SIMULATION OPTIMIZATION SYSTEMS
Research Laboratory

ADVANCES IN OPTIMIZATION OF CIRCUITS AND SYSTEMS USING RECENT MINIMAX AND ℓ_1 ALGORITHMS

W. Kellermann

SOS-86-2-T

March 1986

McMASTER UNIVERSITY
Hamilton, Canada L8S 4L7
Department of Electrical and Computer Engineering

ADVANCES IN OPTIMIZATION OF CIRCUITS AND SYSTEMS USING RECENT MINIMAX AND ℓ_1 ALGORITHMS

W. Kellermann

SOS-86-2-T

March 1986

ADVANCES IN OPTIMIZATION OF CIRCUITS AND SYSTEMS USING RECENT MINIMAX AND ℓ_1 ALGORITHMS

ADVANCES IN OPTIMIZATION OF CIRCUITS AND SYSTEMS USING RECENT MINIMAX AND ℓ_1 ALGORITHMS

By

WITOLD KELLERMANN, M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

December 1985

DOCTOR OF PHILOSOPHY (1985)
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Advances in Optimization of Circuits and Systems Using Recent Minimax and ℓ_1 Algorithms

AUTHOR: Witold Kellermann, M.Sc. (E.E.)
(Technical University of Wroclaw)

SUPERVISOR: J.W. Bandler, Professor, Department of Electrical and Computer Engineering
B.Sc. (Eng.), Ph.D., D.Sc. (Eng.) (University of London)
D.I.C., (Imperial College)
P.Eng. (Province of Ontario)
C.Eng., F.I.E.E. (United Kingdom)
Fellow, I.E.E.E.

NUMBER OF PAGES: xvi, 202

ABSTRACT

This thesis concerns itself with computer-aided techniques for design centering, tolerancing and tuning, fault location and model parameter identification from measurements.

Since many of the engineering system problems discussed in this thesis are formulated as optimization problems we examine algorithms and techniques for nonlinear optimization. Our attention is focused on minimax and ℓ_1 algorithms since many formulations of engineering system problems exploit the characteristic features of these two norms.

A novel approach for worst-case network design is proposed and an algorithm for the fixed tolerance problem embodying worst-case search and selection of sample points is presented.

The features of the ℓ_1 norm in the tuning problem are discussed in detail and explained using necessary conditions for optimality of the nonlinear ℓ_1 problem with nonlinear constraints. Regular and singular ℓ_1 problems are defined and a criterion for determining a singularity present in the ℓ_1 problem is formulated.

New formulations using the ℓ_1 norm are given for fault isolation and model parameter identification in analog circuits.

Practical engineering problems have been solved illustrating the wide applicability of the concepts used and the robustness of the algorithms employed.

A new algorithm for minimizing the cardinality of a set subject to nonlinear, nondifferentiable constraints is presented and illustrated by solving the best mechanical alignment problem. The load shedding and generation rescheduling problem in power systems is formulated using the ℓ_1 norm. The formulation is tested

on 6-bus and 26-bus power systems. A general microwave multiplexer design procedure exploiting exact network sensitivities is introduced and illustrated by designing 5-channel, 11 GHz and 12-channel, 12 GHz multiplexers.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Dr. J.W. Bandler for his encouragement, continued assistance, expert guidance and supervision throughout the course of this work. He also thanks Dr. C.M. Crowe and Dr. M.A. El-Kady, members of his Supervisory Committee for their continuing interest and useful suggestions. The interaction with Dr. El-Kady resulted in new applications, particularly in the power systems area (Chapter 6) and the best mechanical alignment problem (Chapter 5).

The author is particularly grateful to Dr. K. Madsen of the Technical University of Denmark for making available his Fortran codes for minimax and ℓ_1 optimization and for fruitful technical discussions.

The development of new results in the design of microwave multiplexers (Chapter 7) is attributable in part to the very strong initial impetus provided in this direction by Com Dev Ltd. of Cambridge, Ontario.

It is the author's pleasure to acknowledge discussions with Dr. W. M. Zuberek, currently with the Memorial University of Newfoundland. Dr. Zuberek contributed ideas to the best mechanical alignment problem and to the new approach to worst-case circuit design.

The author acknowledges useful and inspiring discussions with his colleagues S. Daijavad, S.H. Chen, Q.J. Zhang and Dr. A.E. Salama.

The assistance of Ms. M.L. Renault of McMaster University in preparing some of the numerical examples is also gratefully acknowledged.

Thanks are due to Mr. S. Pan of McMaster University for adapting the ℓ_1 package to the Texas Instruments Professional Computer and for running some of the test problems.

The financial assistance provided by the Natural Sciences and Engineering Research Council of Canada through Grants A7239 and G1135, the Department of Electrical and Computer Engineering through a Teaching Assistantship and the Ministry of Colleges and Universities through an Ontario Graduate Scholarship is gratefully acknowledged.

Thanks go to the Engineering Word Processing Centre for their patience and expert typing of this manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF TABLES

LIST OF TABLES (continued)

1

INTRODUCTION

The increasing size and complexity of physical man-made engineering systems necessitate the use of computers in all aspects of the design, production and maintenance processes. A corresponding need has developed for efficient and powerful computer-aided techniques for thorough study and optimal realization of the above mentioned processes.

Computer-aided design (CAD) techniques are now well established for design centering, tolerance optimization, yield maximization, cost minimization and the rapidly increasing range of applications includes electronic circuits, power systems, microwave systems and mechanical systems.

Computer-aided design is often treated together with computer-aided manufacturing (CAM). We are not including CAM in this thesis, since CAM starts from data, preferably machine-readable data, that is produced in the design process, but CAM is not part of the design process itself.

Computer-aided testing (CAT) techniques, which originated from the area of digital circuits, are primarily associated for analog circuits with the problems of fault location, model parameter identification from measurements and postproduction tuning.

Recently, the term computer-aided engineering (CAE) has been used most frequently for turnkey software and hardware systems for electronic systems and component design. It has also been used in a more general sense to include a broad set of system analysis tools applicable in many engineering disciplines.

Most of the discussion in this thesis focuses on computer-aided engineering system problem solving, in which key elements are formulations of the problems, algorithms for solving the problems and software implementing the methods proposed.

Many of the subproblems associated with overall problems in different engineering disciplines are similar to each other. This motivates us to utilize a conceptual framework developed over the past ten years for design centering, tolerancing and tuning (DCTT) and for fault analysis, postproduction tuning and model parameter identification from measurements. Our aim is to provide a set of methods and techniques for solving these problems which employ recent optimization algorithms with the emphasis on nonlinear minimax and ℓ_1 algorithms. Our attention is focused on minimax and ℓ_1 algorithms since many formulations of engineering system problems exploit the characteristic features of these two norms.

We do not presume to be able to solve all problems associated with any overall engineering system. Applications of the methods and techniques proposed will be immediately apparent in many cases. Often it will also occur that familiarity with the concepts and techniques will clarify certain problem aspects which have been obscured or unrecognized.

In Chapter 2 previous work in the area of design centering, tolerancing and tuning is reviewed. We consider the relevant fundamental concepts and definitions commonly used in the DCTT literature. Three general formulations of the optimal DCTT problem are given and some important special cases are described in more detail. We provide also an adequate state-of-the-art review of algorithms for DCTT problems.

Chapter 3 deals with the use of minimax optimization techniques in computer-aided engineering. A critical review of the existing minimax algorithms is given together with a comparison of minimax algorithms using the classical three-section transmission-line transformer (Bandler and Macdonald 1969a). The Hald and Madsen algorithm (Hald and Madsen 1981) is treated in some detail and its performance is demonstrated on regular and singular problems. A detailed description of the algorithm is given in Appendix A. A novel approach to worst-case network design is proposed and an algorithm for the fixed tolerance problem embodying worst-case search and selection of sample points is presented.

Chapter 4 covers the use of ℓ_1 optimization techniques in computer-aided engineering. Previous work in the area of nonlinear ℓ_1 optimization is briefly reviewed. The Hald and Madsen algorithm (Hald and Madsen 1985) for nonlinear ℓ_1 optimization is presented in some detail in Appendix C. The features of the ℓ_1 norm in the tuning problem are discussed in detail and explained using necessary conditions for optimality of the nonlinear ℓ_1 problem with nonlinear constraints. Regular and singular ℓ_1 problems are defined and a criterion for determining a singularity present in the ℓ_1 problem is formulated. New formulations using the ℓ_1 norm are given for fault isolation and model parameter identification in analog circuits.

The next three chapters, Chapter 5, 6 and 7, contain major applications of the concepts and methods described earlier. In each case, a difficult engineering problem has been solved illustrating the wide applicability of the methods proposed and the robustness of the algorithms used.

Chapter 5 describes a minimax approach to the best mechanical alignment problem. A new algorithm for minimizing the cardinality of a set subject to nonlinear,

nondifferentiable constraints is presented and illustrated by solving practical problems.

Load shedding and generation rescheduling in power systems using the ℓ_1 norm is treated in Chapter 6. A new formulation for the problem is proposed and tested using 6-bus and 26-bus systems.

In Chapter 7, optimal design of microwave multiplexing networks is formulated as an optimization problem. A general multiplexer optimization procedure exploiting exact network sensitivities is illustrated by designing 11 GHz, 5-channel and 12 GHz, 12-channel multiplexers.

We conclude in Chapter 8 along with some suggestions for further research.

The author contributed substantially to the following original developments presented in this thesis:

(1) An algorithm for the fixed tolerance problem embodying worst-case search and selection of sample points.

(2) Mixed programming formulation of the tuning problem ensuring that the solution gives the minimum number of tunable parameters.

(3) A new formulation for fault isolation in analog circuits using the ℓ_1 norm and an exact penalty function.

(4) An algorithm for minimizing the cardinality of a set subject to nonlinear, nondifferentiable constraints.

(5) A formulation of the load shedding problem in power systems using the ℓ_1 norm.

(6) A procedure for optimal design of microwave multiplexers using ℓ_1 and minimax optimization.

2

DESIGN CENTERING, TOLERANCING AND TUNING (DCTT) - A REVIEW

2.1 INTRODUCTION

The development of new design procedures and techniques can, in general, be characterized as an attempt to include in the design process as many factors which may influence the performance of a manufactured design as possible. With readily available and ever increasing computing power at hand, computer-aided designers are dealing with more realistic problems. We should not, however, rely only on the computing power of modern machines since besides economical considerations there exist physical limits to what is practically achievable. Bremermann (1962) determined by simple physical considerations that "... no data processing system whether artificial or living can process more than $2x10^{47}$ bits per second per gram of its mass".

In the classical design problem we are interested in finding one single point in the design parameter space which satisfies the design specifications. This kind of solution is impractical from the manufacturing point of view since there is a number of factors which influence the performance of a manufactured design.

Phenomena associated with the design of circuits and which can be considered are (Bandler and Rizk 1979):

(a) manufacturing tolerances (i.e., the actual value of the design variable outcome may lie within an interval with a certain probability density function);

(b) model uncertainties; equivalent circuits are used to model actual circuits and the parameters of equivalent circuits usually have uncertainties associated with them;

(c) parasitic effects; these parasitics can substantially alter the ideal circuit performance and should be taken into consideration where possible; they are marked in many analog electrical circuits (active, high frequency, etc.);

(d) environmental uncertainties; some circuits have to meet stringent specifications for a variety of different environmental conditions; military and telephone equipment, for example, often has to be designed for extreme temperatures;

(e) mismatched terminations; network terminations or loads may be substantially different from ideal;

(f) material uncertainty; uncertainties exist in the materials used to fabricate the circuits.

Taking into account in the design process the above mentioned factors, if at all possible, is usually in conflict with the feasible or acceptable computational effort involved. Therefore, a successful design procedure is usually a compromise between the complexity of the model and the computational cost to produce a design satisfying all specifications.

In the next section we consider the relevant fundamental concepts and definitions commonly used in the DCTT literature.

2.2 FUNDAMENTAL CONCEPTS AND DEFINITIONS

The mathematical formulation of an approach which embodies centering, tolerancing and tuning in a unified manner was presented by Bandler, Liu and Tromp (1976a).

A design consists of design data of the nominal point $\boldsymbol{\phi}^0$, the tolerance vector $\boldsymbol{\varepsilon}$ and the tuning vector $\mathbf{t}$, where

$$\boldsymbol{\phi}^0 \triangleq \begin{bmatrix} \phi_1^0 \\ \phi_2^0 \\ . \\ . \\ . \\ \phi_n^0 \end{bmatrix}, \quad \boldsymbol{\varepsilon} \triangleq \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix} \quad \text{and} \quad \mathbf{t} \triangleq \begin{bmatrix} t_1 \\ t_2 \\ . \\ . \\ . \\ t_n \end{bmatrix}, \tag{2.1}$$

n is the number, for example, of network parameters which may be indexed by

$$I_\phi \triangleq \{1, 2, ..., n\}. \tag{2.2}$$

We will assume that the parameters can be varied continuously and that the parameters can be chosen independently. Extra conditions such as discretization and imposed parameter bounds may be treated as constraints, see Bandler, Liu and Chen (1975). Some of the parameters can be set to zero or held constant.

An outcome $\{\boldsymbol{\phi}, \boldsymbol{\varepsilon}, \boldsymbol{\mu}\}$ of a design $\{\boldsymbol{\phi}, \boldsymbol{\varepsilon}, \mathbf{t}\}$ implies a point in the parameter space given by

$$\boldsymbol{\phi} = \boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}, \tag{2.3}$$

where

$$\mathbf{E} \triangleq \begin{bmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \\ & & \ddots & \\ & & & \varepsilon_n \end{bmatrix}, \tag{2.4}$$

and $\boldsymbol{\mu} \in R_\mu$. R_μ is a set of multipliers determined from realistic situations of the tolerance spread. We consider

$$R_\mu \triangleq \{\boldsymbol{\mu} \mid -1 \leq \mu_i \leq -a_i,\ a_i \leq \mu_i \leq 1,\ i \in I_\phi\}, \tag{2.5}$$

where

$$0 \leq a_i \leq 1. \tag{2.6}$$

The most commonly used continuous range is obtained by setting a_i to zero. Unless otherwise stated, the case

$$R_\mu \triangleq \{\boldsymbol{\mu} \mid -1 \leq \mu_i \leq 1,\quad i \in I_\phi\}, \tag{2.7}$$

is considered (Bandler and Liu 1974).

The tolerance region R_ε, as described by Butler (1971) and Bandler (1972, 1974), is a set of points defined by (2.3) for all $\boldsymbol{\mu} \in R_\mu$. In the case of $-1 \leq \mu_i \leq 1$, $i \in I_\phi$,

$$R_\varepsilon \triangleq \{\boldsymbol{\phi} \mid \phi_i = \phi_i^0 + \varepsilon_i \mu_i,\ -1 \leq \mu_i \leq 1,\quad i \in I_\phi\}, \tag{2.8}$$

which is a convex regular polytope of n dimensions with sides of length $2\varepsilon_i$, $i \in I_\phi$, and centered at $\boldsymbol{\phi}^0$. The extreme points of R_ε are obtained by setting $\mu_i = \pm 1$. Thus, the set of vertices may be defined as

$$R_v \triangleq \{\boldsymbol{\phi} \mid \phi_i = \phi_i^0 + \varepsilon_i \mu_i,\ \mu_i \in \{-1, 1\},\quad i \in I_\phi\}. \tag{2.9}$$

The number of points in R_v is 2^n. Let each of these points be indexed by $\boldsymbol{\phi}^i$, $i \in I_v$, where

$$I_v \triangleq \{1, 2, \ldots, 2^n\}. \tag{2.10}$$

Thus, $R_v = \{\boldsymbol{\phi}^1, \boldsymbol{\phi}^2, \ldots, \boldsymbol{\phi}^{2^n}\}$.

