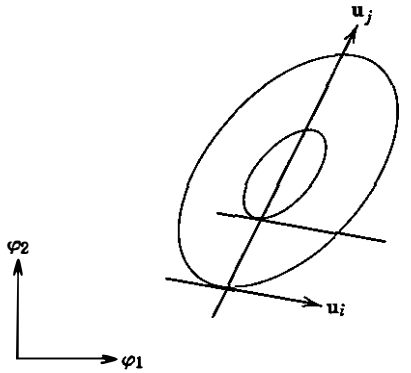

FIGURE 6-18 An illustration of two conjugate directions.

function in a finite number of iterations. It is evident that one linear minimization along each direction in turn locates the minimum.

Three well-known methods which use conjugate directions are the *conjugate gradient* method described by Fletcher and Reeves [35], the Fletcher-Powell-Davidon method [27], and the Powell-Zangwill method [28, 33] which does not require derivatives (See also references [1, 8]).

The Conjugate Gradient Method

The direction of search $\mathbf{s}^j$ is given by [35]

$$\mathbf{s}^j = -\nabla U^j + \beta^j \mathbf{s}^{j-1} \tag{6.94}$$

where

$$\beta^j = \frac{(\nabla U^j)^T \nabla U^j}{(\nabla U^{j-1})^T \nabla U^{j-1}}, \tag{6.95}$$

and, initially, $\beta^0 = 0$. Thus the first iteration is in the direction of steepest descent. Apart from round-off errors, the procedure will terminate at the minimum of a quadratic in at most k iterations. In general, however, it is recommended that $k+1$ iterations be completed before restarting the procedure.

The Fletcher-Powell-Davidon Method

Redefining $\mathbf{H}$ as any positive definite matrix, we have [27]

$$\mathbf{s}^j = -\mathbf{H}^j \nabla U^j. \tag{6.96}$$

Note that $\mathbf{H}^j$ is the jth approximation to the *inverse* of the Hessian matrix. Initially, $\mathbf{H}^0$ is the unit matrix, and again we have the steepest descent direction.

$\mathbf{H}$ is continually updated using first derivative information such that

$$\boldsymbol{\phi}^{j+1} - \boldsymbol{\phi}^j = \mathbf{H}^{j+1} \mathbf{g}^j \tag{6.97}$$

where

$$\mathbf{g}^j = \nabla U^{j+1} - \nabla U^j.$$

The following updating procedure is used:

$$\mathbf{H}^{j+1} = \mathbf{H}^j + \frac{\Delta\boldsymbol{\phi}^j \Delta\boldsymbol{\phi}^{jT}}{\Delta\boldsymbol{\phi}^{jT} \mathbf{g}^j} - \frac{\mathbf{H}^j \mathbf{g}^j \mathbf{g}^{jT} \mathbf{H}^j}{\mathbf{g}^{jT} \mathbf{H}^j \mathbf{g}^j} \tag{6.98}$$

where

$$\Delta\boldsymbol{\phi}^j = \alpha^j \mathbf{s}^j,$$

and α^j is found by a one-dimensional search (Section 6.5).

Fletcher and Powell prove by induction that if $\mathbf{H}^j$ is positive definite then $\mathbf{H}^{j+1}$ is also positive definite. $\mathbf{H}^0$, being the unit matrix, is clearly positive definite. On a quadratic function it is further proved that $\mathbf{H}^k$ is the inverse of the Hessian matrix and $\nabla U^k = \mathbf{0}$, apart from round-off errors. Both the proof of convergence and success in practice depend on accurate location of the minimum in the linear searches. If necessary, $\mathbf{H}$ may be reset to the unit matrix.

This method is still generally acknowledged to be the best general purpose gradient optimization method.

6.8 LEAST *p*th APPROXIMATION

The material in this section could equally well have been treated under gradient methods. It is useful, however, to distinguish between these problems since special techniques are available for least *p*th approximation.

For objective functions in the form of (6.29) and (6.30) we can write

$$\nabla U = \int_{\psi_l}^{\psi_u} \mathrm{Re}\{p\,|e(\phi, \psi)|^{p-2} e^*(\phi, \psi)\, \nabla e(\phi, \psi)\}\, d\psi \tag{6.99}$$

for the continuous case and

$$\nabla U = \sum_{i \in I} \mathrm{Re}\{p\,|e_i(\phi)|^{p-2} e_i^*(\phi)\, \nabla e_i(\phi)\} \tag{6.100}$$

for the discrete case. If the appropriate derivatives, namely ∇e, are available, we could proceed to optimize with a suitable gradient method (Section 6.7).

In more complicated situations we can envisage a linear combination of functions in the form (6.29) and (6.30), for example,

$$U = \alpha_1 U_1 + \alpha_2 U_2 + \cdots. \tag{6.101}$$

Simultaneous approximation of more than one response specification might be posed in this way (See Section 6.9). The factors α_1, α_2, etc. would be given values commensurate with the importance of U_1, U_2, etc.

Temes and Zai [15, 36] have extended the well-known least squares method of Gauss [6, 8, 14] to a *least* p*th method*. Since the former method falls out as a special case, the latter method will be briefly described. For definiteness, assume the objective function is of the form (with real $e_i(\phi)$)

$$U = \sum_{i=1}^{n} [e_i(\phi)]^p \tag{6.102}$$

where $n > k$ and p is any positive even integer. Then

$$\nabla U = \sum_{i=1}^{n} p e_i^{p-1}\, \nabla e_i \tag{6.103}$$

and

$$\mathbf{H} = \nabla(\nabla U)^T = \sum_{i=1}^{n} [p e_i^{p-1} \nabla(\nabla e_i)^T + p(p-1) e_i^{p-2} \nabla e_i (\nabla e_i)^T]. \quad (6.104)$$

Now assume that the first term may be neglected in comparison with the second. This really corresponds to a linearization of $e_i(\phi)$. Then

$$\mathbf{H} \approx \sum_{i=1}^{n} p(p-1) e_i^{p-2} \nabla e_i (\nabla e_i)^T.$$

This can be rewritten as

$$\mathbf{H} \approx p(p-1)\mathbf{A}^T\mathbf{B}\mathbf{A} \quad (6.105)$$

where

$$\mathbf{A} \triangleq [\nabla e_1 \quad \nabla e_2 \quad \cdots \quad \nabla e_n]^T$$

and

$$\mathbf{B} \triangleq \begin{bmatrix} e_1^{p-2} & 0 & \cdots & 0 \\ 0 & e_2^{p-2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & e_n^{p-2} \end{bmatrix}.$$

Letting

$$\boldsymbol{\epsilon} \triangleq [e_1^{p-1} \quad e_2^{p-1} \quad \cdots \quad e_n^{p-1}]^T,$$

(6.103) becomes

$$\nabla U = p\mathbf{A}^T\boldsymbol{\epsilon}. \quad (6.106)$$

Using the step given by the Newton method (6.90),

$$\Delta\phi = -(p-1)^{-1}(\mathbf{A}^T\mathbf{B}\mathbf{A})^{-1}\mathbf{A}^T\boldsymbol{\epsilon}. \quad (6.107)$$

Under suitable conditions, it can be shown that $\Delta\phi$ points in the downhill direction. The modified Newton procedure

$$\phi^{j+1} = \phi^j - \alpha^j(p-1)^{-1}(\mathbf{A}^T\mathbf{B}\mathbf{A})^{-1}\mathbf{A}^T\boldsymbol{\epsilon} \quad (6.108)$$

is recommended where α^j is chosen to minimize U^{j+1}.