The tuning region $R_t(\boldsymbol{\mu})$ is defined as the set of points (see Bandler, Liu and Tromp 1976a)

$$\boldsymbol{\phi} = \boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu} + \mathbf{T}\boldsymbol{\rho} \tag{2.11}$$

for all $\boldsymbol{\rho} \in R_\rho$, where

$$\mathbf{T} \triangleq \begin{bmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{bmatrix}. \tag{2.12}$$

Some of the common examples of R_ρ are

$$R_\rho \triangleq \{\boldsymbol{\rho} \mid -1 \le \rho_i \le 1,\ \ i \in I_\phi\}, \tag{2.13}$$

or in the case of one-way tuning or irreversible trimming,

$$R_\rho \triangleq \{\boldsymbol{\rho} \mid 0 \le \rho_i \le \quad 1,\ i \in I_\phi\}, \tag{2.14}$$

or

$$R_\rho \triangleq \{\boldsymbol{\rho} \mid -1 \le \rho_i \le 0,\ i \in I_\phi\}. \tag{2.15}$$

The constraint region R_c is defined as (Butler 1971, Bandler 1972, 1974),

$$R_c \triangleq \{\boldsymbol{\phi} \mid g_i(\boldsymbol{\phi}) \ge 0,\ i \in I_c\}, \tag{2.16}$$

where

$$I_c \triangleq \{1,2,\ldots,m_c\}. \tag{2.17}$$

is the index set for the performance specifications and parameter constraints. R_c is assumed to be nonempty.

The definitions and concepts presented are illustrated in Fig. 2.1 by a two-dimensional example.

2.3 THREE FORMULATIONS OF THE OPTIMAL DCTT PROBLEM

2.3.1 General Nonlinear Programming Formulation with Differentiable Constraints

The first general formulation of the optimal DCTT problem was given by Bandler, Liu and Tromp (1976a). The problem was stated as follows: obtain a set of

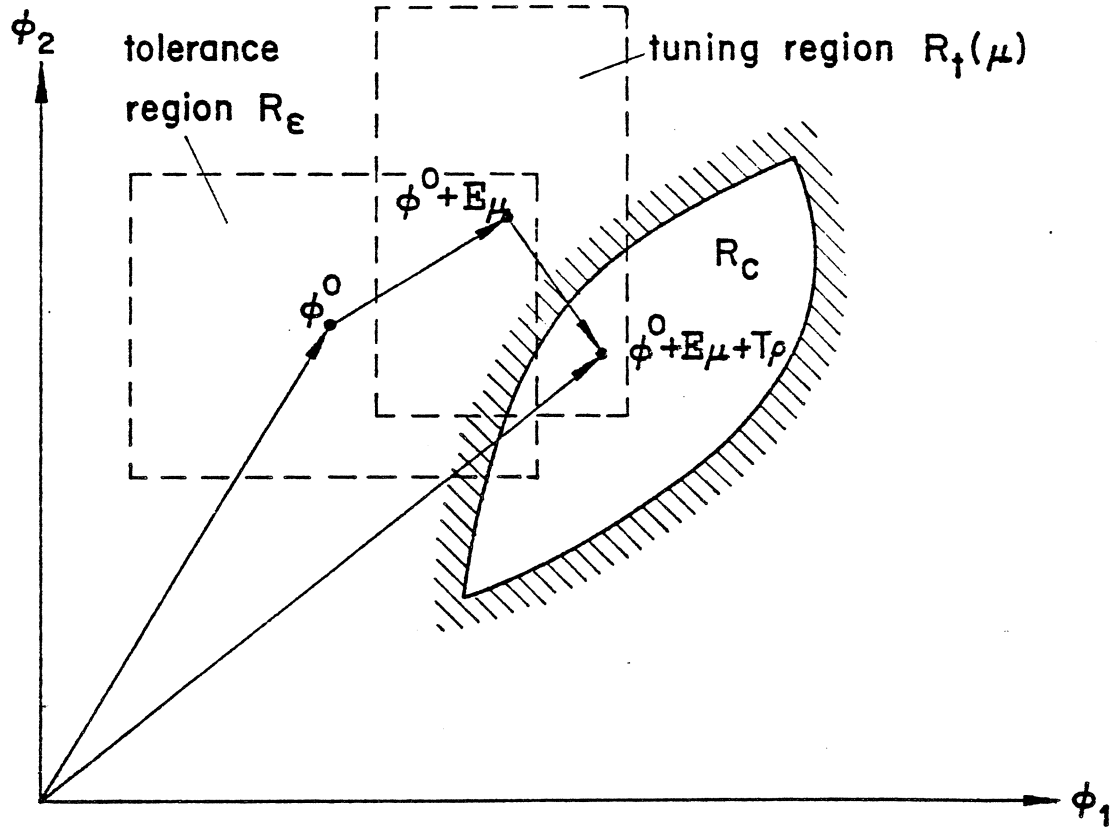


Fig. 2.1 Illustration of concepts in design centering, tolerancing and tuning. (Bandler, Liu and Tromp 1976a).

optimal design values $\{\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \mathbf{t}\}$ such that any outcome $\{\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \boldsymbol{\mu}\}$, $\boldsymbol{\mu} \in R_\mu$, may be tuned into R_c for some $\boldsymbol{\rho} \in R_\rho$. It was formulated as the nonlinear programming problem:

$$\text{minimize } C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \mathbf{t}) \tag{2.18}$$

subject to

$$\boldsymbol{\phi} \in R_c \ (\text{or } \mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}), \tag{2.19}$$

where

$$\boldsymbol{\phi} = \boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu} + \mathbf{T}\boldsymbol{\rho} \tag{2.20}$$

and constraints $\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \mathbf{t} \geq \mathbf{0}$, for all $\boldsymbol{\mu} \in R_\mu$ and some $\boldsymbol{\rho} \in R_\rho$. C is an appropriate function chosen to represent a reasonable approximation to known component cost data.

Stated in an abstract sense, the worst-case solution of the problem must satisfy

$$R_t(\boldsymbol{\mu}) \cap R_c \neq \varnothing \tag{2.21}$$

for all $\boldsymbol{\mu} \in R_\mu$, where $\varnothing$ denotes an empty set.

They also discussed the geometrical structure of the problem and introduced some important special cases obtained by separating the components into effectively tuned and effectively toleranced parameters. They proved that the solution of the reduced problem is the solution of the original one under certain conditions.

2.3.2 General Nonlinear Programming Formulation with Nondifferentiable Constraints

Polak and Sangiovanni-Vincentelli (1979) formulated the DCTT problem as a mathematical programming problem in the form

$$\text{minimize } C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \mathbf{t})$$

subject to

$$\min_{i \in I_c} \quad \min_{\mu \in R_\mu} \quad \max_{\rho \in R_\rho} \quad g_i(\phi) \geq 0 \tag{2.22}$$

and the constraints ϕ^0, ε, $t \geq 0$, where ϕ is given in (2.20). They demonstrated that their formulation is equivalent to the one in Bandler, Liu and Tromp (1976a). They suggested a new algorithm which deals with the nondifferentiable constraints (2.22). The algorithm solves the problem as a sequence of approximating problems with $R_\mu^j \subset R_\mu$ as a discrete set. They showed that, under certain conditions, the accumulation points of the sequence of stationary points of the approximating problems are stationary points of the original problem.

2.3.3 Formulation Based on Generalized Least pth Function

A different formulation was presented by Bandler and Abdel-Malek (1978a). They introduced a generalized least pth function to convert a tolerance and tuning problem to an equivalent tolerance problem. An expanded constraint region, namely the tunable constraint region R_{ct}, replaces the original region R_c (see Fig. 2.2). The region is given for $p = \infty$ by

$$R_{ct} \triangleq \{\phi \mid \max_{\rho \in R_\rho} \quad \min_{i \in I_c} \; g_i(\phi + T\rho) \geq 0\}, \tag{2.23}$$

where ϕ is given by (2.3). The authors based some definitions of yield upon R_{ct} and described worst-case design and worst-case centering.

2.4 SPECIAL CASES AND OBJECTIVE FUNCTIONS

Several objective functions (or cost functions) have been proposed (Pinel and Roberts 1972, Pinel 1973, Bandler 1974, Karafin 1974, Bandler and Liu 1974). In practice, a suitable modeling problem would have to be solved to determine the cost-tolerance-tuning relationship. We assume that the nominal parameter values,

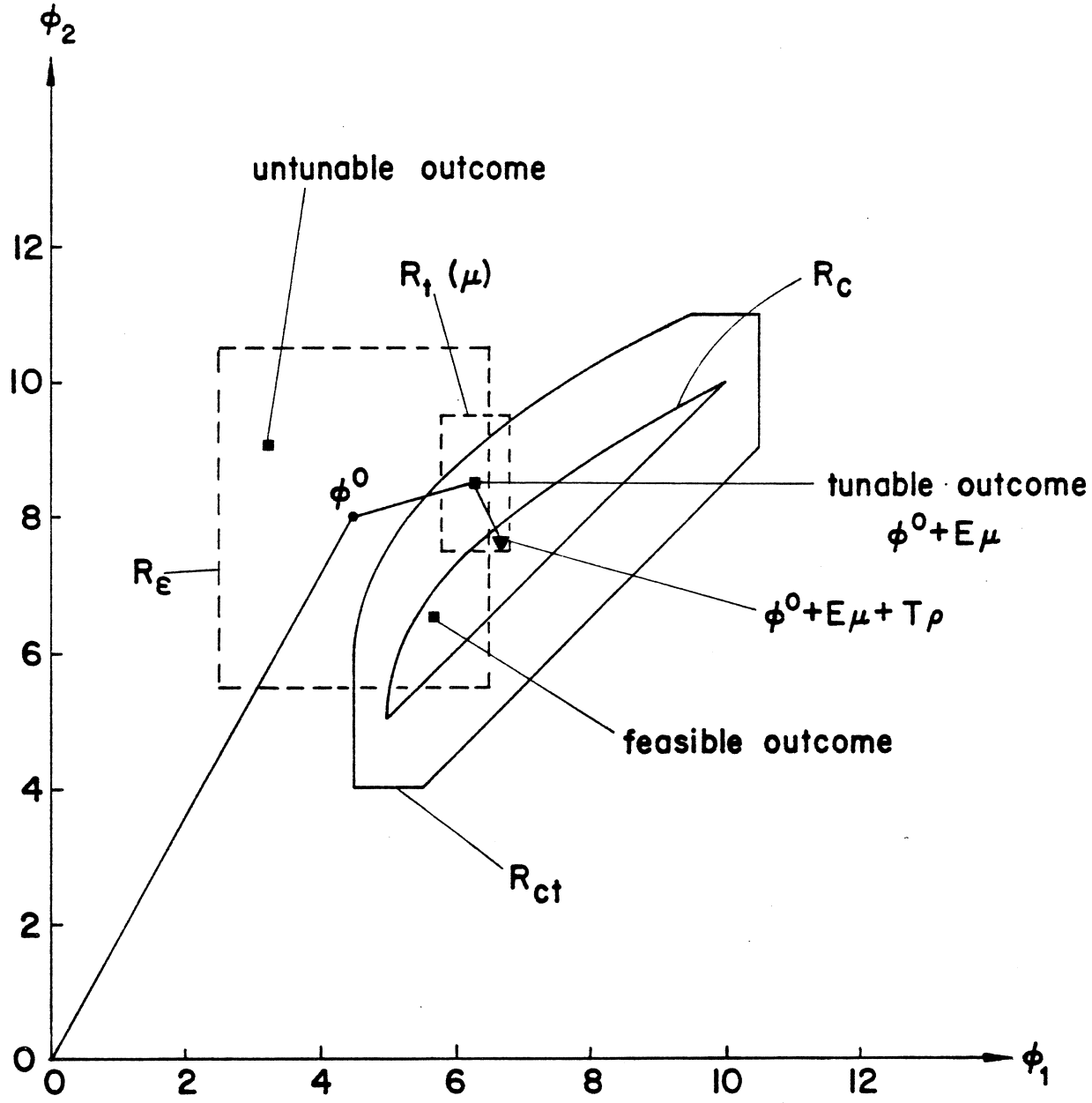


Fig. 2.2 Geometric interpretation of the tolerance problem equivalent to the tolerance-tuning problem. (Bandler and Abdel-Malek 1978a).

tolerances and tuning ranges (either absolute or relative) are the main variables and that the cost of the design is the sum of the cost of the individual components.

Suitable objective functions will be, for example, of the form

$$C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}, \mathbf{t}) = \sum_{i=1}^{n} \left(c_i \frac{\phi_i^0}{\varepsilon_i} + c_i' \frac{t_i}{\phi_i^0} \right), \tag{2.24}$$

where the c_i and c_i' are nonnegative constants. These may be set to zero if the corresponding element is not to be toleranced or tuned, respectively.

The special cases considered may be defined mathematically as a zero tolerance problem (ZTP), a fixed tolerance problem (FTP) and a variable tolerance problem (VTP) (Schjaer-Jacobsen and Madsen 1979, Bandler and Rizk 1979).

Schjaer-Jacobsen and Madsen (1979) define the problems in terms of a set of m nonlinear differentiable functions of n real variables. In this presentation we define those problems using notation and concepts directly related to the design problem (Bandler and Kellermann 1983).

We do not include tuning in this chapter since it will be considered in much more detail than other problems in Chapter 4.

2.4.1 The Zero Tolerance Problem (Centering Problem)

In this problem we have $\boldsymbol{\varepsilon} = \mathbf{0}$ and $\mathbf{t} = \mathbf{0}$. We want to find the nominal design $\boldsymbol{\phi}^0$ satisfying the design specifications $\mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}$, where $\boldsymbol{\phi} = \boldsymbol{\phi}^0$. The problem is a pure centering problem in which a feasible, centered nominal design is found if $R_c \neq \varnothing$. The solution may be useful at the initial stage of a design process when the designer has no prior experience with the problem and an initial rough approximation gives some insight.

The problem can be conveniently formulated in minimax form as

$$\underset{\boldsymbol{\phi}^0}{\text{minimize}}\ F(\boldsymbol{\phi}^0), \tag{2.25}$$

subject to

$$\boldsymbol{\phi}^0 \geq \mathbf{0}, \tag{2.26}$$

where

$$F(\boldsymbol{\phi}^0) \triangleq \max_{i \in I_c} (-g_i(\boldsymbol{\phi}^0)). \tag{2.27}$$

2.4.2 The Fixed Tolerance Problem

Here we have $\boldsymbol{\varepsilon} = \text{const} \neq \mathbf{0}$. We want to find $\boldsymbol{\phi}^0$, the center of the tolerance region R_ε, when the manufacturing tolerances on the components are fixed. This is basically a centering problem and can be formulated in minimax form as

$$\underset{\boldsymbol{\phi}^0}{\text{minimize}}\ F(\boldsymbol{\phi}^0) \tag{2.28}$$

subject to

$$\boldsymbol{\phi}^0 \geq \mathbf{0}, \tag{2.29}$$

where

$$\boldsymbol{\phi} = \boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu} \text{ for all } \boldsymbol{\mu} \in R_\mu \tag{2.30}$$

and

$$F(\boldsymbol{\phi}^0) = \max_{i \in I_c} (-g_i(\boldsymbol{\phi})). \tag{2.31}$$

Under certain assumptions (one-dimensional convexity of R_c, see Fig. 2.3) it is sufficient to choose only the vertices of R_ε to form appropriate minimax functions.

2.4.3 The Variable Tolerance Problem

In this problem we have $\boldsymbol{\varepsilon} \neq \text{const}$, $\mathbf{t} = \mathbf{0}$. The manufacturing tolerances are considered as variables instead of as being fixed.

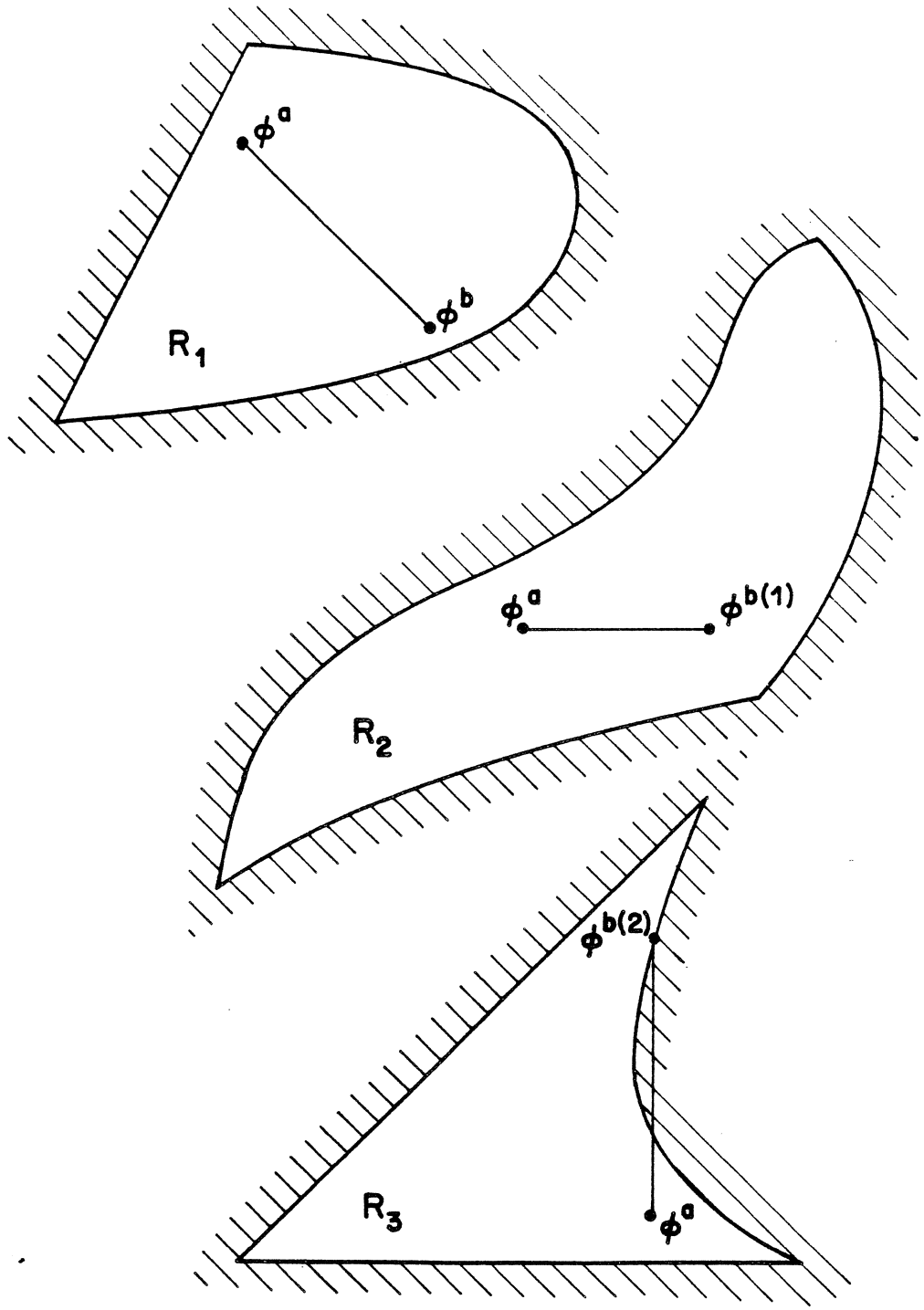


Fig. 2.3 Illustrations of convex, one-dimensionally convex and nonconvex regions (Liu 1975).

The design problem can be formulated as

$$\underset{\boldsymbol{\phi}^0,\,\boldsymbol{\varepsilon}}{\text{minimize}}\; C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}) \tag{2.32}$$

subject to

$$\boldsymbol{\phi}^0 \geq \mathbf{0}\,, \tag{2.33}$$

$$\boldsymbol{\varepsilon} \geq \mathbf{0}\,, \tag{2.34}$$

$$\mathbf{g}(\boldsymbol{\phi}) \geq \mathbf{0}\,, \tag{2.35}$$

where $\boldsymbol{\phi}$ is given by (2.3).

The objective function C is directly related to the component cost, and generally possesses the properties

$$C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}) \to \text{const} \;\text{ as }\; \boldsymbol{\varepsilon} \to \infty\,,$$

$$C(\boldsymbol{\phi}^0, \boldsymbol{\varepsilon}) \to \infty \qquad \text{as } \varepsilon_i \to 0\,.$$

A common form of this objective is

$$\sum_{i=1}^{n} c_i \frac{\phi_i^0}{\varepsilon_i}\,, \tag{2.36}$$

where the c_i's are positive constant weights.

2.4.4 Generalized Tolerance Assignment Problem

Tromp (1977) has generalized the tolerance assignment problem so that physical tolerances, model uncertainties, external disturbing effects and dependently toleranced parameters can be considered in a unified manner. In essence, the approach begins with the definitions of the k_{oi}-dimensional vector $\boldsymbol{\phi}^{oi}$, the k_i-dimensional vector $\boldsymbol{\phi}^i$ and the $k_{\mu i}$-dimensional vector $\boldsymbol{\mu}^i$ so that $\boldsymbol{\phi}^i$ is a function of $\boldsymbol{\phi}^{oi}$ and $\boldsymbol{\mu}^i$ for all $i = 1,2,\ldots,n$, and $\boldsymbol{\phi}^{oi}$ itself depends on all $\boldsymbol{\phi}^{i-1}$ for $i = 2,3,\ldots,n$.

Input parameters, e.g., the physical parameters available to the manufacturer might be identified as $\boldsymbol{\phi}^1$, whereas $\boldsymbol{\phi}^n$ would be the output vector, e.g.,

the sampled response of a system or the vector of constraints $\mathbf{g}$, which defined the region R_c. The quantities $\boldsymbol{\phi}^2, \ldots, \boldsymbol{\phi}^{n-1}$ can be identified, for example, as intermediate or model parameters. The variables $\boldsymbol{\mu}^i$, $i=1,2,\ldots,n$, create the unavoidable or undesirable fluctuations and generally embody the unknown or intangible. Hence we let

$$\boldsymbol{\phi}^0 \triangleq \begin{bmatrix} \boldsymbol{\phi}^{01} \\ \boldsymbol{\phi}^{02} \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{\phi}^{0n} \end{bmatrix}, \quad \boldsymbol{\phi} \triangleq \begin{bmatrix} \boldsymbol{\phi}^1 \\ \boldsymbol{\phi}^2 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{\phi}^n \end{bmatrix} \text{ and } \boldsymbol{\mu} \triangleq \begin{bmatrix} \boldsymbol{\mu}^1 \\ \boldsymbol{\mu}^2 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{\mu}^n \end{bmatrix}. \tag{2.37}$$

The tolerance region in the $\boldsymbol{\phi}$ space is obviously no longer restricted to be an orthotope in this formulation.

2.5 ALGORITHMS FOR DCTT PROBLEMS

In worst-case design the whole tolerance region has to lie in the constraint region, i.e., it is required that

$$R_\varepsilon \subset R_c .$$

This is design with 100% yield, where yield Y is given by

$$Y \triangleq \frac{\text{number of outcomes which meet specifications}}{\text{total number of outcomes}} .$$

The 2^n vertices of the tolerance region are usually the points considered as candidates for worst-case. There are two main reasons. The first is that it is impractical, or even impossible, to consider explicitly the infinite number of points contained in the tolerance region. The second is that one-dimensional convexity of the constraint region may be assumed. Bandler (1974) proved, in this case, that it is sufficient for worst-case design to require that

$$R_v \subset R_c .$$

An example of a worst case design and a design with yield $< 100\%$ is shown in Fig. 2.4.

2.5.1 Vertex Selection Schemes

For large problems, with a large number of variables, the number of vertices of the tolerance region becomes enormous. Selection schemes which include purging (dropping of constraints or vertices) as well as addition of vertices of the tolerance region during the optimization process alleviate the need for considering 2^n vertices (Bandler, Liu and Chen 1975, Bandler, Liu and Tromp 1976b, Tromp 1977). One of these schemes (Bandler, Liu and Tromp 1976b) is based on iterative solution of necessary conditions for the worst vertex derived from the Kuhn- Tucker conditions. These methods rely on the assumption that the constraint region is one-dimensionally convex.

Schjaer-Jacobsen and Madsen (1979) suggest the application of interval arithmetic for solving the worst-case problem which guarantees that the worst case is always found. By solving the worst-case problem as described by them, no information is gained about where in the tolerance interval the worst case is attained. In their method the one-dimensional convexity assumption is not required, and the worst case can lie at an edge of the tolerance region instead of at a vertex.

2.5.2 Simplicial and Quadratic Approximations

The tolerance problem described earlier implicitly solves the centering problem, in which we are interested in finding a "center" of the constraint region. Another approach is one developed by Director and Hachtel (1977). It is concerned

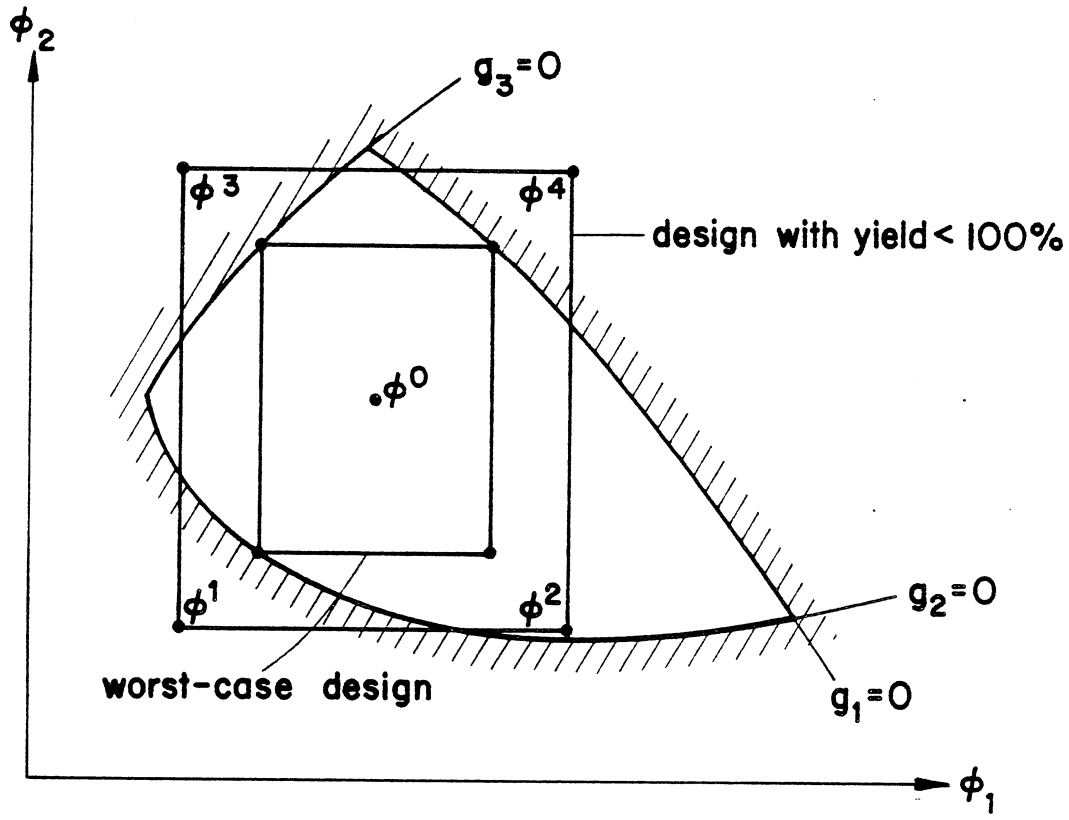


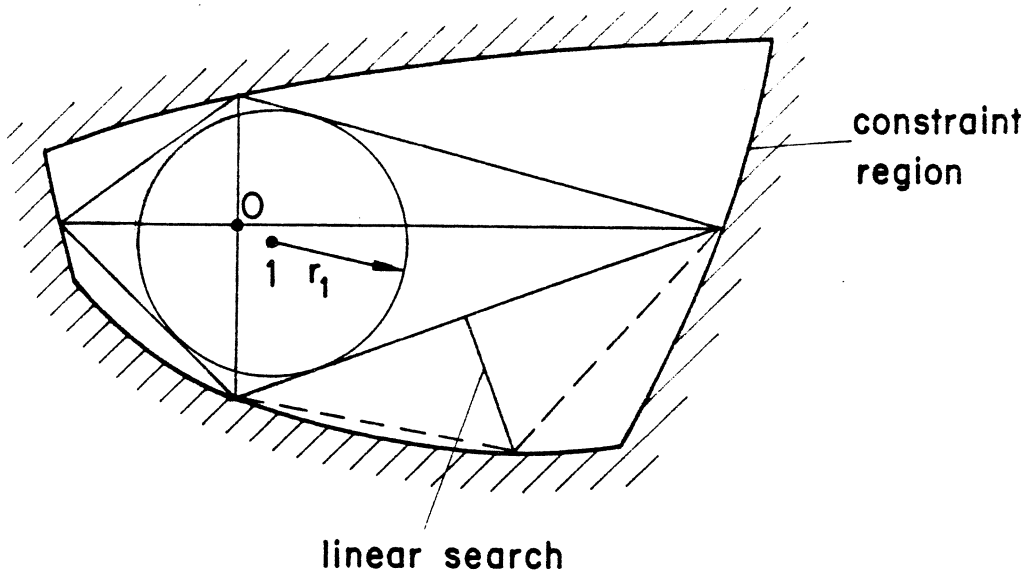
Fig. 2.4 Example of a worst-case design and a design with yield <100 percent. For the worst case design the set of active vertices is $S_{av} = \{1,3,4\}$. These vertices indicate critical regions where constraint violations are most likely to occur for a design with yield <100 percent (Bandler and Abdel-Malek 1978b).

with finding the center of the largest hypersphere inscribed in the constraint region (see Fig. 2.5). In the process, an internal approximation to the region is obtained. The problem of finding the largest hypersphere is solved by a sequence of linear programming problems.

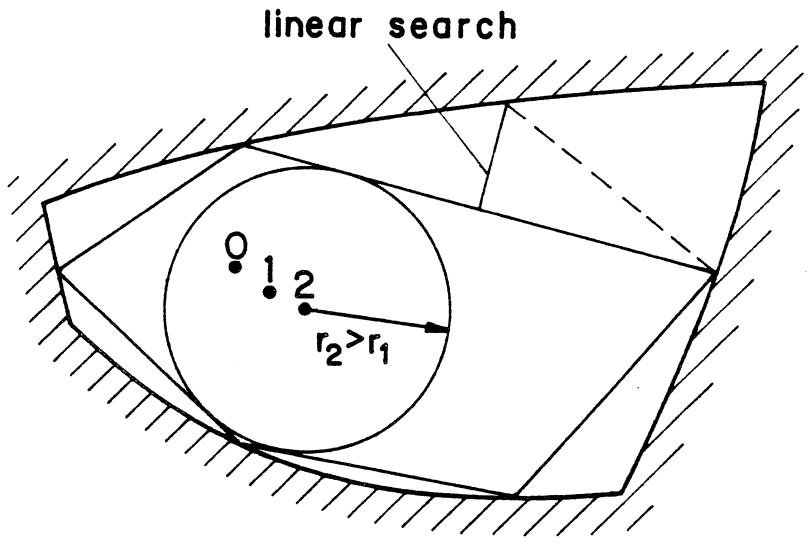
Bandler and Abdel-Malek (see Bandler and Abdel-Malek 1978b, Abdel-Malek and Bandler 1980a, 1980b) proposed a method (in the context of yield optimization) in which approximations are made to both the constraint region and the yield integral over the constraint region. The method assumes that the parameters of the circuit have a joint probability density function which is truncated or adequately represented by an orthotopic truncated distribution over a region with fixed volume but whose position depends on the nominal parameter values. Using the regionalization they approximate the failure rate (1–Y) in two steps. First the intersection of the tolerance region R_ε and the constraint region R_c is approximated with a quadratic approximation. This approximation is updated as the nominal point changes and only generated in those areas where $R_\varepsilon(\phi^0)$ violates the constraint region. Then the quadratic approximation is linearized about the points where the orthotope R_ε intersects the quadratic approximation to $R_\varepsilon \cap R_c$. Because their approximation to the failure rate is analytical it is possible to differentiate it to find the gradient of the yield w.r.t. the nominal point ϕ^0.

2.5.3 Cut-Map Algorithms for Tolerancing and Tuning

Many algorithms for design problems with parameter tolerances assume one- dimensional convexity of the constraint region (or the set of feasible nominal designs). Mayne, Polak and Voreadis (1982) presented an algorithm for the tolerance problem which is suitable for the nonconvex problem.



(a) Initial search for boundary points.



(b) The polytope approximating the boundary of the constraint region after two iterations.