Damping techniques similar to those used in the Gauss method are applicable [8, 14]. Define, for example,

$$U = \sum_{i=1}^{n} [e_i(\phi)]^p + \lambda\, \Delta\phi^T\, \Delta\phi. \quad (6.109)$$

Then

$$\mathbf{H} \approx p(p-1)\mathbf{A}^T\mathbf{B}\mathbf{A} + 2\lambda\mathbf{I}_k \quad (6.110)$$

and

$$\Delta\phi = -p[p(p-1)\mathbf{A}^T\mathbf{B}\mathbf{A} + 2\lambda\mathbf{I}_k]^{-1}\mathbf{A}^T\epsilon \tag{6.111}$$

It may be shown that the convergence and downhill properties are preserved and that for $\lambda > 0$ the step is no larger than the undamped step. As $\lambda \to 0$ the process is undamped, while for $\lambda \to \infty$ the step is in the steepest descent direction. The introduction of α^j to permit a linear search as in (6.108) is also possible.

Example 6-3. An example of least pth approximation [37] compared with minimax approximation is depicted in Figure 6-19. The structure is the

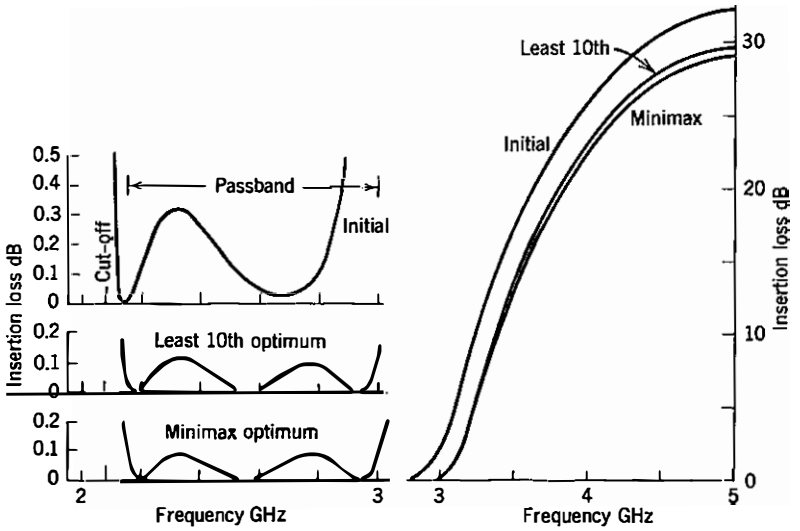

FIGURE 6-19 Example of least 10th approximation compared with minimax approximation in optimizing the passband of the filter of Figure 6-9.

seven-section cascade of transmission lines acting as a filter discussed in Section 6.4. The problem here was to see how small the passband insertion loss could be made under the constraints of the problem (if R_g and R_L were frequency independent, or the lengths were allowed to vary, the answer would be trivial).

A least pth objective function was set up with $p = 10$, using 51 uniformly spaced points in the passband. The objective function was of the form

$$U = \sum_{i=1}^{n} \frac{1}{p} |\rho_i(\phi)|^p.$$

The Fletcher-Powell-Davidon method (Section 6.7) was used, the required first derivatives being obtained from *one* network analysis using the adjoint network method (Section 6.9).

Compare the almost equal-ripple passband response obtained with a maximum insertion loss of about 0.1 dB with the equal-ripple response (maximum insertion loss 0.086 dB) produced by minimax approximation. The latter solution was obtained by Bandler and Lee-Chan [24] using a gradient algorithm with quadratic interpolation used to locate the ripple extrema.

The main conclusion to be reached from this example is that acceptable results can be achieved with relatively moderate values of p. Unless special precautions are taken to avoid ill-conditioning, the use of values of p much greater than 10 is discouraged.

6.9 THE ADJOINT NETWORK METHOD OF GRADIENT EVALUATION

The *adjoint network method* can be used to great advantage in evaluating the gradient vector of objective functions related to gain, insertion loss, reflection coefficient, or any other desired response. A very broad class of networks can be treated by this method. As will be seen, no more than two complete network analyses are required to evaluate the gradient vector regardless of the number of variable parameters.

Director and Rohrer have discussed the concept of the adjoint network and indicated its relevance to automated design of networks in the frequency and time domains [38, 39]. In the frequency domain [39], they considered reciprocal and nonreciprocal, lumped, linear, and time-invariant elements. We will restrict ourselves here to the frequency domain, review Director and Rohrer's results, and extend them to least pth and minimax approximation. Some uniformly distributed elements will also be included [37, 40].

Adjoint Networks And Network Sensitivities

Let

$$\mathbf{v} \triangleq \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_b \end{bmatrix} \tag{6.112}$$

contain all the branch voltages in a network and

$$\mathbf{i} \triangleq \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_b \end{bmatrix} \tag{6.113}$$

contain all the corresponding branch currents (using associated reference directions*). $\mathbf{v}$ and $\mathbf{i}$ must satisfy Kirchhoff's voltage and current laws, respectively. Then Tellegen's theorem states [41]

$$\mathbf{v}^T\mathbf{i} = 0. \tag{6.114}$$

As long as the topologies are the same, $\mathbf{v}$ can refer to one network and $\mathbf{i}$ to another (See the example in Figure 6-20). Let us, therefore, imagine we have

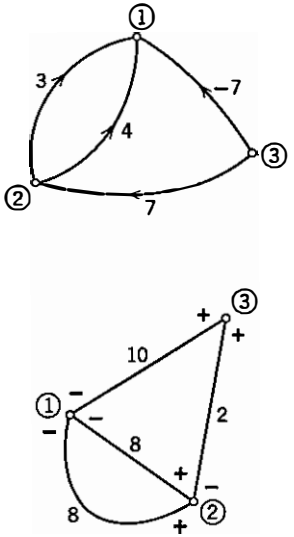

FIGURE 6-20 Illustration of Tellegen's theorem applied to two networks of the same topology. Observe that

$$\mathbf{v}^T\mathbf{i} = 24 + 32 - 70 + 14 = 0.$$

Since the nature of the elements is immaterial, they are replaced by branches.

two networks, the original one which is to be optimized and a topologically equivalent adjoint network. As mentioned earlier we will confine ourselves to a consideration of linear, time-invariant networks in the *frequency domain.* Variables V and I will thus denote phasors associated with the original

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

network, and $\hat{V}$ and $\hat{I}$ the corresponding phasors associated with the adjoint network. By Tellegen's theorem

$$\begin{aligned} \mathbf{V}_B^T \hat{\mathbf{I}}_B &= 0 \\ \mathbf{I}_B^T \hat{\mathbf{V}}_B &= 0 \end{aligned} \tag{6.115}$$

where the subscript B implies that the associated vectors contain all corresponding complex branch voltages and currents. Perturbing elements in the original network we have

$$\Delta \mathbf{V}_B^T \hat{\mathbf{I}}_B = 0 \tag{6.116a}$$

$$\Delta \mathbf{I}_B^T \hat{\mathbf{V}}_B = 0 \tag{6.116b}$$

since Kirchhoff's voltage and current laws must also be applicable to $\Delta \mathbf{V}_B$ and $\Delta \mathbf{I}_B$. Subtracting (6.116b) from (6.116a)

$$\Delta \mathbf{V}_B^T \hat{\mathbf{I}}_B - \Delta \mathbf{I}_B^T \hat{\mathbf{V}}_B = 0. \tag{6.117}$$