Fig. 2.5 Illustration of the simplicial approximation approach (Director and Hachtel 1977).

The design problem requires the solution of the infinite dimensional inequality

$$\mathbf{g}(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}) \geq \mathbf{0} \quad \text{for all} \quad \boldsymbol{\mu} \in R_\mu .$$

Let R_c denote the feasible set for the tolerance problem, i.e., the problem where $\boldsymbol{\phi} = \boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}$ for all $\boldsymbol{\mu} \in R_\mu$. One method of solving such a problem is the outer approximations algorithm described in Mayne, Polak and Trahan (1979). This replaces the infinite-dimensional feasibility problem by an infinite sequence of conventional (finite-dimensional) feasibility problems where $R_\mu{}^i$ is a finite subset of R_μ.

The cut-map algorithm approximates R_c' (the complement of R_c) by the union of a finite number of very simply described regions. Typically, at iteration i, R_c' is approximated by $W_i \triangleq \cup \{B(\boldsymbol{\phi}_j{}^0, r_j) | \; j<i\}$ (or a subset of this set), where $B(\boldsymbol{\phi}_j{}^0, r_j)$ denotes an open ball with center $\boldsymbol{\phi}_j{}^0$ and radius $r_j > 0$, such that $R_c \cap B(\boldsymbol{\phi}_j{}^0, r_j) = \varnothing$.

Clearly $R_c \cap W_i = \varnothing$ so that $R_c \subset W_i'$, that is W_i' is an outer approximation to R_c of a particularly simple kind. The algorithm proceeds by determining any $\boldsymbol{\phi}_i{}^0$ in W_i' and then computing r_i; $r_i = 0$ implies that $\boldsymbol{\phi}_i{}^0$ lies in R_c. Rules for computing r_i and constructing W_i are given and represent extensions of the conceptual algorithms of Eaves and Zangwill (1971).

In Voreadis and Mayne (1982) the idea of cut-map algorithms is extended to the case when tuning is also present. Both algorithms, however, are suitable only for the case when the tolerance and tuning regions are constant.

2.5.4 Function Splitting for the Tolerance Problem

Sophisticated algorithms for the fixed and variable tolerance problem are presented by Brayton, Director, Hachtel and Vidigal (1979). One-dimensional

convexity is assumed. The nondifferentiability of the worst case function

$$\max_{\boldsymbol{\phi} \in R_\varepsilon} f_i(\boldsymbol{\phi}), \text{ or } \max_{\boldsymbol{\phi} \in R_\varepsilon} (-g_i(\boldsymbol{\phi})),$$

$i \in I_c$, is coped with by "function splitting". The idea of function splitting is to treat a single function as if it were simultaneously several different functions. Since the worst case of a single function can occur simultaneously at two different vertices, say $\boldsymbol{\mu}(i_1)$ and $\boldsymbol{\mu}(i_2)$, then it is natural to treat $f(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_1))$ and $f(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_2))$ as two separate functions and to employ both gradients, $\mathbf{f}'(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_1))$ and $\mathbf{f}'(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_2))$, in the optimization algorithm. The worst case function is not differentiable at $\boldsymbol{\phi}^0$ but possesses a generalized gradient which is the convex hull of $\mathbf{f}'(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_1))$ and $\mathbf{f}'(\boldsymbol{\phi}^0 + \mathbf{E}\,\boldsymbol{\mu}(i_2))$. A vertex list, updated at each iteration, defines the function and gradient information supplied to a quadratic program to determine a search direction.

2.5.5 MINMAX-MINBOX Linear Programming Approach

Hachtel, Scott and Zug (1980) proposed an interactive linear programming based method for optimization problems in worst case circuit design and device modeling. They propose a flexible objective MINMAX-MINBOX linear programming approach. The MINMAX linear programming design step, similar to the method developed by Schjaer-Jacobsen and Madsen (1979), asks the user to guess at the effective range of linearity of the specified objective functions, and then produces the minimum (over the n-dimensional design space) of the maximum (over the function indexes) function subject to this "box constraint". The MINBOX linear programming step asks the user to specify desired improved levels for the upper bounded objective functions. The MINBOX LP step either produces the smallest step $\Delta\boldsymbol{\phi}$ which achieves those levels or states that the levels are infeasible.

2.5.6 Related Algorithms

In the design of electronic circuits, as in all types of design, the engineer is faced with making a decision in the presence of competing objectives. Lightner and Director (1981) presented a technique for multiple criterion optimization (MCO). They proposed user oriented weight selection heuristics for the weighted ℓ_∞ solution to the MCO problem and generalized this idea for a family of weighted p-norms.

Vidigal and Director (1982) described a design centering algorithm for nonconvex regions of acceptability which is basically a convergent sequence of subproblems, each of which has a convex region of acceptability. Convergent algorithms exist for the solution of these subproblems, e.g., Director and Hachtel (1977) or Bandler and Abdel-Malek (1978b).

3

MINIMAX OPTIMIZATION TECHNIQUES IN COMPUTER-AIDED ENGINEERING

3.1 INTRODUCTION

A wide class of engineering system problems can be formulated as optimization problems with the objective being the norm of the error functions w.r.t. specified or measured responses of a system.

Many circuit design problems can be formulated naturally as minimax optimization problems. Most commonly, the minimax functions result from lower and/or upper specifications on a performance function of interest. In practice, we form error functions at a finite discrete set of frequencies, for example, and assume that a sufficient number of sample points have been chosen so that the discrete approximation problem adequately approximates the continuous problem. This may result in a large number of minimax functions to be minimized. Therefore, a highly efficient and fast algorithm for minimax optimization is of great importance to many system designers and engineers.

In this chapter, the area of nonlinear minimax optimization is briefly reviewed. The Hald and Madsen algorithm is treated in some detail. The ideas behind the algorithm are explained and illustrated with a microwave circuit example. We also present a comparison of the Hald and Madsen algorithm with other minimax algorithms, using a three-section transmission-line transformer problem.

A novel approach to worst-case tolerance design of circuits is proposed. Previous work in this area has been concentrated on worst-case design techniques

disregarding the source of the minimax functions, i.e., the discretization of a continuous problem. Our approach integrates a search technique for maxima of the response (a technique based on cubic interpolation) with the worst-case search using linearly constrained optimization.

3.2 INDEPENDENT PARAMETERS, PERFORMANCE FUNCTIONS AND SPECIFICATIONS

In electrical circuit design, more than one response function might have to meet given specifications. As an example, a circuit can be designed to meet desired specifications in both the frequency and the time domains (Bandler and Rizk, 1979). Graphically, this situation is shown in Fig. 3.1. In this case, we have more than one independent variable, ψ, namely, ψ^1, ψ^2, ..., ψ^λ, where λ is the number of these independent variables. Accordingly, we have λ response functions $F^1(\boldsymbol{\phi}, \psi^1)$, $F^2(\boldsymbol{\phi}, \psi^2)$, ..., $F^\lambda(\boldsymbol{\phi}, \psi^\lambda)$. In general, we can have λ upper specifications $S_u{}^1(\psi^1)$, $S_u{}^2(\psi^2)$, ..., $S_u{}^\lambda(\psi^\lambda)$ and λ lower specifications, $S_\ell{}^1(\psi^1)$, $S_\ell{}^2(\psi^2)$, ..., $S_\ell{}^\lambda(\psi^\lambda)$. The error functions will be of the form

$$e_u^j(\boldsymbol{\phi}, \psi^j) = w_u^j(\psi^j)\,(F^j(\boldsymbol{\phi}, \psi^j) - S_u^j(\psi^j)), \qquad j = 1, 2, ..., \lambda, \tag{3.1a}$$

$$e_\ell^j(\boldsymbol{\phi}, \psi^j) = w_\ell^j(\psi^j)\,(F^j(\boldsymbol{\phi}, \psi^j) - S_\ell^j(\psi^j)), \qquad j = 1, 2, ..., \lambda, \tag{3.1b}$$

where $w_u{}^j(\psi^j)$ and $w_\ell{}^j(\psi^j)$ are positive weighting functions and the subscripts u and ℓ refer to upper and lower specifications, respectively.

In a typical DCTT problem, the independent variable is the frequency and we are interested in the output response of the circuit at a discrete set of frequency points. Without loss of generality, we consider the following error functions

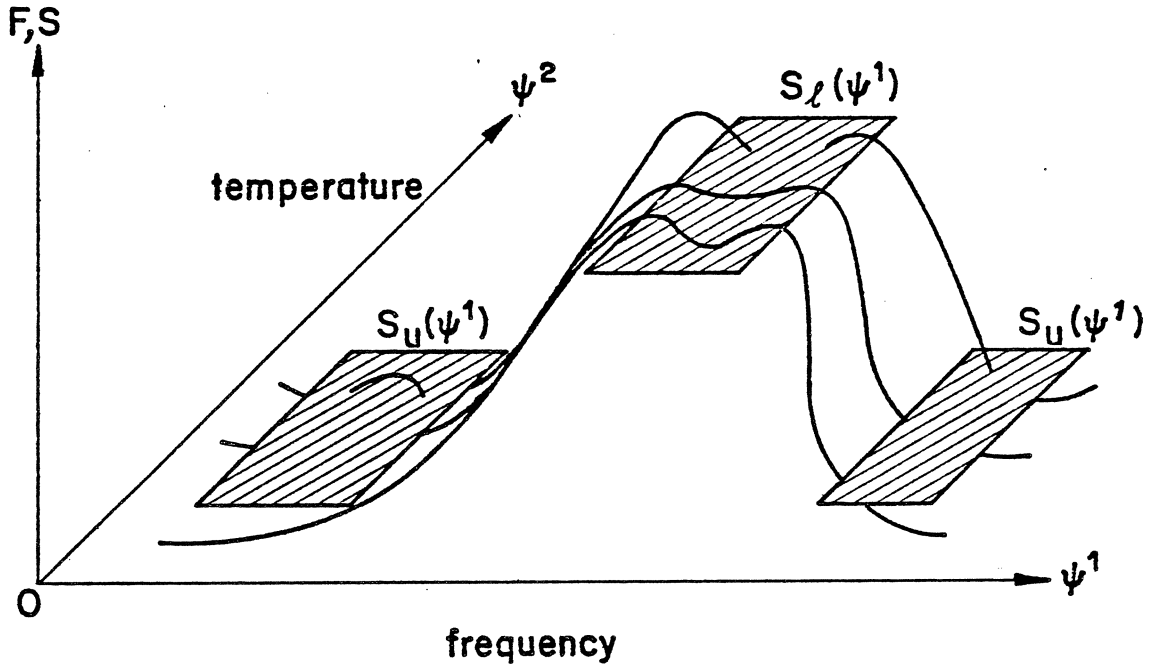


Fig. 3.1 Upper and lower specifications for an amplifier to be designed to operate over a specified temperature range (Bandler and Rizk 1979).

$$e_{ui}(\boldsymbol{\Phi}) \triangleq e_u(\boldsymbol{\Phi}, \psi_i) = w_{ui}(F_i(\boldsymbol{\Phi}) - S_{ui}), \quad i \in I_u, \tag{3.2a}$$

$$e_{\ell i}(\boldsymbol{\Phi}) \triangleq e_\ell(\boldsymbol{\Phi}, \psi_i) = w_{\ell i}(F_i(\boldsymbol{\Phi}) - S_{\ell i}), \quad i \in I_\ell, \tag{3.2b}$$

where

$$F_i(\boldsymbol{\Phi}) \triangleq F(\boldsymbol{\Phi}, \psi_i) \tag{3.2c}$$

I_u and I_ℓ are index sets, not necessarily disjoint. Let

$$f_i \triangleq \begin{cases} e_{uj}, & j \in I_u, \\ -e_{\ell k}, & k \in I_\ell, \end{cases} \quad i \in I_c, \tag{3.3}$$

where

$$I_u \triangleq \{1, 2, \ldots, n_u\}, \tag{3.4}$$

$$I_\ell \triangleq \{1, 2, \ldots, n_\ell\}, \tag{3.5}$$

$$I_c \triangleq \{1, 2, \ldots, m\}, \tag{3.6}$$

and $m = n_u + n_\ell$. The m functions

$$\mathbf{f} = [f_1 \;\; f_2 \;\; \ldots. \;\; f_m]^T \tag{3.7}$$

characterize the circuit performance, which is monitored during the optimization process.

If we let

$$M_f(\boldsymbol{\Phi}) \triangleq \max_{i \in I_c} f_i(\boldsymbol{\Phi}), \tag{3.8}$$

then the sign of M_f indicates whether the specifications are satisfied or violated.

3.3 REVIEW OF MINIMAX ALGORITHMS

3.3.1 Formulation of the Problem

The mathematical formulation of the linearly constrained minimax problem is the following. Let

$$f_j(\mathbf{x}) \triangleq f_j(x_1, x_2, \ldots, x_n), \quad j = 1, 2, \ldots, m,$$

be a set of m nonlinear, continuously differentiable functions. The vector $\mathbf{x} \triangleq [x_1 \; x_2 \; \ldots \; x_n]^T$ is the set of n parameters to be optimized.

We consider the optimization problem

$$\underset{\mathbf{x}}{\text{minimize}} \; F(\mathbf{x}) \triangleq \max_j f_j(\mathbf{x}) , \tag{3.9a}$$

subject to

$$\mathbf{a}_i^T \mathbf{x} + b_i = 0 , \qquad i = 1, 2, \ldots, \ell_{eq} , \tag{3.9b}$$

$$\mathbf{a}_i^T \mathbf{x} + b_i \geq 0, \qquad i = (\ell_{eq} + 1), \ldots, \ell , \tag{3.9c}$$

where $\mathbf{a}_i$ and b_i , $i = 1, 2, \ldots, \ell$, are constants.

3.3.2 Methods Based on Linearization

Over the past fifteen years, this type of problem has been considered by many researchers. Usually only the unconstrained problem is treated, however. But in some of the methods to be described, it is no complication and computationally costless to add linear constraints. Many of the minimax papers in the literature use the objective function

$$\hat{F}(\mathbf{x}) \triangleq \max_j |f_j(\mathbf{x})| ,$$

instead of F. There is no significant difference between these two optimization problems. We prefer (3.9) since it is notationally easier and more general.

One of the earliest methods for solving the minimax problem was that of Osborne and Watson (1969). At the kth iterate, $\mathbf{x}_k$, their method uses a linear approximation of the nonlinear minimax problem, namely,

$$\underset{\mathbf{h}}{\text{minimize}} \; \bar{F}(\mathbf{x}_k, \mathbf{h}) \triangleq \max_j \{f_j(\mathbf{x}_k) + \mathbf{f}_j'(\mathbf{x}_k)^T \mathbf{h}\} , \tag{3.10}$$

where $f_j'(\mathbf{x}_k)$ denotes the gradient vector of f_j w.r.t. $\mathbf{x}$ at the point $\mathbf{x}_k$. The minimizer $\mathbf{h}_k$ of (3.10) is found using linear programming and it is used in a line search. This method may be efficient, but often it is inefficient. No convergence can be guaranteed and the method can even provide convergence to a nonstationary point. Madsen (1975) incorporated trust regions in the Osborne and Watson method. The linearized problem (3.10) is solved subject to a local bound on the variable $\mathbf{h}$, the bound being adjusted during the iteration. No line search is used. This method has been proved to provide convergence to the set of stationary points and has a quadratic final rate of convergence when the solution is regular (Madsen and Schjaer-Jacobsen 1976). However, the rate of convergence may be very slow on singular problems.

The method of Anderson and Osborne (1977) is very similar to that of Madsen. The main difference lies in the way of bounding the step length $\|\mathbf{h}_k\|$. A different approach was used by Bandler and Charalambous (1972). They presented an approach utilizing efficient unconstrained gradient minimization techniques in conjunction with least pth objective functions employing extremely large values of p. Charalambous and Conn (1978) apply an active set strategy to obtain a direction for a line search.

All of these methods are first-order methods, i.e., the search is based on first-order derivatives only. Therefore, all of these methods have problems with singular solutions and the rate of convergence may be very slow.