Figure 6-21 shows N-port original and adjoint elements characterized in terms of open-circuit impedance matrices $\mathbf{Z}$ and $\hat{\mathbf{Z}}$, respectively. Letting $\mathbf{V}$, $\mathbf{I}$,

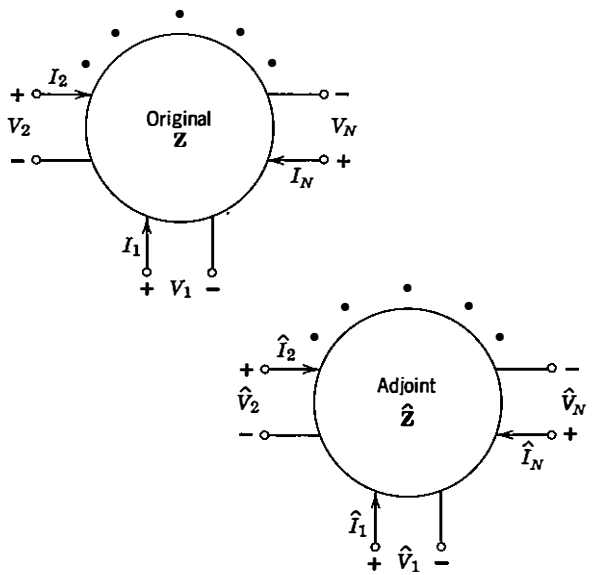

FIGURE 6-21 Original and adjoint elements represented by impedance matrices. In general, many such elements suitably connected form the original and adjoint networks.

$\hat{\mathbf{V}}$, and $\hat{\mathbf{I}}$ denote N-element vectors containing the relevant port variables

$$\mathbf{V} = \mathbf{ZI} \tag{6.118}$$

$$\hat{\mathbf{V}} = \hat{\mathbf{Z}}\hat{\mathbf{I}}. \tag{6.119}$$

Perturbing the parameters in the original element and neglecting higher-order terms

$$\Delta\mathbf{V} = \Delta\mathbf{ZI} + \mathbf{Z}\,\Delta\mathbf{I}. \tag{6.120}$$

As indicated by Figure 6-22, the port variables can be thought of as equivalent branch variables, so that, substituting (6.120) into (6.117) we see that

$$(\mathbf{I}^T\,\Delta\mathbf{Z}^T + \Delta\mathbf{I}^T\mathbf{Z}^T)\hat{\mathbf{I}} - \Delta\mathbf{I}^T\hat{\mathbf{V}}$$

reduces to

$$\mathbf{I}^T\,\Delta\mathbf{Z}^T\hat{\mathbf{I}} \tag{6.121}$$

if

$$\hat{\mathbf{Z}} \equiv \mathbf{Z}^T. \tag{6.122}$$

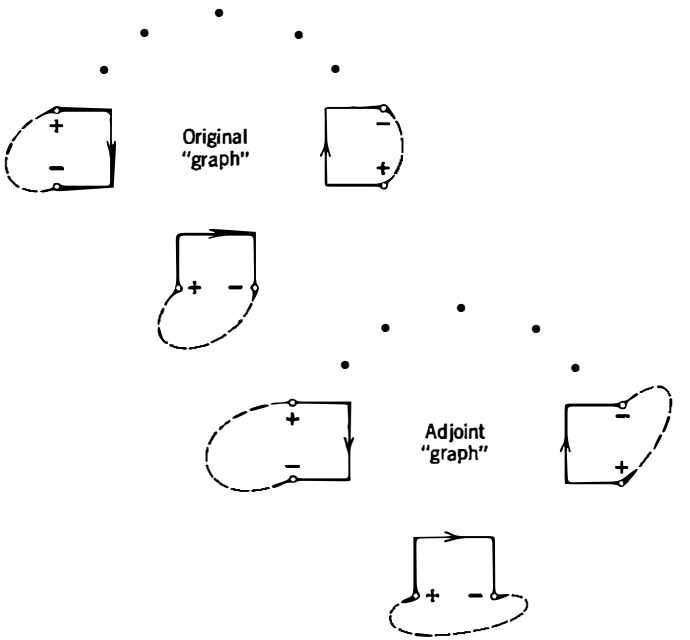

FIGURE 6-22 Representation of the elements of Figure 6-21 for application of Tellegen's theorem. Some ith equivalent branch might consist of an impedance z_{ii} in series with voltage generators of value $z_{ij}I_j$, $j = 1, 2, \ldots$. See, for example, Figure 6-25.

This defines the adjoint element. Observe that expression (6.121), the only term in (6.117) relating to the N-port element, does not contain $\Delta\mathbf{I}$ or $\Delta\mathbf{V}$. Further, note that the adjoint of a reciprocal element is identical to the original, since $\mathbf{Z}^T = \mathbf{Z}$.

Next define voltage and current excitation vectors and response vectors as in Figure 6-23. In keeping with the present notation, the hat " $\wedge$ " will distinguish the corresponding quantities for the adjoint network. Terms in (6.117) associated with the excitations and responses are

$$\Delta\mathbf{V}_V^T\hat{\mathbf{I}}_V - \Delta\mathbf{I}_V^T\hat{\mathbf{V}}_V + \Delta\mathbf{V}_I^T\hat{\mathbf{I}}_I - \Delta\mathbf{I}_I^T\hat{\mathbf{V}}_I$$

which reduces to

$$-\Delta\mathbf{I}_V^T\hat{\mathbf{V}}_V + \Delta\mathbf{V}_I^T\hat{\mathbf{I}}_I \tag{6.123}$$

since $\Delta\mathbf{V}_V$ and $\Delta\mathbf{I}_I$ become zero when the excitations are held fixed.

Clearly, any network may be thought of as consisting of the interconnection of a number of multiport elements. Thus, several terms of the form of expression (6.121) can appear in (6.117).

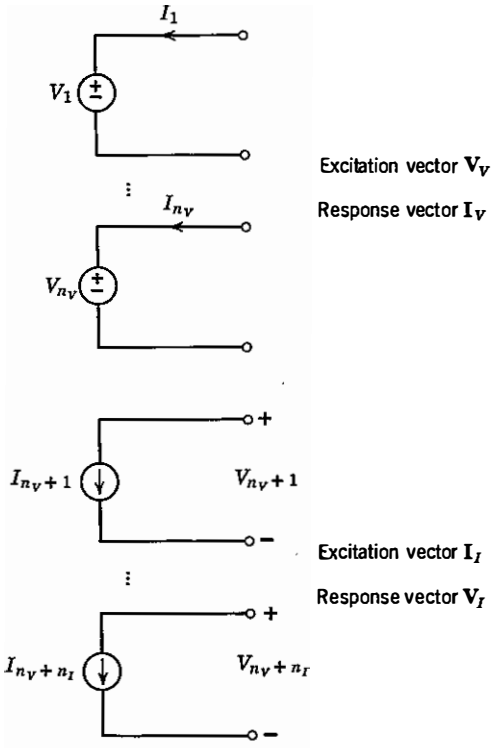

FIGURE 6-23 Port excitation and response vectors.

For an admittance matrix representation we can show that

$$-\mathbf{V}^T \Delta \mathbf{Y}^T \hat{\mathbf{V}} \tag{6.124}$$

corresponds to expression (6.121). $\mathbf{Y}^T$ is the admittance matrix of the adjoint. Things are slightly more complicated for the hybrid matrix. If we take

$$\begin{bmatrix} \mathbf{I}_a \\ \mathbf{V}_b \end{bmatrix} = \begin{bmatrix} \mathbf{Y} & \mathbf{A} \\ \mathbf{M} & \mathbf{Z} \end{bmatrix} \begin{bmatrix} \mathbf{V}_a \\ \mathbf{I}_b \end{bmatrix}, \tag{6.125}$$

then the corresponding relation for the adjoint is

$$\begin{bmatrix} \hat{\mathbf{I}}_a \\ \hat{\mathbf{V}}_b \end{bmatrix} = \begin{bmatrix} \mathbf{Y}^T & -\mathbf{M}^T \\ -\mathbf{A}^T & \mathbf{Z}^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_a \\ \hat{\mathbf{I}}_b \end{bmatrix}. \tag{6.126}$$

The expression corresponding to (6.121) can be shown to be

$$[\mathbf{V}_a^T \quad \mathbf{I}_b^T] \begin{bmatrix} -\Delta \mathbf{Y}^T & \Delta \mathbf{M}^T \\ -\Delta \mathbf{A}^T & \Delta \mathbf{Z}^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_a \\ \hat{\mathbf{I}}_b \end{bmatrix}. \tag{6.127}$$

To summarize the results of the above discussion, we note that (6.117) can be written in the form

$$\Delta \mathbf{I}_V^T \hat{\mathbf{V}}_V - \Delta \mathbf{V}_I^T \hat{\mathbf{I}}_I = \mathbf{G}^T \Delta \boldsymbol{\phi} \tag{6.128}$$

where $\mathbf{G}$ is a vector of sensitivity components related to the adjustable parameters of the network, namely $\boldsymbol{\phi}$. Equation (6.128) basically relates changes in port responses due to changes in element values.

Figure 6-24 shows the results of a direct application of the formulas (6.121) and (6.124) to three commonly used elements. Table 6-2 summarizes sensitivity expressions for some commonly used lumped and distributed elements. An element consisting of a single branch is simply viewed as a one-port element.

Consider, for example, a uniformly distributed line (Figure 6-25) having characteristic impedance Z_0 and electrical length θ. Since the element is reciprocal:

$$\hat{\mathbf{Z}} = \mathbf{Z}^T = \mathbf{Z} = Z_0 \begin{bmatrix} \coth\theta & \operatorname{csch}\theta \\ \operatorname{csch}\theta & \coth\theta \end{bmatrix}. \tag{6.129}$$

Invoking expression (6.121) we obtain

$$\begin{aligned} \mathbf{I}^T \Delta \mathbf{Z}^T \hat{\mathbf{I}} &= \mathbf{I}^T \left(\Delta Z_0 \begin{bmatrix} \coth\theta & \operatorname{csch}\theta \\ \operatorname{csch}\theta & \coth\theta \end{bmatrix} - \frac{Z_0 \, \Delta\theta}{\sinh\theta} \begin{bmatrix} \operatorname{csch}\theta & \coth\theta \\ \coth\theta & \operatorname{csch}\theta \end{bmatrix} \right)^T \hat{\mathbf{I}} \\ &= \left(\frac{\Delta Z_0}{Z_0} \mathbf{Z}\mathbf{I} - \frac{\Delta\theta}{\sinh\theta} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{Z}\mathbf{I} \right)^T \hat{\mathbf{I}} = \frac{\Delta Z_0}{Z_0} \mathbf{V}^T \hat{\mathbf{I}} - \frac{\Delta\theta}{\sinh\theta} \mathbf{V}^T \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \hat{\mathbf{I}}. \end{aligned} \tag{6.130}$$

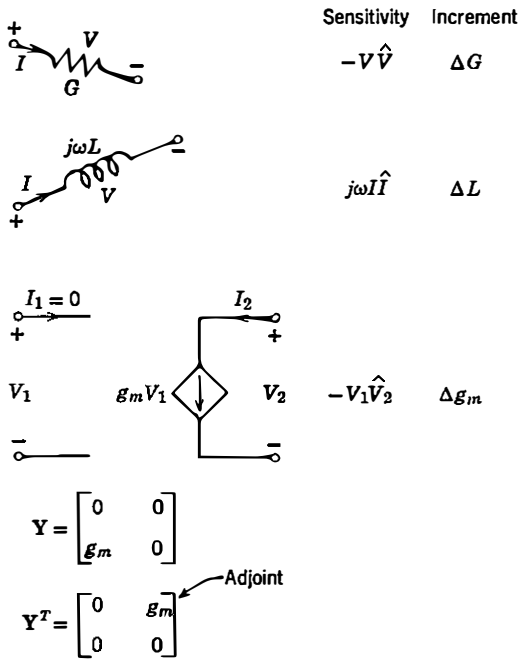

FIGURE 6-24 Sensitivities for three common elements: a resistor of conductance G, an inductor of inductance L and a voltage-controlled current source with transfer conductance g_m.

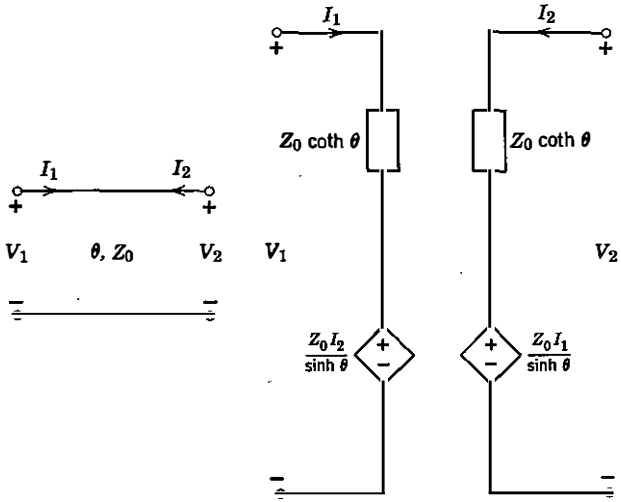

FIGURE 6-25 Uniform line and convenient representation.

Observe that the sensitivities shown in Table 6-2 depend on 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 to evaluate them. Note that if there is no excitation at a port, the appropriate source is set to zero. If the response at a port is of no interest, the appropriate adjoint excitation should be zero. Elements or parameters not to be varied are simply not represented in $\mathbf{G}$ or ϕ.

An Application to Minimax Approximation

Consider the situation depicted in Figure 6-26. Suppose we are given the problem: minimize a positive independent variable U subject to

$$U \geq f(\phi, \omega_i) \triangleq |\rho(\phi, j\omega_i)|^2, \qquad \omega_i \in \Omega_d \tag{6.131}$$

where ρ is the input reflection coefficient, and Ω_d is a discrete set of frequencies in the band of interest. This problem then is effectively to minimize the maximum magnitude of the reflection coefficient over a band [37, 42]. Now

$$\rho = \frac{Z_{in} - R_g}{Z_{in} + R_g} = 1 - \frac{2R_g}{Z_{in} + R_g} = 1 + \frac{2R_g I_g}{V_g} \tag{6.132}$$

so that

$$\begin{aligned} \nabla f(\phi, \omega_i) &= \mathrm{Re}\{2\rho^*(\phi, j\omega_i)\, \nabla \rho(\phi, j\omega_i)\} \\ &= \mathrm{Re}\left\{\frac{4R_g}{V_g}\, \rho^*(\phi, j\omega_i)\, \nabla I_g(\phi, j\omega_i)\right\}. \end{aligned} \tag{6.133}$$

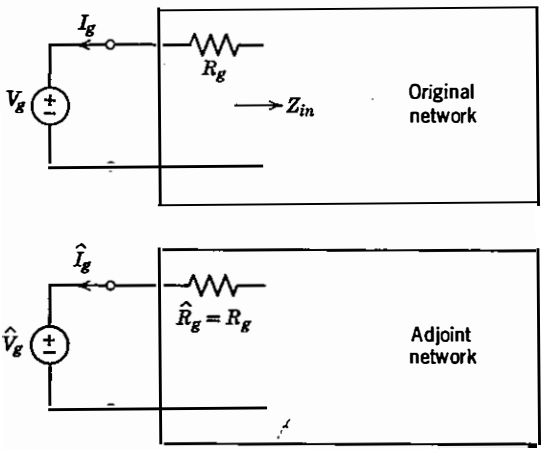

FIGURE 6-26 Possible original and adjoint networks for design on the reflection coefficient basis.