3.3.3 Methods Using Second-Order Information

In order to overcome this problem, some second-order (or approximate second-order) information must be used. Hettich (1976) was the first who proposed doing this. He used a Newton iteration for solving a set of equations which expresses

the necessary conditions for an optimum. However, Hettich's method is only local. It is required that the initial point be close to the solution and that the set of active functions (and constraints) is known. Han (1981) suggested nonlinear programming techniques for solving the minimax problem. He uses a nonlinear programming formulation of the minimax problem which is solved via successive quadratic programming (Powell 1978). A line search is incorporated using the minimax objective function as merit function. Overton (1979) uses an approach similar to Han's, but solves equality constrained quadratic problems and uses a specialized line search.

The method of Watson (1979) is very similar to the method of Hald and Madsen (1981). It switches between a first- and a second-order method. The main differences between the Hald and Madsen algorithm and the Watson method are the following. Watson requires the user to provide exact first- and second-order derivatives whereas Hald and Madsen require only first-order derivatives. Furthermore, Watson fails to define a suitable set of criteria for switching between the first-order and the second-order methods. The Hald and Madsen method has guaranteed convergence to the set of stationary points whereas Watson's method has no such property. It can even provide convergence to a nonstationary point.

3.4 THE HALD AND MADSEN MINIMAX ALGORITHM IN SYSTEM DESIGN

3.4.1 General Description

The Hald and Madsen algorithm for nonlinear minimax optimization (Hald and Madsen 1981) is a combination of the first-order method of Madsen (1975) and an approximate second-order method. The first-order method provides fast convergence

to the neighbourhood of a solution. If the solution is singular (see Madsen and Schjaer-Jacobsen 1976), then the rate of convergence becomes very slow and a switch is made to the second-order method. Here, a quasi-Newton method is used to solve a set of nonlinear equations that necessarily hold at a local solution of (3.9). This method has superlinear final convergence. Several switches between the two methods may take place and the switching criteria ensure the global convergence of the combined method. The user of this algorithm is required to supply function values and first-order derivatives whereas the necessary second derivative estimates are generated by the algorithm.

For this thesis, we have used the MMLC version of the algorithm (Bandler and Zuberek 1982), based on the earlier implementation due to Hald (1981).

The algorithm is described in more detail in Appendix A, where the two methods, namely, the first-order method (denoted Method 1) and the approximate second-order method (denoted Method 2) are presented and the switching conditions between the two methods are given.

3.4.2 Performance of the Algorithm on Regular and Singular Problems

When the solution is singular, the final rate of convergence of Method 1 can be very slow. Consider the example of Fig. 3.2 in two variables, where the two functions are active at the solution $\mathbf{z}$ (i.e., $f_j(\mathbf{z}) = F(\mathbf{z})$ for two values of j). Figure 3.2 shows contours for a two-dimensional singular minimax problem arising from optimization of a two-section 10:1 transmission-line transformer, where the minimax functions correspond to the reflection coefficient sampled at 11 normalized frequencies with respect to 1 GHz (0.5, 0.6, ..., 1.4, 1.5). The optimization variables are characteristic impedances Z_1 and Z_2. Section lengths ℓ_1 and ℓ_2 are kept constant

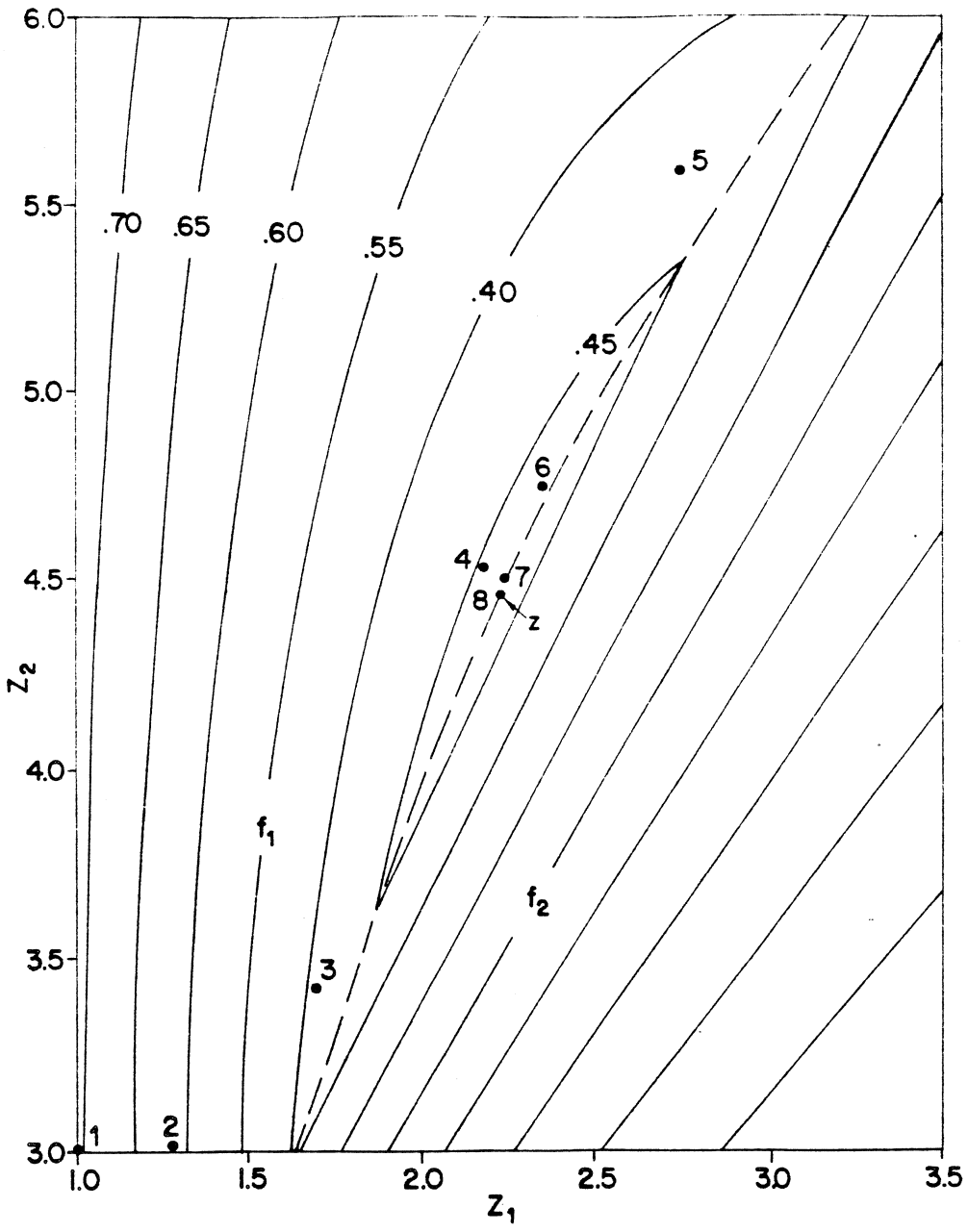


Fig. 3.2 Two-dimensional singular minimax problem arising from optimization of a two-section 10:1 transmission-line transformer with optimization parameters Z_1 and Z_2. The first 6 iterations are performed using Method 1 of the algorithm. Iterations 7 and 8 are performed using Method 2. The total number of iterations (function evaluations) to reach the solution with the accuracy of 10^{-6} is 11 (0.49s on Cyber 170/815). If Method 2 is not used 25 iterations (1.14s of CPU time) are required to reach the solution.

at their optimal value ℓ_q, which is the quarter wavelength at the center frequency. According to Madsen and Schjaer-Jacobsen (1976), this is a singular problem. Above the dotted line, F is equal to one of the functions f_j, $F(\mathbf{x}) = f_1(\mathbf{x})$, and below the dotted line, F is equal to another function, $F(\mathbf{x}) = f_2(\mathbf{x})$. At the dotted line, $f_1(\mathbf{x}) = f_2(\mathbf{x}) = F(\mathbf{x})$ and this line represents the bottom of a valley.

If f_1 and f_2 are different, then there is a kink at the bottom of the valley and a method based on linearization, such as Method 1, will provide fast convergence to this kink, as illustrated by the first three iterands in Fig. 3.2. After the dotted line has been reached, however, the convergence towards $\mathbf{z}$ can be slow because the iterands have to follow a curve which passes the solution $\mathbf{z}$ in a smooth manner (with no kink). Therefore, a method based on first-order derivatives only is not sufficient, in general, to give fast convergence. Some second-order (or approximate second-order) information is needed. The first six iterations are performed using Method 1 of the algorithm. Iterations 7 and 8 are performed using Method 2. The total number of iterations (function evaluations) to reach the solution with the accuracy of 10^{-6} is 11 (0.49 seconds on the Cyber 170/815). If Method 2 is not used at all, 25 iterations (1.14 seconds of CPU time) are required to reach the solution.

Notice that if three functions were equal at a minimum of a two-dimensional problem, then no smooth curve through the solution exists and Method 1 provides fast (quadratic) convergence to the solution. This is illustrated in Fig. 3.3, which shows contours for the same two-section 10:1 transmission-line transformer problem. However, the optimization variables are now ℓ_1/ℓ_q and Z_1. Characteristic impedance Z_2 and section length ℓ_2/ℓ_q are kept at their optimal values ($\ell_2/\ell_q = 1$, $Z_2 = 4.47213$). Here, the problem is regular and five iterations are sufficient to reach the vicinity of the solution. In the figure, the first five iterations shown are performed

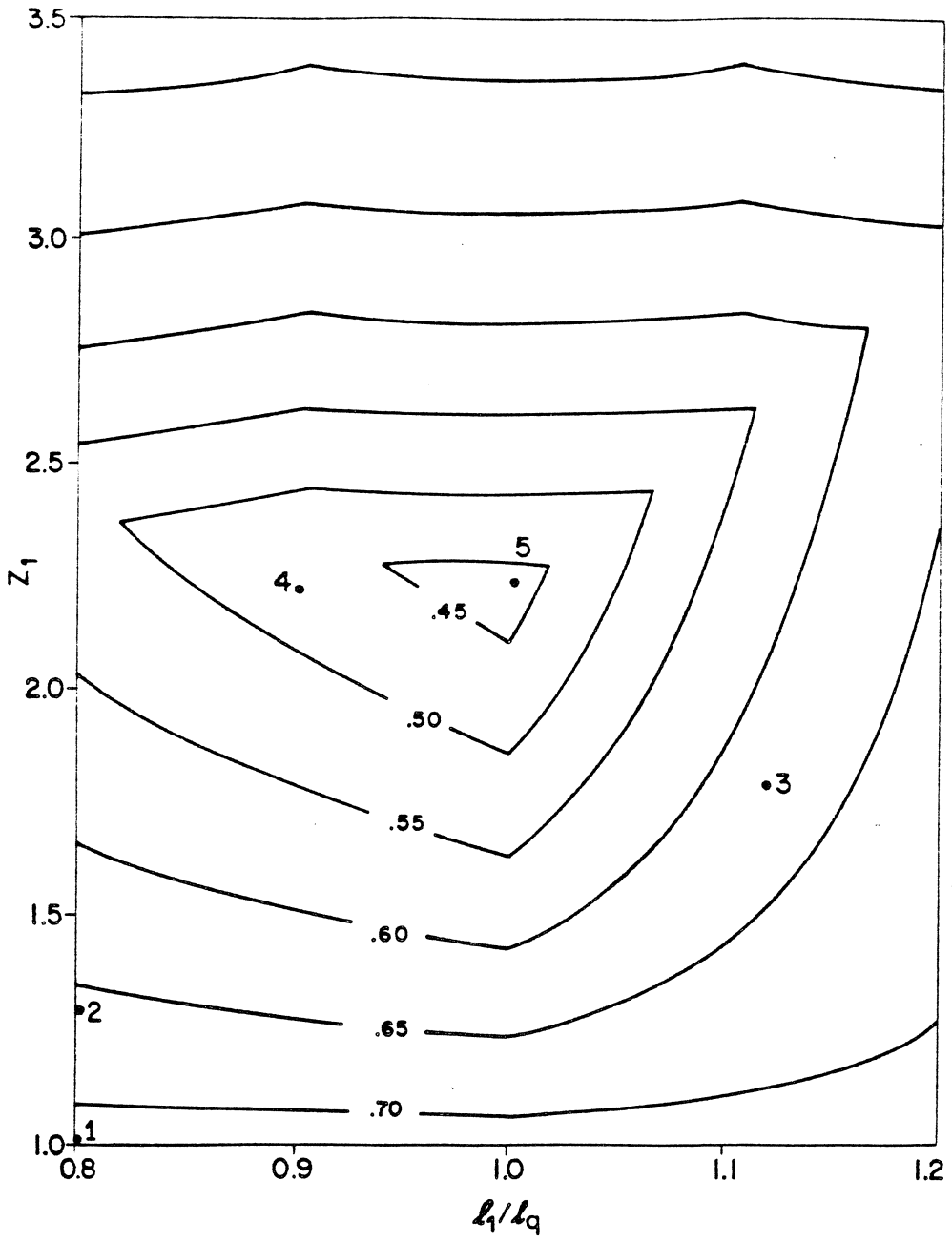


Fig. 3.3 Two-dimensional regular minimax problem arising from optimization of a two-section 10:1 transmission-line transformer with optimization parameters ℓ_1/ℓ_q and Z_1. The first 5 iterations shown are performed using Method 1. The total number of iterations to reach the solution with the accuracy of 10^{-6} is 8 (0.37s of CPU time on Cyber 170/815).

using Method 1. The total number of iterations to reach the solution with the accuracy of 10^{-6} is 8 (0.37 seconds of CPU time on the Cyber 170/815).

To show the influence of the parameters DX (initial step length of the iterative algorithm) and KEQS (the number of successive iterations with identical sets of active residual functions that is required before a switch to Stage 2 is made), the optimization has been performed several times for different values of DX and KEQS. The resulting numbers of residual function evaluations required to achieve the accuracy EPS = 10^{-6}, as well as the number of shifts to Stage 2, are summarized in Table 3.1 (the numbers of shifts are given in parentheses).

It can be observed that the increasing values of KEQS correspond to slightly increased numbers of function evaluations. Moreover, too small and too large values of DX require more residual function evaluations because of adjustments which are performed by the algorithm. From other experiments, it was observed that the increasing values of KEQS correspond, generally, to smaller numbers of shifts to Stage 2 (some too early shifts are eliminated).

3.5 COMPARISON OF MINIMAX ALGORITHMS FOR CIRCUIT DESIGN

3.5.1 The Test Problem

To compare the performance of minimax algorithms, a three-section, 100-percent relative bandwidth 10:1 transmission-line transformer problem has been chosen (see Fig. 3.4). It is a special case of an N-section transmission-line transformer. Originally studied by Bandler and Macdonald (1969a, 1969b), this type of test problem is now widely considered.

TABLE 3.1

THE INFLUENCE OF THE CONTROLLING PARAMETERS DX AND KEQS ON THE NUMBER OF FUNCTION EVALUATIONS

	KEQS		
DX	2	3	4
0.1	21(2)	23(2)	24(2)
0.25	19(2)	18(2)	19(2)
0.5	18(2)	20(2)	22(2)
0.75	18(2)	18(2)	20(2)
1.0	21(2)	22(2)	23(2)

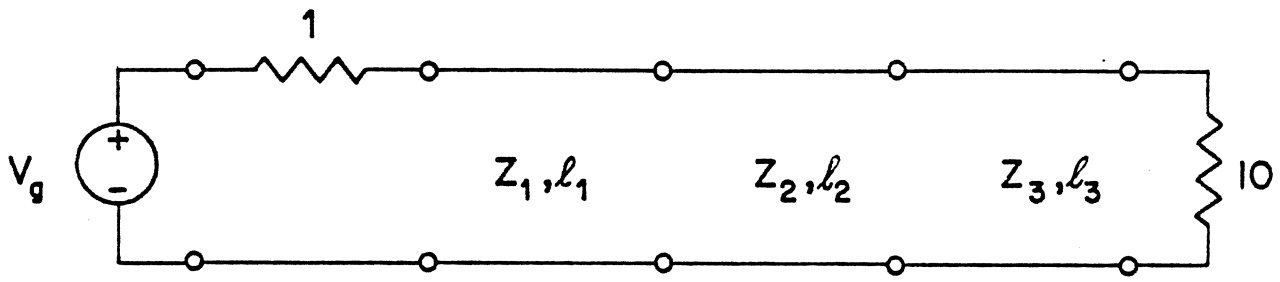


Fig. 3.4 Three-section, 10:1 transmission-line transformer used as a test problem to compare the performance of minimax algorithms.