TABLE 6-2 Sensitivity Expressions for Some Lumped and Distributed Elements

Element	Equation		Sensitivity (component of **G**)	Increment (component of $\Delta\Phi$)
	Original	Adjoint		
Resistor	$V = RI$ $I = GV$	$\hat{V} = R\hat{I}$ $\hat{I} = G\hat{V}$	$I\hat{I}$ $-V\hat{V}$	ΔR ΔG
Inductor	$V = j\omega LI$ $I = \frac{1}{j\omega}\Gamma V$	$\hat{V} = j\omega L\hat{I}$ $\hat{I} = \frac{1}{j\omega}\Gamma\hat{V}$	$j\omega I\hat{I}$ $-\frac{1}{j\omega}V\hat{V}$	ΔL $\Delta\Gamma$
Capacitor	$V = \frac{1}{j\omega}SI$ $I = j\omega CV$	$\hat{V} = \frac{1}{j\omega}S\hat{I}$ $\hat{I} = j\omega C\hat{V}$	$\frac{1}{j\omega}I\hat{I}$ $-j\omega V\hat{V}$	ΔS ΔC
Transformer	$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} 0 & n \\ -n & 0 \end{bmatrix}\begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$	$\begin{bmatrix} \hat{V}_1 \\ \hat{I}_2 \end{bmatrix} = \begin{bmatrix} 0 & n \\ -n & 0 \end{bmatrix}\begin{bmatrix} \hat{I}_1 \\ \hat{V}_2 \end{bmatrix}$	$V_2\hat{I}_1 + I_1\hat{V}_2$	Δn
Gyrator	$\mathbf{V} = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}\mathbf{I}$	$\hat{\mathbf{V}} = \begin{bmatrix} 0 & -\alpha \\ \alpha & 0 \end{bmatrix}\hat{\mathbf{I}}$	$I_1\hat{I}_2 - I_2\hat{I}_1$	$\Delta\alpha$
Voltage controlled voltage source	$\begin{bmatrix} I_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mu & 0 \end{bmatrix}\begin{bmatrix} V_1 \\ I_2 \end{bmatrix}$	$\begin{bmatrix} \hat{I}_1 \\ \hat{V}_2 \end{bmatrix} = \begin{bmatrix} 0 & -\mu \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{V}_1 \\ \hat{I}_2 \end{bmatrix}$	$V_1\hat{I}_2$	$\Delta\mu$
Voltage controlled current source	$\mathbf{I} = \begin{bmatrix} 0 & 0 \\ g_m & 0 \end{bmatrix}\mathbf{V}$	$\hat{\mathbf{I}} = \begin{bmatrix} 0 & g_m \\ 0 & 0 \end{bmatrix}\hat{\mathbf{V}}$	$-V_1\hat{V}_2$	Δg_m

Table 6-2 *Continued*

Current controlled voltage source	$\mathbf{V} = \begin{bmatrix} 0 & 0 \\ r_m & 0 \end{bmatrix}\mathbf{I}$	$\hat{\mathbf{V}} = \begin{bmatrix} 0 & r_m \\ 0 & 0 \end{bmatrix}\hat{\mathbf{I}}$	$I_1 \hat{I}_2$	Δr_m
Current controlled current source	$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}\begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$	$\begin{bmatrix} \hat{V}_1 \\ \hat{I}_2 \end{bmatrix} = \begin{bmatrix} 0 & -\beta \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{I}_1 \\ \hat{V}_2 \end{bmatrix}$	$-I_1 \hat{V}_2$	$\Delta\beta$
Short circuited uniformly distributed line	$V = Z_0 \tanh \theta I$	$\hat{V} = Z_0 \tanh \theta \hat{I}$	$\tanh \theta I \hat{I}$ $Z_0 \operatorname{sech}^2 \theta I \hat{I}$	ΔZ_0 $\Delta\theta$
	$I = Y_0 \coth \theta V$	$\hat{I} = Y_0 \coth \theta \hat{V}$	$-\coth \theta V \hat{V}$ $Y_0 \operatorname{csch}^2 \theta V \hat{V}$	ΔY_0 $\Delta\theta$
Open circuited uniformly distributed line	$V = Z_0 \coth \theta I$	$\hat{V} = Z_0 \coth \theta \hat{I}$	$\coth \theta I \hat{I}$ $-Z_0 \operatorname{csch}^2 \theta I \hat{I}$	ΔZ_0 $\Delta\theta$
	$I = Y_0 \tanh \theta V$	$\hat{I} = Y_0 \tanh \theta \hat{V}$	$-\tanh \theta V \hat{V}$ $-Y_0 \operatorname{sech}^2 \theta V \hat{V}$	ΔY_0 $\Delta\theta$
Uniformly distributed line	$\mathbf{V} = Z_0 \begin{bmatrix} \coth\theta & \operatorname{csch}\theta \\ \operatorname{csch}\theta & \coth\theta \end{bmatrix}\mathbf{I}$	same as original network equation but with $\hat{\mathbf{V}}$ and $\hat{\mathbf{I}}$ replacing $\mathbf{V}$ and $\mathbf{I}$ respectively	$\frac{1}{Z_0}\mathbf{V}^T\hat{\mathbf{I}}$ $-\frac{1}{\sinh\theta}\mathbf{V}^T\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\hat{\mathbf{I}}$	ΔZ_0 $\Delta\theta$
	$\mathbf{I} = Y_0 \begin{bmatrix} \coth\theta & -\operatorname{csch}\theta \\ -\operatorname{csch}\theta & \coth\theta \end{bmatrix}\mathbf{V}$		$-\frac{1}{Y_0}\mathbf{I}^T\hat{\mathbf{V}}$ $-\frac{1}{\sinh\theta}\mathbf{I}^T\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\hat{\mathbf{V}}$	ΔY_0 $\Delta\theta$

Table 6-2 *Continued*

Lossless transmission line	$\mathbf{V} = -jZ_0\begin{bmatrix}\cot\beta l & \csc\beta l\\ \csc\beta l & \cot\beta l\end{bmatrix}\mathbf{I}$	same as original network equation but with $\hat{\mathbf{V}}$ and $\hat{\mathbf{I}}$ replacing $\mathbf{V}$ and $\mathbf{I}$ respectively	$\frac{1}{Z_0}\mathbf{V}^T\hat{\mathbf{I}}$	ΔZ_0
			$-\frac{\beta}{\sin\beta l}\mathbf{V}^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{\mathbf{I}}$	Δl
	$\mathbf{I} = -jY_0\begin{bmatrix}\cot\beta l & -\csc\beta l\\ -\csc\beta l & \cot\beta l\end{bmatrix}\mathbf{V}$		$-\frac{1}{Y_0}\mathbf{I}^T\hat{\mathbf{V}}$	ΔY_0
			$-\frac{\beta}{\sin\beta l}\mathbf{I}^T\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\hat{\mathbf{V}}$	Δl
Uniform RC line	as for uniformly distributed line with $Z_0 = \sqrt{\frac{R}{sC}}$ and $\theta = \sqrt{sRC}$		$\frac{1}{2R}\mathbf{V}^T\begin{bmatrix}1 & -\frac{\theta}{\sinh\theta}\\ -\frac{\theta}{\sinh\theta} & 1\end{bmatrix}\hat{\mathbf{I}}$	ΔR
			$-\frac{1}{2C}\mathbf{V}^T\begin{bmatrix}1 & \frac{\theta}{\sinh\theta}\\ \frac{\theta}{\sinh\theta} & 1\end{bmatrix}\hat{\mathbf{I}}$	ΔC

From (6.128)

$$\Delta I_g \hat{V}_g = \mathbf{G}^T \Delta\phi. \tag{6.134}$$

Hence

$$\Delta I_g = \left(\frac{1}{\hat{V}_g}\mathbf{G}^T\right)\Delta\phi = \nabla I_g^T \Delta\phi,$$

so that

$$\nabla I_g = \frac{1}{\hat{V}_g}\mathbf{G}, \tag{6.135}$$

and, finally,

$$\nabla f(\phi, \omega_i) = \mathrm{Re}\left\{\frac{4R_g}{V_g \hat{V}_g}\rho^*(\phi, j\omega_i)\mathbf{G}(\phi, j\omega_i)\right\}. \tag{6.136}$$

Observe that we are at liberty to set $\hat{V}_g = V_g$. If the original network is reciprocal so that the adjoint network is identical to the original, we need perform only one network analysis to obtain $\nabla f(\phi, \omega_i)$.

An Application to Least *p*th Approximation

It can be shown that if there are n_V independent voltage sources and n_I independent current sources

$$\mathbf{G} = \sum_{i=1}^{n_V} \hat{V}_i \nabla I_i - \sum_{i=n_V+1}^{n_V+n_I} \hat{I}_i \nabla V_i. \tag{6.137}$$

Suppose we are given the objective function [37, 39, 42],

$$U = \sum_{i=1}^{n_V+n_I} \int_{\Omega} |e_i(\phi, j\omega)|^p \, d\omega, \tag{6.138}$$

where Ω defines a frequency range of interest and where $e_i(\phi, j\omega)$ is an ith function of the form of (6.21) such that

$$F_i(\phi, j\omega) \triangleq \begin{cases} I_i(\phi, j\omega), & i = 1, 2, \ldots, n_V \\ V_i(\phi, j\omega), & i = n_V + 1, \ldots, n_V + n_I. \end{cases} \tag{6.139}$$

Equation (6.138) thus represents a summation of functions of the form of (6.101). The specified functions $S_i(j\omega)$ correspond to desired response currents and voltages. In general, $F_i(\phi, j\omega)$, $S_i(j\omega)$, and hence $e_i(\phi, j\omega)$ may be complex. Now, from (6.99)

$$\nabla U = \sum_{i=1}^{n_V+n_I} \int_{\Omega} \mathrm{Re}\{p\,|e_i(\phi, j\omega)|^{p-2} w_i(\omega) e_i^*(\phi, j\omega) \nabla F_i(\phi, j\omega)\} \, d\omega. \tag{6.140}$$

Comparing (6.137), (6.139), and (6.140), we see that if the adjoint network excitations are taken as

$$p\,|e_i(\phi,j\omega)|^{p-2}w_i(\omega)e_i^*(\phi,j\omega) = \begin{cases} \hat{V}_i(j\omega) & i = 1, 2, \ldots, n_V \\ -\hat{I}_i(j\omega) & i = n_V + 1, \ldots, n_V + n_I, \end{cases} \tag{6.141}$$

then

$$\nabla U = \int_{\Omega} \mathrm{Re}\{\mathbf{G}\}\, d\omega. \tag{6.142}$$

The corresponding expression for the discrete case is

$$\nabla U = \sum_{\Omega_d} \mathrm{Re}\{\mathbf{G}\} \tag{6.143}$$

where Ω_d is the discrete set of frequencies.

An Application to Group Delay Computation

In group delay computations we are essentially interested in sensitivities with respect to frequency ω [43]. This parameter is different from others that we have considered in that it is common throughout the network. Specifically, let us distinguish variables associated with some jth element of an n-element network by the subscript j. Then, assuming only ω is varied, (6.128) can be written as

$$\Delta\mathbf{I}_V^T\hat{\mathbf{V}}_V - \Delta\mathbf{V}_I^T\hat{\mathbf{I}}_I = \sum_{j=1}^{n} [\mathbf{V}_{aj}^T\mathbf{I}_{bj}^T] \begin{bmatrix} -\Delta\mathbf{Y}_j^T & \Delta\mathbf{M}_j^T \\ -\Delta\mathbf{A}_j^T & \Delta\mathbf{Z}_j^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_{aj} \\ \hat{\mathbf{I}}_{bj} \end{bmatrix}, \tag{6.144}$$

if each element, for complete generality, is characterized by an appropriate hybrid matrix. Using the rule that $\Delta x = (\partial x/\partial\omega)\,\Delta\omega$, where x is any quantity depending on ω, (6.144) can be more appropriately written

$$\sum_{i=1}^{n_V} \hat{V}_i \frac{\partial I_i}{\partial\omega} - \sum_{i=n_V+1}^{n_V+n_I} \hat{I}_i \frac{\partial V_i}{\partial\omega} = \sum_{j=1}^{n} G_{\omega j} \tag{6.145}$$

where

$$G_{\omega j} \triangleq [\mathbf{V}_{aj}^T\mathbf{I}_{bj}^T] \begin{bmatrix} -\dfrac{\partial\mathbf{Y}_j^T}{\partial\omega} & \dfrac{\partial\mathbf{M}_j^T}{\partial\omega} \\ -\dfrac{\partial\mathbf{A}_j^T}{\partial\omega} & \dfrac{\partial\mathbf{Z}_j^T}{\partial\omega} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_{aj} \\ \hat{\mathbf{I}}_{bj} \end{bmatrix}. \tag{6.146}$$

If, in particular, the kth port is to be investigated, and this happens to be a current-excited port,* then (6.145) reduces to

$$-\hat{\mathbf{I}}_k \frac{\partial V_k}{\partial \omega} = \sum_{j=1}^{n} G_{\omega j} \tag{6.147}$$

if all adjoint excitations except $\hat{I}_k$ are set to zero. Evaluation of the sensitivity expression $G_{\omega j}$ is accomplished by the results of two network analyses. The sensitivity formulas from Table 6-2 may be used if appropriate, since

$$\begin{bmatrix} -\dfrac{\partial \mathbf{Y}_j^T}{\partial \omega} & \dfrac{\partial \mathbf{M}_j^T}{\partial \omega} \\ -\dfrac{\partial \mathbf{A}_j^T}{\partial \omega} & \dfrac{\partial \mathbf{Z}_j^T}{\partial \omega} \end{bmatrix} = \sum_r \begin{bmatrix} -\dfrac{\partial \mathbf{Y}_j^T}{\partial \phi_r} & \dfrac{\partial \mathbf{M}_j^T}{\partial \phi_r} \\ -\dfrac{\partial \mathbf{A}_j^T}{\partial \phi_r} & \dfrac{\partial \mathbf{Z}_j^T}{\partial \phi_r} \end{bmatrix} \frac{\partial \phi_r}{\partial \omega}, \tag{6.148}$$

where the subscript r denotes some rth parameter in the jth element with respect to which a sensitivity expression is already available.