The problem is to minimize the maximum reflection coefficient of this matching network. A detailed discussion on the formulation of direct minimax response objectives is presented in Bandler (1969).

Formally, the problem is to

$$\underset{\mathbf{x}}{\text{minimize}}\ F(\mathbf{x}) = \max_{[0.5\ ,\ 1.5]} |\rho(\mathbf{x}, \omega)|\ , \tag{3.11}$$

where

$$\mathbf{x} = [\ell_1/\ell_q \quad Z_1 \quad \ell_2/\ell_q \quad Z_2 \quad \ell_3/\ell_q \quad Z_3]^T\ .$$

The minimax functions represent the modulus of the reflection coefficient sampled at the 11 normalized frequencies ω (w.r.t. 1 GHz) {0.5, 0.6, 0.7, 0.77, 0.9, 1.0, 1.1, 1.23, 1.3, 1.4, 1.5}. The known quarter-wave solution is given by $\ell_1 = \ell_2 = \ell_3 = \ell_q$, $Z_1 = 1.63471$, $Z_2 = 3.16228$, $Z_3 = 6.11729$, where ℓ_q is the quarter wavelength at the center frequency, namely,

$$\ell_q = 7.49484 \text{ cm for 1 GHz}\ .$$

The corresponding maximum reflection coefficient is 0.19729. Two starting points have been used

$$\mathbf{x}_0^1 = [0.8 \quad 1.5 \quad 1.2 \quad 3.0 \quad 0.8 \quad 6.0]^T\ ,$$

$$\mathbf{x}_0^2 = [1.0 \quad 1.0 \quad 1.0 \quad 3.16228 \quad 1.0 \quad 10.0]^T\ .$$

Gradient vectors with respect to section lengths and characteristic impedances are obtained using the adjoint network method.

3.5.2 Performance of the Algorithms

Table 3.2 shows the performance of selected minimax algorithms on the test problem. Table 3.2 also shows results obtained using the algorithm published in Bandler, Kellermann and Madsen (1985b), with a cubic interpolation search for maxima of the response. Using this technique, the number of sample points can be

TABLE 3.2

COMPARISON OF ALGORITHMS FOR THE THREE-SECTION TRANSFORMER

(NUMBER OF FUNCTION EVALUATIONS)

Algorithm		Starting Point $\mathbf{x}_0{}^1$	Starting Point $\mathbf{x}_0{}^2$
Bandler, Kellermann and Madsen (1985b)[1]		18 *	21 **
Hald and Madsen (1981)		21	46
Agnew (1981)	Alg. I	23	64
	Alg. II	20	109
Bandler and Charalambous (1973)		95	155
Charalambous and Conn (1975)		162	67
Conn (1979)		67	80
Madsen (1975)		253	707
Madsen and Schjaer-Jacobsen (1976)		29	48
Bandler, Kellermann and Madsen (1985b)[2]		15 +	22 ++

Execution times on Cyber 170/815 in seconds are * 0.6 , ** 0.7 , + 0.57 , ++ 0.85

"Active" frequency points selected by the cubic interpolation search
0.50000, 0.76999, 1.23001, 1.50000

1 without cubic interpolation

2 with cubic interpolation

reduced from 11 to 4 and we do not have to know in advance the location of frequency points corresponding to the maxima of the response. More information on the cubic interpolation search technique is given in Section 3.6 in the context of a new approach to worst-case design of circuits.

The results published by Hald and Madsen (1981) correspond to the combined method as described here except that the PSB (Powell's symmetric Broyden) formula was used for updating the Jacobian in Method 2. Numerical results published in Bandler, Kellermann and Madsen (1985b) indicate that the use of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula as described in Appendix A is significantly better (see Table 3.2).

3.6 WORST-CASE NETWORK DESIGN

3.6.1 Preliminary Remarks

In this section we will formulate the fixed tolerance problem (FTP) (Bandler, Liu and Tromp 1976a, Schjaer-Jacobsen and Madsen 1979) embodying worst-case search and the selection of sample points for the discrete approximation of a continuous problem. As mentioned in the introduction, the discretization of a continuous problem may result in a large number of minimax functions to be minimized. The size of the problem increases dramatically if we want to consider tolerances on network parameters since for each frequency point selected to represent the response 2^n (n is the number of network parameters) minimax functions have to be created if we want to consider all vertices of the tolerance region.

A number of methods have been proposed for solving the worst-case problem. Schjaer-Jacobsen and Madsen (1979) suggest the application of interval

arithmetic. Bandler, Liu and Chen (1975) and Tromp (1977) described methods which rely on the assumption that the functions considered are one-dimensionally convex.

Our approach to the fixed tolerance problem is a double iterative algorithm. For each outer iteration of minimization first a search using cubic interpolation is done to determine frequency points which are candidates for active functions and then a number (equal to the number of selected minimax functions) of inner loop optimizations are performed to determine the worst case for each of the minimax functions.

The advantage of our approach is that the worst-case search (done by means of linearly constrained optimization) and the actual minimization are linked together such that each worst-case calculation affects immediately the outer iteration of minimization.

3.6.2 Cubic Interpolation Search Technique

The cubic interpolation technique allows us to consider the minimum number of frequency points to adequately approximate the continuous problem. In many cases the discretization of a continuous problem may not be adequate to give the continuous minimax solution. As illustrated in Fig. 3.5, the solution obtained using uniformly spaced sample points may not be optimal in the continuous minimax sense since some of the peaks of the response (or error function) would be missed. One way to overcome this difficulty is to use densely spaced sample points. This, however, may result in a prohibitively large number of minimax functions to be optimized. Therefore, it is desirable to develop a technique to locate the maxima of the response w.r.t. frequency and to track these maxima during the optimization process as they shift along the frequency axis due to the changes in the values of optimization

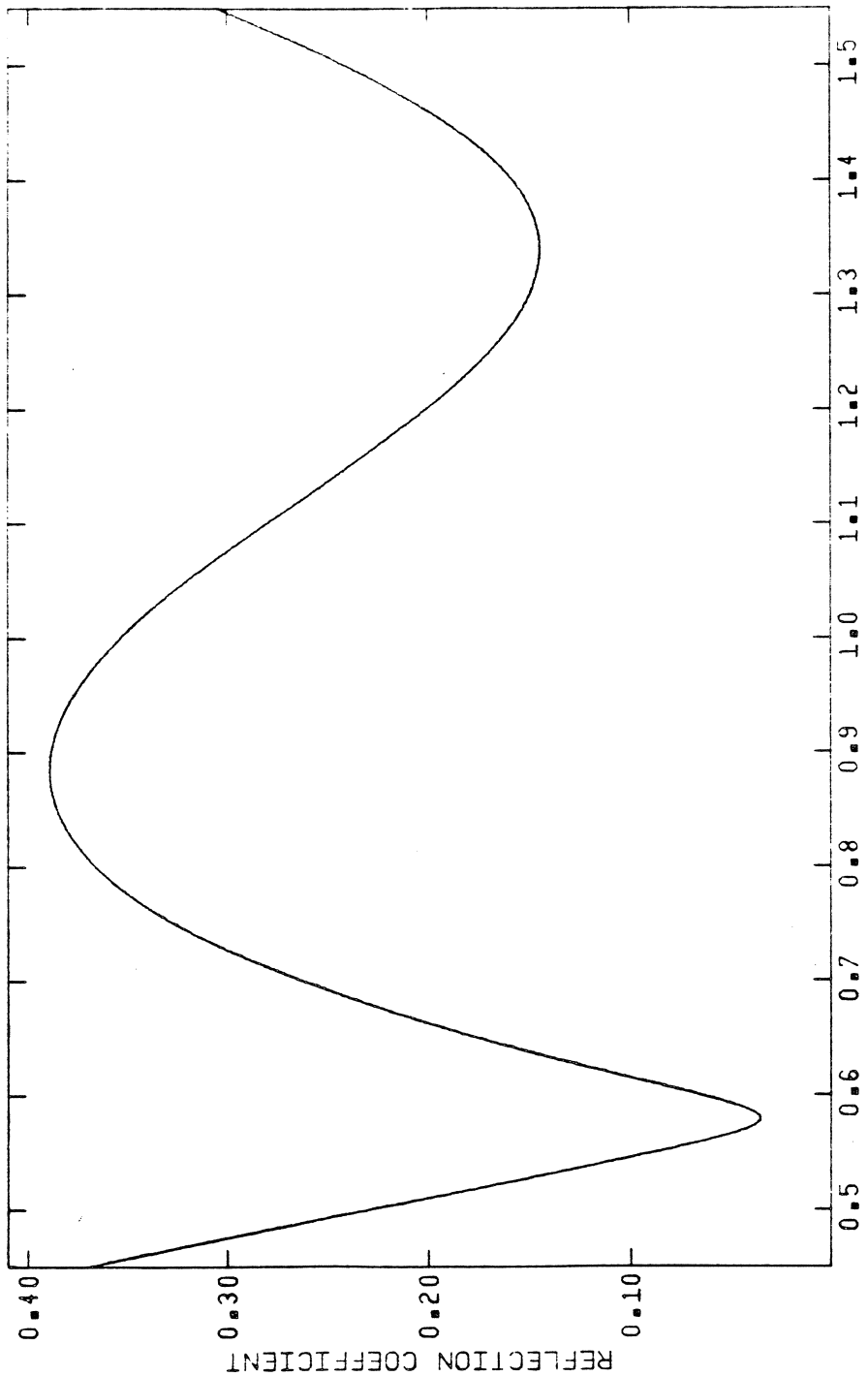


Fig. 3.5a Response of the three-section transmission-line transformer at the starting point $\mathbf{x}_1^0$. The initial sample points are 0.5, 0.8, 1.2 and 1.5. The uniformly spaced (with the step length of 0.1) search points are 0.5, 0.6, ..., 1.5. The sample points selected by the algorithm using cubic interpolation are 0.5, 0.884, 1.2 and 1.5. The edges of the frequency interval are kept as the sample points. The initial sample point 0.8 has been replaced by the point 0.884 since a maximum has been detected between 0.8 and 0.9. The sample point 1.2 has not been changed since no maximum has been found between points 1.2 and 1.3.

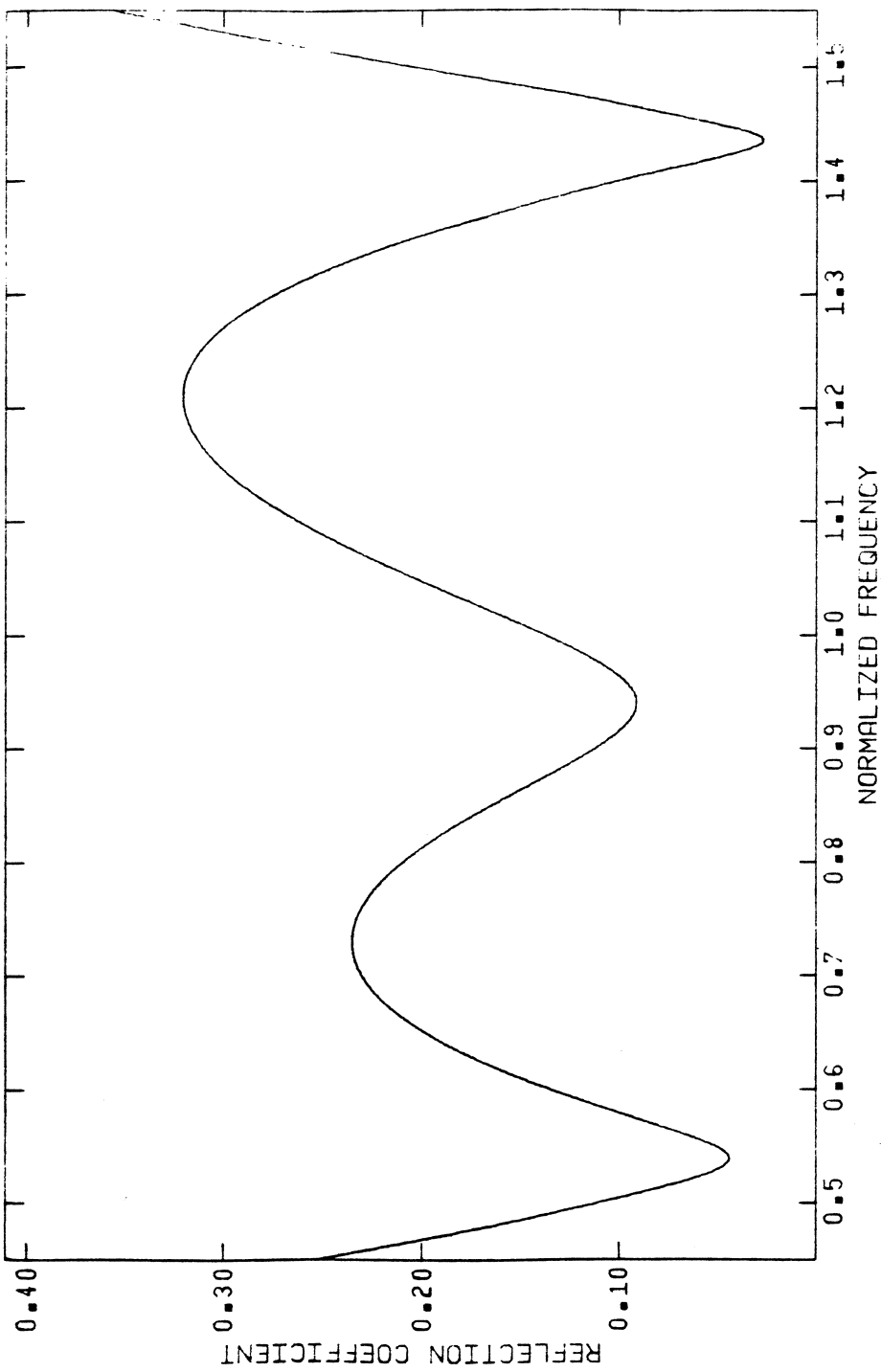


Fig. 3.5b Response of the three-section transmission-line transformer after the first iteration of optimization. The sample points selected by the cubic interpolation search (with the step length of 0.1) are 0.5, 0.729, 1.210 and 1.5. Now the initial sample point 1.2 has been replaced by 1.210 since a maximum has been detected between points 1.2 and 1.3 . The sample point 0.884 has been replaced by a new sample point 0.729 resulting from the maximum between 0.7 and 0.8.