Consider, for example, $\theta = j\omega l/c = j\beta l$ where c is the velocity of propagation. Then the ω-sensitivity of a lossless transmission line is jl/c times the θ-sensitivity shown in Table 6-2. Consider an inductor as a second example. The lefthand side of (6.148) reduces immediately to jL using $Z = j\omega L$.

Finally, to compute the group delay $T_G(\omega)$ we note that

$$T_G(\omega) = -\mathrm{Im}\left\{\frac{1}{V_k}\frac{\partial V_k}{\partial \omega}\right\}, \tag{6.149}$$

where it is assumed that all sources have constant, frequency-independent phase angles. For convenience, letting the excitation $\hat{I}_k = 1/V_k$,

$$T_G(\omega) = \mathrm{Im}\left\{\sum_{j=1}^{n} G_{\omega j}\right\}. \tag{6.150}$$

Equation (6.150) is also valid for calculations of group delay if the kth port is a voltage-excited port.* All one has to remember is to set all adjoint excitations to zero except $\hat{V}_k$ which is set to $-1/I_k$.

Extensions and Other Applications

An important point to remember about the adjoint network method is that the analysis of the adjoint, in general, can take considerably less effort than the analysis of the original network. If $\mathbf{Y}_n$ is, for example, the nodal admittance matrix of the original network, and its inverse $\mathbf{Y}_n^{-1}$ has been computed, then we can use the result $(\mathbf{Y}_n^T)^{-1} = (\mathbf{Y}_n^{-1})^T$. For a further discussion of possible computational efficiency, the reader is referred to Director [44].

*The value of the excitation could, of course, be zero.

Extensions to second-order sensitivities have been formulated [45], including group-delay sensitivities [43]. Of particular interest to filter designers are the recent applications of the adjoint network concept to the computation of dissipation-induced loss distortion in both lumped and distributed networks [46, 47]. Further extensions include the exploitation of the adjoint network concept in first- and second-order sensitivity computation using wave variables rather than voltages and currents [48, 49, 50]. These results should also be of interest to filter designers.

6.10 SUMMARY

A wide range of topics in the field of computer-aided circuit optimization has been discussed. Formulations and methods suitable for automated design, when the classical approach is inappropriate, have been stressed. The formulation of objective functions from design objectives has been discussed, including least pth and minimax. Methods of dealing with parameter and response constraints by means of transformations or penalties have been considered in some detail. Minimax approximation through linear programming and nonlinear programming has been discussed. Efficient one-dimensional methods and multidimensional gradient and direct search methods have been reviewed. Least pth approximation has been considered, with emphasis on gradient methods of solution. Finally, the adjoint network method of evaluating derivatives for design in the frequency domain was reviewed.

Most computer centers should have linear programming routines, and at least one efficient gradient algorithm, available as library programs, and possibly other methods also. It is hoped that this chapter has gone a reasonable way towards helping the network designer formulate his problems effectively so that he can take full advantage of the available computer programs.

ACKNOWLEDGEMENT

The cooperation of R. E. Seviora of the University of Toronto, particularly in the section on adjoint networks, is much appreciated. Dr. E. Della Torre of McMaster University provided considerable constructive criticism.

PROBLEMS

6.1 (a) Prove that ΔU is maximized in the direction of ∇U for a given step size.
(b) Use the multidimensional Taylor series expansion to show that a turning point of a convex differentiable function is a global minimum.

6.2 (a) If $g(\phi)$ is concave, verify that $g(\phi) \geq \mathbf{0}$ describes a convex feasible region.
(b) Under what conditions could equality constraints be included in convex programming?

6.3 Find suitable transformations for the following constraints so that we can use unconstrained optimization.
(a) $0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq \phi_i \leq \cdots \leq \phi_k$.
(b) $0 < l \leq \phi_2/\phi_1 \leq u$
$\phi_1 > 0$
$\phi_2 > 0$.

6.4 Derive (6.99) and (6.100).

6.5 Derive the sensitivity expression (6.124) from first principles.

6.6 Derive the entries of Table 6-2 relating to:
(a) A voltage controlled voltage source.
(b) An open-circuited uniformly distributed line.
(c) A uniform RC line.

6.7 Verify that the adjoint network may be characterized by the hybrid matrix description in (6.126).

6.8 Obtain the adjoint network in terms of an *ABCD* or chain matrix characterization of a two-port. Find sensitivity expressions in these terms for some of the entries of Table 6-2.

6.9 Consider the problem of minimizing

$$U = \phi_3(\phi_1 + \phi_2)^2$$

subject to

$$g_1 = \phi_1 - \phi_2^2 \geq 0$$
$$g_2 = \phi_2 \geq 0$$
$$h = (\phi_1 + \phi_2)\phi_3 - 1 = 0.$$

Is this a convex programming problem? Formulate it for solution by the sequential unconstrained minimization method. Starting with a feasible point, show how the constrained minimum is approached as the parameter $r \to 0$. Draw a contour sketch to illustrate the process. Are the conditions for a constrained minimum satisfied?

6.10 For the linear function

$$F(\boldsymbol{\phi}, \boldsymbol{\psi}) = \sum_{i=1}^{k} \phi_i f_i(\boldsymbol{\psi}),$$

(a) Formulate the discrete minimax approximation of $S(\boldsymbol{\psi})$ by $F(\boldsymbol{\phi}, \boldsymbol{\psi})$ as a linear programming problem, assuming $\boldsymbol{\phi}$ to be unconstrained.
(b) Assuming an objective function of the form of (6.102), derive ∇U and $\mathbf{H}$ (Note that a polynomial is a special case).

6.11 Verify (6.137).

6.12 Formulate the design of a notch filter in terms of inequality constraints, given the following requirements. The attenuation should not exceed A_1 dB over the frequency range 0 to ω_1, and A_2 dB over the range ω_2 to ω_3, with $0 < \omega_1 < \omega_2 < \omega_3$. At ω_0, where $\omega_1 < \omega_0 < \omega_2$, the attenuation must exceed A_0 dB.

6.13 Devise an algorithm for finding the extrema of a well-behaved multimodal function of one variable (Figure 6-2), such as the passband response of a filter.

6.14 Discuss the scaling effects of the transformation $\phi_i = \exp \phi_i'$ (Table 6-1).

6.15 (a) Are the necessary conditions for a constrained *minimum* satisfied anywhere along the boundary of the feasible region in Figure 6-1?
(b) What about the conditions for a constrained *maximum*?

6.16 Suppose we have to minimize

$$U = \sum_{\omega_i \in \Omega_d} [L(\omega_i) - S(\omega_i)]^p$$

where $L(\omega_i)$ is the insertion loss in dB of a filter between R_g and R_L, $S(\omega_i)$ is the desired insertion loss between R_g and R_L, Ω_d is a set of discrete frequencies ω_i, and p is an even positive integer. Obtain an expression relating ∇U to $\mathbf{G}(j\omega_i)$ where the elements of $\mathbf{G}$ might be as in Table 6-2. Assume convenient values for the excitations of the original and adjoint networks.

REFERENCES

[1] R. Fletcher, "A review of methods for unconstrained optimization," *Optimization*, R. Fletcher, Ed., Academic Press, New York, 1969.

[2] E. M. L. Beale, "Nonlinear programming," *Digital Computer User's Handbook*, M. Klerer and G. A. Korn, Ed., McGraw-Hill, New York, 1967.

[3] P. Wolfe, "Methods of nonlinear programming," *Nonlinear Programming*, J. Abadie, Ed., John Wiley, New York, 1967.

[4] W. I. Zangwill, *Nonlinear Programming*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

[5] M. J. Box, "A comparison of several current optimization methods, and the use of transformations in constrained problems," *Computer J.*, **9**, 67–77 (May, 1966).

[6] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory and Techniques*, **MTT-17**, 533–552 (Aug., 1969).

[7] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley, New York, 1968.

[8] J. Kowalik and M. R. Osborne, *Methods for Unconstrained Optimization Problems*, Elsevier, New York, 1968.

[9] F. A. Lootsma, "Logarithmic programming: a method of solving nonlinear-programming problems," *Philips Res. Repts.*, **22**, 329–344 (June, 1967).

[10] J. Bracken and G. P. McCormick, *Selected Applications of Nonlinear Programming*, John Wiley, New York, 1968.

[11] H. W. Kuhn and A. W. Tucker, "Non-linear programming," *Proc. 2nd Symp. on Math. Statistics and Probability. Berkeley, Calif.*, University of California Press, 481–493, 1951.

[12] D. C. Handscomb, Ed., *Methods of Numerical Approximation*, Pergamon, Oxford, 1966.

[13] G. C. Temes and J. A. C. Bingham, "Iterative Chebyshev approximation technique for network synthesis," *IEEE Trans. Circuit Theory*, **CT-14**, 31–37 (March, 1967).

[14] G. C. Temes and D. A. Calahan, "Computer-aided network optimization the state-of-the-art," *Proc. IEEE*, **55**, 1832–1863 (Nov., 1967).

[15] G. C. Temes, "Optimization methods in circuit design," *Computer Oriented Circuit Design*, F. F. Kuo and W. G. Magnuson, Jr., Ed., Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

[16] L. S. Lasdon and A. D. Waren, "Optimal design of filters with bounded, lossy elements," *IEEE Trans. Circuit Theory*, **CT-13**, 175–187 (June, 1966).

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

[18] Y. Ishizaki and H. Watanabe, "An iterative Chebyshev approximation method for network design," *IEEE Trans. Circuit Theory*, **CT-15**, 326–336 (Dec., 1968).

[19] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the nonlinear case," *Computer J.*, **12**, 63–68 (Feb., 1969).

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

[21] J. W. Bandler, "Computer optimization of inhomogeneous waveguide transformers," *IEEE Trans. Microwave Theory and Techniques*, **MTT-17**, 563–571 (Aug., 1969).

[22] J. W. Bandler and P. A. Macdonald, "Optimization of microwave networks by razor search," *IEEE Trans. Microwave Theory and Techniques*, **MTT-17**, 552–562 (Aug., 1969).

[23] H. J. Carlin and O. P. Gupta, "Computer design of filters with lumped-distributed elements or frequency variable terminations," *IEEE Trans. Microwave Theory and Techniques*, **MTT-17**, 598-604 (Aug., 1969).

[24] J. W. Bandler and A. G. Lee-Chan, "Gradient razor search method for optimization," *1971 International Microwave Symp., Digest of Technical Papers*, 118–119 (May, 1971).

[25] J. W. Bandler, "Conditions for a minimax optimum," *IEEE Trans. Circuit Theory*, **CT-18**, 476-479 (July, 1971).

[26] M. J. Box, D. Davies, and W. H. Swann, *Non-linear Optimization Techniques*, Oliver and Boyd, Edinburgh, 1969.

[27] R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," *Computer J.*, **6**, 163–168 (June, 1963).

[28] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Computer J.*, **7**, 155–162 (July, 1964).

[29] J. W. Bandler and P. A. Macdonald, "Cascaded noncommensurate transmission-line networks as optimization problems," *IEEE Trans. Circuit Theory*, **CT-16**, 391–394 (Aug., 1969).

[30] R. Hooke and T. A. Jeeves, "'Direct search' solution of numerical and statistical problems," *J. ACM*, **8**, 212–229 (April, 1961).

[31] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer J.*, **7**, 308–313 (Jan., 1965).

[32] H. H. Rosenbrock, "An automatic method for finding the greatest or least value of a function," *Computer J.*, **3**, 175–184 (Oct., 1960).

[33] W. I. Zangwill, "Minimizing a function without calculating derivatives," *Computer J.*, **10**, 293–296 (Nov., 1967).

[34] R. Fletcher, "Function minimization without evaluating derivatives—a review," *Computer J.*, **8**, 33–41 (April, 1965).

[35] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Computer J.*, **7**, 149–154 (July, 1964).

[36] G. C. Temes and D. Y. F. Zai, "Least *p*th approximation," *IEEE Trans. Circuit Theory*, **CT-16**, 235–237 (May, 1969).

[37] J. W. Bandler and R. E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory and Techniques*, **MTT-18**, 1159–1170 (Dec., 1970).

[38] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, **CT-16**, 318–323 (Aug., 1969).

[39] S. W. Director and R. A. Rohrer, "Automated network design—the frequency-domain case," *IEEE Trans. Circuit Theory*, **CT-16**, 330–337 (Aug., 1969).

[40] J. W. Bandler and R. E. Seviora, "Computation of sensitivities for noncommensurate networks," *IEEE Trans. Circuit Theory*, **CT-18**, 174–178 (Jan., 1971).

[41] C. A. Desoer and E. S. Kuh, *Basic Circuit Theory*, McGraw-Hill, New York, 1969, Chapter 9.

[42] R. E. Seviora, M. Sablatash and J. W. Bandler, "Least *p*th and minimax objectives for automated network design," *Electronics Letters*, **6**, 14–15 (Jan., 1970).

[43] G. C. Temes, "Exact computation of group delay and its sensitivities using adjoint-network concept," *Electronics Letters*, **6**, 483–485 (July, 1970).

[44] S. W. Director, "*LU* factorization in network sensitivity calculations," *IEEE Trans. Circuit Theory*, **CT-18**, 184–185 (Jan., 1971).

[45] G. A. Richards, "Second-derivative sensitivity using the concept of the adjoint network," *Electronics Letters*, **5**, 398–399 (Aug., 1969).

[46] G. C. Temes and R. N. Gadenz, "Simple technique for the prediction of dissipation-induced loss distortion," *Electronics Letters*, **6**, 836–837 (Dec., 1970).

[47] R. N. Gadenz and G. C. Temes, "Computation of dissipation-induced loss distortion in lumped/distributed networks," *Electronics Letters*, **7**, 258–260 (May, 1971).

[48] J. W. Bandler and R. E. Seviora, "Sensitivities in terms of wave variables," *Proc. 8th Annual Allerton Conf. on Circuit and System Theory*, 379–387 (Oct., 1970).

[49] J. W. Bandler and R. E. Seviora, "Wave sensitivities of networks," *IEEE Trans. Microwave Theory and Techniques*, **MTT-20**, 138–147 (Feb., 1972).

[50] J. W. Bandler and R. E. Seviora, "Computation of equivalent wave source using the adjoint network," *Electronics Letters*, **7**, 235–236 (May, 1971).