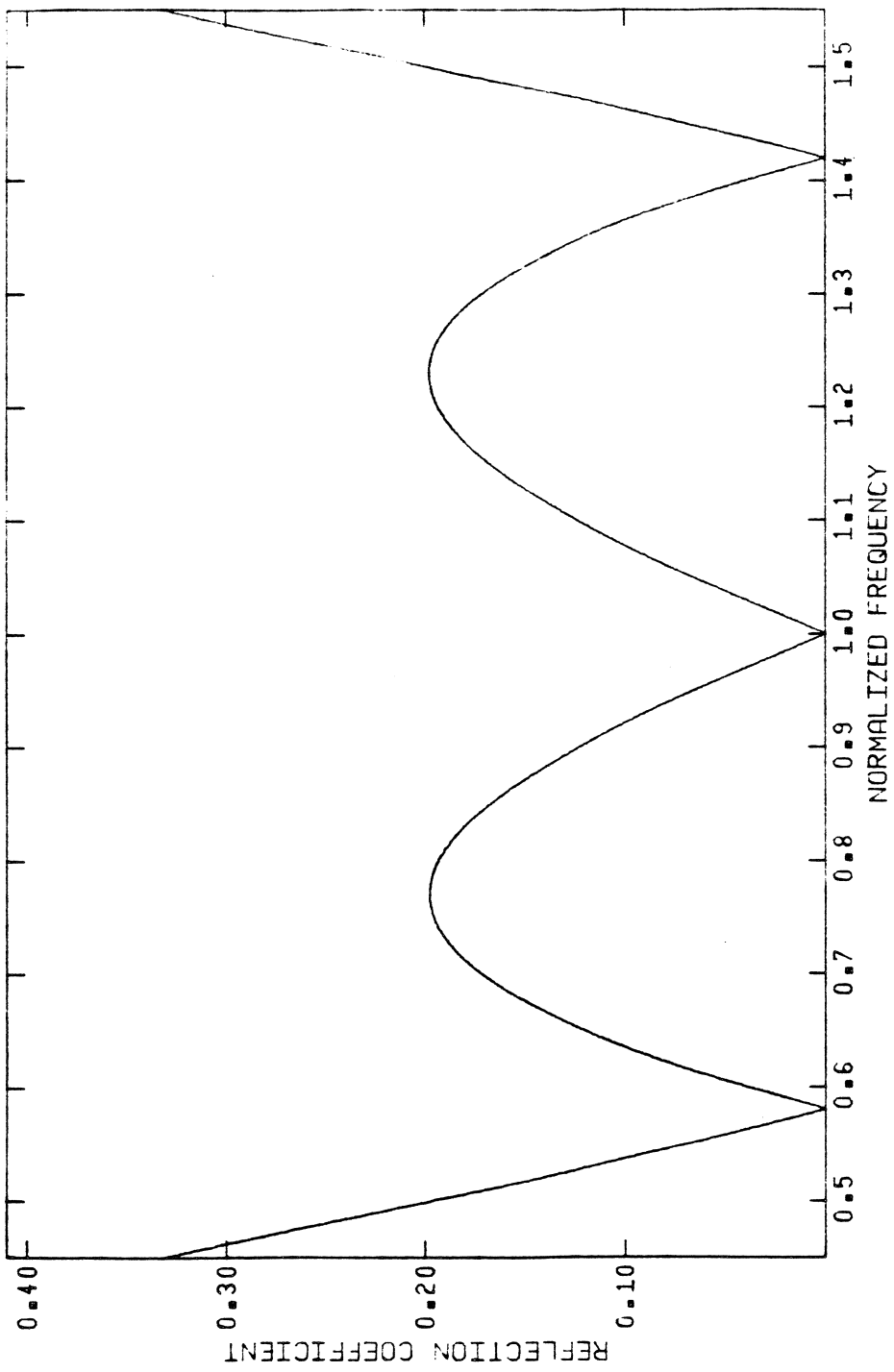


Fig. 3.5c Response of the three-section transformer at the solution. The sample points selected by the cubic interpolation search (with the step length of 0.1) before the last iteration of optimization are 0.5, 0.770, 1.230 and 1.5.

parameters. Such a technique has been developed by Bandler and Chen (1984). It is based on the cubic interpolation formulas of Fletcher and Powell (1963). For convenient reference the formulas are given in Appendix B.

3.6.3 Fixed Tolerance Problem Embodying Worst-Case Search and Selection of Sample Points

We consider a set of m nonlinear functions

$$f_j(\boldsymbol{\phi}^0) \triangleq f(\boldsymbol{\phi}^0, \omega_j)\,, \quad j \in J \triangleq \{1,2,..,m\}\,, \tag{3.12}$$

where ω_j, $j \in J$, is an independent parameter (frequency). The number of functions m is equal to

$$m = m_{max} + 2\,,$$

where m_{max} is the number of the maxima of the response and 2 represents the edges of the frequency interval $[\omega_\ell, \omega_u]$.

The fixed tolerance problem can be defined on the basis of the worst-case objective function (Schjaer-Jacobsen and Madsen 1979) as that of determining

$$\min_{\boldsymbol{\phi}^0} F(\boldsymbol{\phi}^0) = \min_{\boldsymbol{\phi}^0} \max_{j\in J} \max_{\boldsymbol{\phi}\in R_\varepsilon} f_j(\boldsymbol{\phi})\,. \tag{3.13}$$

For each outer iteration of minimization w.r.t. $\boldsymbol{\phi}^0$ m frequency points are determined (by a search technique based on cubic interpolation) and m linearly constrained optimizations are performed to find the worst cases.

At the kth outer iteration of minimization we have an approximation $\boldsymbol{\phi}_k^0$ of the solution and we solve m linearly constrained optimizations, where the jth problem, $j \in J$, is

$$\underset{\boldsymbol{\phi}_k}{\text{minimize}} \quad (-f_j(\boldsymbol{\phi}_k)) \tag{3.14a}$$

subject to

$$(\phi_i^0)_k - \varepsilon_i \le (\phi_i)_k \le (\phi_i^0)_k + \varepsilon_i , \qquad i = 1, 2, \ldots, n. \tag{3.14b}$$

Once $\boldsymbol{\phi}_k^*$ for the jth function is determined we can identify whether the worst-case occurred at a vertex using the following criteria.

Let

$$(y_i)_k = |(\phi_i^0)_k - (\phi_i^*)_k| . \tag{3.15}$$

If $|(y_i)_k - \varepsilon_i| \le 10^{-5}$, then the worst-case occurred at a vertex, for which μ_i, $i \in I$, are easy to determine

$$\mu_i = \begin{cases} -1 & \text{if} \quad (\phi_i^*)_k \le (\phi_i^0)_k \\ +1 & \text{otherwise} \end{cases} \tag{3.16}$$

The function values f_j, $j \in J$, and the gradients of f_j, $j \in J$, which are returned to the outer iteration are evaluated at a point $(\boldsymbol{\phi}_j^*)_k$, i.e., where the jth worst-case occurred.

3.6.4 Illustration of the Approach: Three-Section Transmission-Line Transformer

The three-section transmission-line transformer is used to illustrate the approach and its validity for worst-case design. Numerical results are summarized in Table 3.3. As expected the nominal parameter values are different from the values obtained for the nominal design problem. The location of the two internal maxima of the response has also changed as compared to the nominal design problem. Each linearly constrained optimization to determine worst-case for the particular frequency with the accuracy 10^{-3} requires about 4 iterations of the algorithm.

TABLE 3.3

FIXED TOLERANCE PROBLEM FOR THE

THREE-SECTION 10:1 TRANSFORMER

Number of Minimax Functions	4
Number of Variables	6
Required Accuracy of the Solution	10^{-5}
Assumed Tolerances	5%
Step Size in the Cubic Interpolation Search	0.1
Solution Vector	$\ell_1/\ell_q = 0.96373$
	$Z_1 = 1.67797$
	$\ell_2/\ell_q = 0.98720$
	$Z_2 = 3.22493$
	$\ell_3/\ell_q = 0.96483$
	$Z_3 = 6.04817$
"Active" Frequency Points	0.50000
	0.78726
	1.27242
	1.50000
Maximum Reflection Coefficient	0.33589
Number of Function Evaluations	32
Execution Time on Cyber 170/815 (in seconds)	8.1

3.7 CONCLUDING REMARKS

In this chapter, we have considered minimax optimization techniques in computer-aided design of engineering systems. The area of nonlinear minimax optimization has been briefly reviewed. Comparison made on the classical three-section transmission-line transformer problem shows clearly that the algorithm of Hald and Madsen is better than the other algorithms in terms of the number of function evaluations required to reach the solution with a desired accuracy.

We have presented a novel approach to worst-case tolerance design of circuits integrating a cubic interpolation based search technique for maxima of the response with the worst-case search using linearly constrained optimization. The validity of the approach has been demonstrated by solving a fixed tolerance problem for the three-section transmission-line transformer. We emphasize that our approach does not require the designer to know in advance the location of frequency points corresponding to the maxima of the response and significantly reduces the number of sample points adequately approximating the continuous response. This aspect of the approach is particularly important since it can reduce the number of minimax functions for which the worst cases have to be found.

4

ℓ_1 OPTIMIZATION TECHNIQUES IN COMPUTER-AIDED ENGINEERING

4.1 INTRODUCTION

The problem of minimizing the ℓ_1 norm of a set of nonlinear functions arises in a variety of areas. The most popular application of the ℓ_1 norm is the problem of approximating a function to data that might be contaminated with some wild points or gross errors. In this case the minimization of the ℓ_1 norm residual is superior to using other norms ℓ_p with $p > 1$ (Bartels and Conn 1981).

The number of applications of the ℓ_1 norm to circuit and system problems is increasing. The ℓ_1 norm has been successfully used to isolate the most likely faulty elements in fault isolation techniques for linear analog circuits (Bandler, Biernacki, Salama and Starzyk 1982).

Another important application of the ℓ_1 norm is the functional approach to post-production tuning (Bandler and Salama 1985), where the ℓ_1 type of objective function is used to select the number of tunable parameters needed to tune all possible outcomes of a manufactured design.

As the number of applications of the ℓ_1 norm to circuit and system problems is increasing so is the importance of fast and efficient nonlinear ℓ_1 optimization algorithms to circuit designers and engineers. We present a brief review of the existing ℓ_1 algorithms and concentrate in more detail on the Hald and Madsen algorithm.

The problem of tunable parameter selection in optimal DCTT is considered with the emphasis on the tuning problem at the design stage. The necessary

conditions for optimality of the nonlinear ℓ_1 problem with nonlinear constraints are derived and discussed in connection with the tuning problem.

We define regular and singular ℓ_1 problems and formulate a criterion for determining a singularity present in the ℓ_1 problem.

The properties of the Hald and Madsen ℓ_1 optimization algorithm are applied to fault isolation in linear analog circuits under an insufficient number of independent voltage measurements. A new formulation of the problem, a formulation based on an exact penalty function, is proposed and illustrated by a simple resistive network example.

In this chapter we also present a formulation using the ℓ_1 norm for model parameter identification problems and illustrate it with a 6th order multi-coupled cavity narrow band-pass filter.

4.2 REVIEW OF ℓ_1 ALGORITHMS

4.2.1 Formulation of the Problem

The optimization problem to be considered has the following mathematical formulation. Let

$$f_j(\mathbf{x}) \triangleq f_j(x_1, x_2, \ldots, x_n), \quad j = 1, 2, \ldots, m,$$

be a set of m nonlinear, continuously differentiable functions. The vector $\mathbf{x} \triangleq [x_1 \ x_2 \ldots x_n]^T$ is the set of n parameters to be optimized. We consider the following problem:

$$\underset{\mathbf{x}}{\text{minimize}} \quad F(\mathbf{x}) \triangleq \sum_{j=1}^{m} |f_j(\mathbf{x})| \tag{4.1a}$$

subject to

$$\mathbf{a}_i^T \mathbf{x} + b_i = 0, \qquad i = 1, 2, \ldots, \ell_{eq}, \tag{4.1b}$$

$$\mathbf{a}_i^T \mathbf{x} + b_i \geq 0, \qquad i = \ell_{eq} + 1, \ldots, \ell, \tag{4.1c}$$

where $\mathbf{a}_i$ and b_i, $i = 1, 2, \ldots, \ell$, are constants. This is called the linearly constrained ℓ_1 problem.

4.2.2 Algorithms for the Nonlinear ℓ_1 Problem

The problem (4.1) is, in principle, very similar to the linearly constrained minimax problem where the objective function is $F(\mathbf{x}) \triangleq \max|f_j(\mathbf{x})|$. Therefore, many of the algorithms for solving the minimax problem may be revised into algorithms for solving (4.1) and vice versa. For this reason most of the methods mentioned below have minimax counterparts. A survey of minimax algorithms has recently been given in Bandler, Kellermann and Madsen (1985b).

Most of the methods for minimizing the ℓ_1 function solve only the unconstrained problem (i.e. (4.1) with $\ell = 0$). For the type of methods to be described in this chapter, however, it is no complication and computationally costless to add the linear constraints.

One of the first attempts to solve the ℓ_1 problem was published in the paper of Osborne and Watson (1971). The method is iterative and at the kth iterate $\mathbf{x}_k$ the following linear approximation of the nonlinear ℓ_1 problem is used:

$$\underset{\mathbf{h}}{\text{minimize}} \quad F(\mathbf{x}_k, \mathbf{h}) \triangleq \sum_{j=1}^{m} |f_j(\mathbf{x}_k) + \mathbf{f}_j'(\mathbf{x}_k)^T \mathbf{h}| . \tag{4.2}$$

This linear model problem is solved using linear programming. The direction $\mathbf{h}_k$ found is then used in a line search. This method has quadratic final convergence under special circumstances but normally the final convergence is much slower. The global convergence properties of this method are rather poor, and like the Gauss-Newton method for nonlinear least squares (which is similar) the Osborne and Watson method may provide fast convergence to a nonstationary point, e.g., a point which is not a local minimum.

The more recent papers on the ℓ_1 problem use some second-order information. Most of the methods require that the user supply exact second (as well as first)

derivatives. To the best of our knowledge the method of Hald and Madsen (1985) is the first which uses approximate second-order information (i.e., it is a second-order method, but the user supplies only first derivatives). The methods of the next paragraph use exact second-order information.

El-Attar, Vidyasagar and Dutta (1979) use a sequence of smooth problems approximating the (nondifferentiable) ℓ_1 problem. Each of the smooth problems is solved by standard techniques and the sequence of solutions will often converge to a solution of the ℓ_1 problem. However, this kind of method may have severe ill-conditioning problems near an ℓ_1 solution because a nondifferentiable function with a kink is approximated by smooth functions. This gives curvatures in the smooth functions which tend to infinity as the ℓ_1 solution is approached. Murray and Overton (1981) use a nonlinear programming formulation of the ℓ_1 problem and apply successive quadratic programming. A special line search algorithm is used to obtain a reduction in the ℓ_1 objective function. The algorithm of McLean and Watson (1980) is a hybrid method like the method of Hald and Madsen (1985). It combines a first-order method based on (4.2) using trust regions with a Newton iteration. The first-order method is intended to be used initially, and close to a solution the Newton method should be used. This method often converges rapidly to a solution but the rules for switching between the two stages do not guarantee convergence. In fact the method may converge to a nonstationary point.

The linearly constrained ℓ_1 problem may be formulated as a nonlinear programming problem. Then it can be solved by standard techniques from that field. When Powell's (Powell 1978) method for nonlinear programming is applied to the ℓ_1 problem we obtain a method which in its final stages is very similar to the Hald and Madsen method. It can be shown that in the neighbourhood of a local solution of (4.1) their method generates the same points as Powell's method. However, in the latter method a quadratic program must be solved in every iteration, whereas Hald and

Madsen have to solve only a set of linear equations in the neighbourhood of a solution. Therefore, the computational effort used per iteration with their method is normally much smaller.

4.3 THE HALD AND MADSEN ALGORITHM FOR NONLINEAR ℓ_1 OPTIMIZATION

The algorithm to be described in this section is based on the work of Hald and Madsen (1985). It is a hybrid method combining a first-order method with an approximate second-order method. The first-order method is a robust trust region method which provides convergence to the neighbourhood of a solution. It is based on linear model problems of the type (4.2). These are solved subject to the constraints of the original problem (4.1) and a bound on the step length $\|\mathbf{h}\|$. The latter bound reflects the neighbourhood of the iterate $\mathbf{x}_k$ in which the kth model function (see (4.2)) is a good approximation to the nonlinear ℓ_1 function. If the solution approached by the first-order method is "singular" (see below) then a higher-order method must be used in order to obtain a fast ultimate rate of convergence. Therefore a switch is made to a quasi-Newton method that solves a set of nonlinear equations that necessarily hold at a solution of (4.1). This method has superlinear final convergence. Several switches between the first-order and the quasi-Newton method may take place. The reason for allowing this is that the latter method works only close to a solution, so if it is started too early a switch back to the (more robust) trust region method is necessary. Notice that the user of this algorithm is required to supply function values and first-order derivatives, whereas the necessary second derivative information is generated by the algorithm.

The algorithm is described in more detail in Appendix C, where the two methods, namely, the first-order method (denoted Method 1) and the approximate

second-order method (denoted Method 2) are presented and the switching conditions between the two methods are given.

4.4 NECESSARY CONDITIONS FOR OPTIMALITY OF THE NONLINEAR ℓ_1 PROBLEM WITH NONLINEAR CONSTRAINTS

As indicated in the introduction the ℓ_1 objective function plays an important role in many circuit and systems problems. Since many of these problems are formulated as ℓ_1 optimization problems with nonlinear constraints, it may be useful to derive necessary conditions for optimality and use them to get some insight into the features of the ℓ_1 norm in engineering problems.

In the following derivation we use the approach taken by El-Attar, Vidyasagar and Dutta (1979) and extend it to the case of nonlinear constraints.

The nonlinear ℓ_1 problem with nonlinear constraints may be stated as

$$\underset{\mathbf{x}}{\text{minimize}} \quad F(\mathbf{x}) \triangleq \sum_{j=1}^{m} |f_j(\mathbf{x})| . \tag{4.3a}$$

subject to

$$g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m_c , \tag{4.3b}$$

where g_i: $R^n \rightarrow R$ are, in general, nonlinear constraints.

Problem (4.3) can be transformed into the following nonlinear programming problem

$$\underset{\mathbf{x}, \mathbf{y}}{\text{minimize}} \quad F(\mathbf{x}, \mathbf{y}) \triangleq \sum_{i=1}^{m} y_i . \tag{4.4a}$$

subject to

$$y_i - f_i(\mathbf{x}) \geq 0 , \quad i = 1, 2, \ldots, m , \tag{4.4b}$$

$$y_i + f_i(\mathbf{x}) \geq 0 , \quad i = 1, 2, \ldots, m , \tag{4.4c}$$

$$g_i(\mathbf{x}) \geq 0 , \quad i = 1, \ldots, m_c , \tag{4.4d}$$

where the $f_i(\mathbf{x})$, $g_i(\mathbf{x})$ are as in (4.3) and F: $R^{n+m} \rightarrow R$ is a new objective.

The gradient of the objective function is

$$\mathbf{F}' = \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix}, \tag{4.5}$$

where $\mathbf{u} = [1\ 1\ \ldots\ 1]^T$ is an m-dimensional vector representing the gradient w.r.t. $\mathbf{y}$ and $\mathbf{0} = [0\ 0\ \ldots\ 0]^T$ is an n-dimensional vector representing the gradient w.r.t. $\mathbf{x}$.

Suppose that $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to (4.4), then

$$y_i^* = |f_i(\mathbf{x}^*)|, \quad i = 1, 2, \ldots, m. \tag{4.6}$$

Define the sets

$$I(\mathbf{x}^*) \triangleq \{i \mid f_i(\mathbf{x}^*) > 0\}, \tag{4.7}$$

$$J(\mathbf{x}^*) \triangleq \{i \mid f_i(\mathbf{x}^*) < 0\}, \tag{4.8}$$

$$Z(\mathbf{x}^*) \triangleq \{i \mid f_i(\mathbf{x}^*) = 0\}, \tag{4.9}$$

$$A(\mathbf{x}^*) \triangleq \{i \mid g_i(\mathbf{x}^*) = 0\}. \tag{4.10}$$

The gradients of the active constraints for the problem (4.4) are given by

$$\begin{bmatrix} \mathbf{e}_i \\ -\mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix}, \quad i \in I(\mathbf{x}^*), \tag{4.11}$$

$$\begin{bmatrix} \mathbf{e}_i \\ \mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix}, \quad i \in J(\mathbf{x}^*), \tag{4.12}$$

$$\begin{bmatrix} \mathbf{e}_i \\ -\mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix}, \begin{bmatrix} \mathbf{e}_i \\ \mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix}, \quad i \in Z(\mathbf{x}^*), \tag{4.13}$$

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{g}_i'(\mathbf{x}^*) \end{bmatrix}, \quad i \in A(\mathbf{x}^*), \tag{4.14}$$

where $\mathbf{e}_i$ is an m-dimensional vector with 1 in the ith position and zeros elsewhere.

By applying the Kuhn-Tucker conditions we get the following necessary conditions for optimality

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix} + \sum_{i\in I} \lambda_i \begin{bmatrix} -\mathbf{e}_i \\ \mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix} + \sum_{i\in J} \lambda_i \begin{bmatrix} -\mathbf{e}_i \\ -\mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix} +$$

$$\sum_{i\in Z} \left\{ \lambda_i \begin{bmatrix} -\mathbf{e}_i \\ \mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix} + \gamma_i \begin{bmatrix} -\mathbf{e}_i \\ -\mathbf{f}_i'(\mathbf{x}^*) \end{bmatrix} \right\} + \sum_{i\in A} \lambda_i \begin{bmatrix} \mathbf{0} \\ -\mathbf{g}_i'(\mathbf{x}^*) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \tag{4.15}$$

where $\lambda_i \geq 0$ and $\gamma_i \geq 0$ are the corresponding multipliers.

Splitting equation (4.15) we get

$$\begin{aligned} &\sum_{i\in I} \lambda_i \mathbf{f}_i'(\mathbf{x}^*) + \sum_{i\in J} \lambda_i \left(-\mathbf{f}_i'(\mathbf{x}^*)\right) + \sum_{i\in Z} (\lambda_i - \gamma_i) \mathbf{f}_i'(\mathbf{x}^*) + \\ &\quad + \sum_{i\in A} \lambda_i \left(-\mathbf{g}_i'(\mathbf{x}^*)\right) = \mathbf{0}, \\ &\lambda_i = 1, \ i \in I, \\ &\lambda_i = 1, \ i \in J, \\ &\left.\begin{aligned} &\lambda_i + \gamma_i = 1 \\ &\gamma_i \geq 0, \lambda_i \geq 0 \end{aligned}\right\}, i \in Z, \\ &\lambda_i \geq 0, \ i \in A, \end{aligned} \tag{4.16}$$

or

$$\begin{aligned} &\sum_{i\notin Z} \sigma_i \mathbf{f}_i'(\mathbf{x}^*) + \sum_{i\in Z} \delta_i \mathbf{f}_i'(\mathbf{x}^*) = \sum_{i\in A} \lambda_i \mathbf{g}_i'(\mathbf{x}^*), \\ &-1 \leq \delta_i \leq 1, \ i \in Z, \\ &\lambda_i \geq 0, \ i \in A, \end{aligned} \tag{4.17}$$

where

$$\sigma_i \triangleq \operatorname{sign} f_i(\mathbf{x}^*). \tag{4.18}$$

The necessary conditions for optimality of the nonlinear ℓ_1 problem indicate that zero functions $f_i(\mathbf{x}^*)$, $i \in Z$, play an important role in the characteristics of the ℓ_1 problem. They are called active ℓ_1 functions since they characterize an ℓ_1 solution in a similar way as the active functions in the nonlinear minimax problem. At an ℓ_1 solution all the functions are naturally divided into two sets, one containing zero functions and the other nonzero functions. This fact may be very useful in many

engineering problems which require some kind of a discrete selection or classification process if only we can formulate the problem in such a way that zero (or nonzero) functions at an ℓ_1 solution have a unique interpretation in terms of a practical engineering problem.

Once again an analogy can be drawn to the minimax problem in which minimax functions (error functions) result from lower and/or upper specifications on a performance function of interest and the sign of the minimax objective indicates whether the specifications are satisfied or violated.

4.5 REGULAR AND SINGULAR ℓ_1 PROBLEMS

Madsen and Schjaer-Jacobsen (1976) presented a theoretical treatment of singularities in nonlinear minimax optimization problems, which allows for a classification of regular and singular problems. Similar concepts apply to nonlinear ℓ_1 problems (Bandler, Kellermann and Madsen 1985c).

In order to simplify the notation assume that the first r functions are active in the nonlinear ℓ_1 problem, i.e., $Z = \{1, 2, ..., r\}$ (see (4.9) for the definition of Z).

Definition 1

The unconstrained ℓ_1 problem is singular with respect to the solution $\mathbf{x}^*$ if the matrix

$$\mathbf{D}_a \triangleq [\mathbf{f}_1'(\mathbf{x}^*)\ \mathbf{f}_2'(\mathbf{x}^*) \ldots \mathbf{f}_r'(\mathbf{x}^*)] \tag{4.19}$$

has rank less than n. Otherwise the problem is regular.

When constraints are present, active constraints play the same role as functions whose values are zero at the solution.

Assume that the first m_a constraints are active in the nonlinear ℓ_1 problem, i.e., $A = \{1, 2, ..., m_a\}$ (see (4.10) for the definition of A).

Definition 2

The constrained ℓ_1 problem is singular with respect to the solution $\mathbf{x}^*$ if the matrix

$$\mathbf{D} \triangleq [\mathbf{f}_1'(\mathbf{x}^*)\ \mathbf{f}_2'(\mathbf{x}^*) \ldots \mathbf{f}_r'(\mathbf{x}^*)\ \mathbf{g}_1'(\mathbf{x}^*)\ \mathbf{g}_2'(\mathbf{x}^*) \ldots \mathbf{g}_{m_a}'(\mathbf{x}^*)] \tag{4.20}$$

has rank less than n. Otherwise the problem is singular.

Normally a problem is regular if the (total) number of zero functions and active constraints is at least n.

Consider a two-dimensional example with two functions and no constraints

$$\begin{aligned} f_1(\mathbf{x}) &= (x_1 - 1)^2 + x_2^2, \\ f_2(\mathbf{x}) &= x_1^2 - x_2. \end{aligned} \tag{4.21}$$

Contours of F near the solution are given in Fig. 4.1. The solution is at $\mathbf{x}^* = [0.59\ 0.35]^T$, where $f_1(\mathbf{x}^*) = 0.28927 > 0$ and $f_2(\mathbf{x}^*) = 0$. This is an example of a singular ℓ_1 problem since the number of functions whose values are zero at the solution is less than n (the number of variables).

4.6 TUNABLE PARAMETER SELECTION IN THE OPTIMAL DCTT PROBLEM

4.6.1 Approaches to the Tuning Problem

Computer-aided designers have approached the tuning problem in two ways, each emphasizing one distinct facet. Before production, at the design stage, one can consider tuning as an integral part of the design process (Bandler, Liu and Tromp (1976a), Polak and Sangiovanni-Vincentelli (1979)), the objective being to relax the tolerances on the circuit components and compensate for the uncertainties in the model parameters. It is often necessary to introduce tuning parameters in order to obtain a feasible design. Sometimes, this is the case when tolerances are fixed and there is no solution to the FTP satisfying all design specifications. The integral

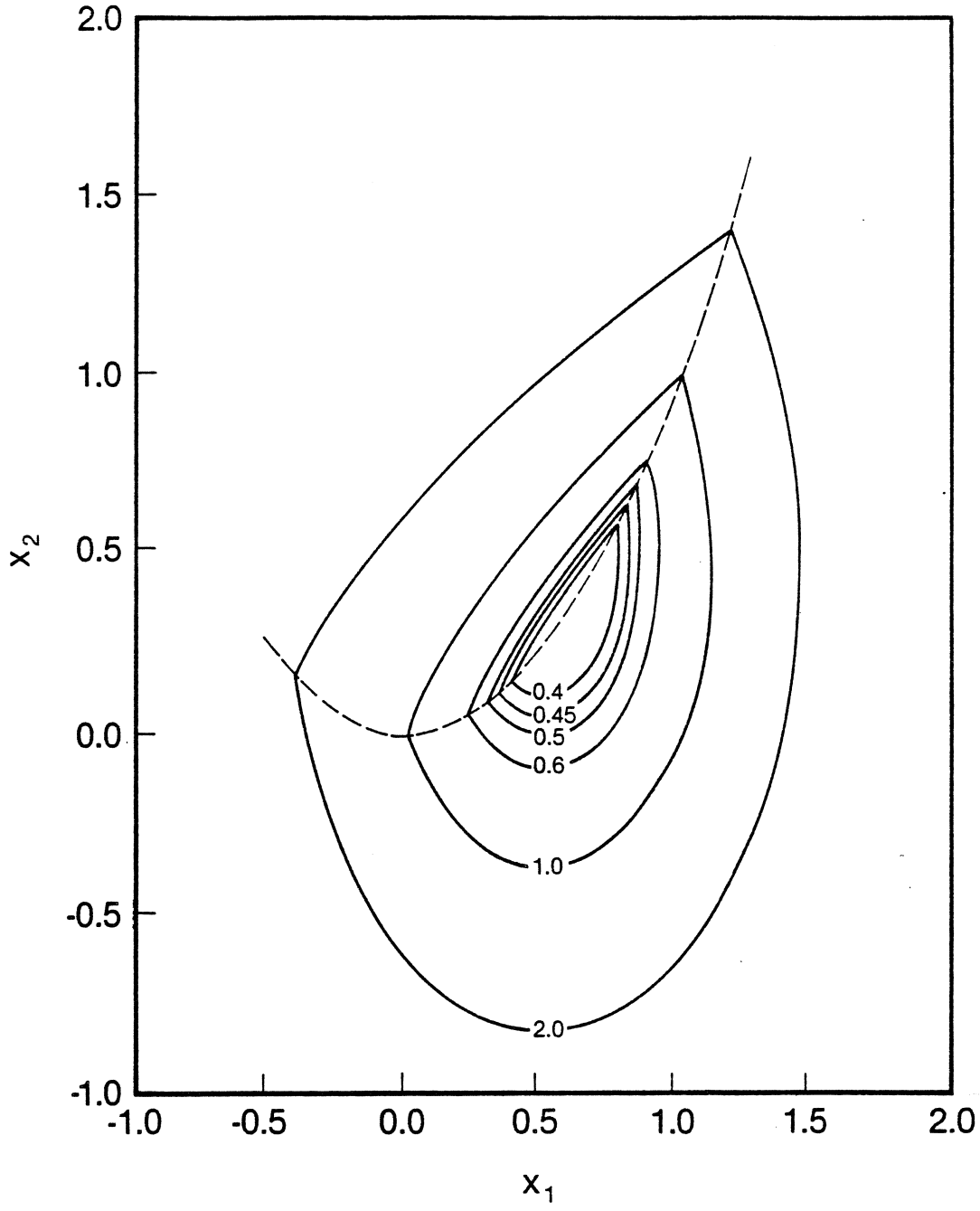


Fig. 4.1 Contours for a singular two-dimensional ℓ_1 problem (4.21).

design problem is formulated and solved using optimization such that the essential demand of production cost reduction is optimally met. The solution of the design problem provides the manufacturer with the allowed design tolerances and the tunable parameters. At the design stage we can distinguish two types of tuning problems:

a) tuning with a fixed set of tunable parameters,

b) tuning with a variable set of tunable parameters.

Problem a) is basically a centering problem (if the tolerances are fixed) and the solution gives us tuning ranges for the tunable parameters. Problem b) involves the selection of tunable parameters at the design stage. Introducing tunable parameters is usually expensive and we want to find the necessary tunable parameters (and their ranges) with the objective of minimizing the cost of introducing tunable parameters. Ideally, we would like to find the minimum number of tunable parameters which are necessary to satisfy the specifications.

In the final stages of production, the manufactured circuit is usually tested to check whether or not it meets the design specifications. Tolerances on circuit components, parasitic effects and uncertainties in the circuit model cause deviations in the manufactured circuit performance, and violation of the design specifications may result (Bandler and Salama 1981). Post-production tuning is usually needed and the tuning assignment problem arises. Here, it is required to find the necessary changes in the tunable parameters to adjust the manufactured circuit to satisfy the design requirements. The post-production tuning problem has been a problem of significant interest amongst computer-aided designers which resulted in a number of algorithms (Adams and Manaktala 1975, Lüder and Kaiser 1976, Lopresti 1977, Alajajian, Trick and El-Masry 1980, Bandler and Biernacki 1980, Schockley and Morris 1983). Most of these algorithms utilize network sensitivities and first-order approximations.

On the contrary, tuning at the design stage has not been given its proper place in computer-aided design. In this chapter, therefore, we examine some aspects of the tuning problem at the design stage.

4.6.2 Tuning at the Design Stage with a Fixed Set of Tunable Parameters

Taking the tuning process into account at the design stage we separate the design components into tunable elements and nontunable elements (Bandler and Biernacki 1980).

Let

$$\boldsymbol{\Phi}_t \triangleq \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \Phi_k \end{bmatrix} \tag{4.22}$$

define the tunable design elements and

$$\boldsymbol{\Phi}_r \triangleq \begin{bmatrix} \Phi_{k+1} \\ \Phi_{k+2} \\ \cdot \\ \cdot \\ \cdot \\ \Phi_n \end{bmatrix} \tag{4.23}$$

define the nontunable elements.

Once the tunable components have been chosen the problem is to find the nominal point $\boldsymbol{\Phi}^0$, tuning ranges of the tunable parameters t_i, $i \in I_t$, and settings of the tunable parameters ρ_i^r, $i \in I_t$, $r \in I_v$, for all worst-cases, such that any worst-case can be tuned into the feasible region of the problem.

The problem is basically a centering problem and can be formulated in minimax form as

$$\underset{\boldsymbol{\phi}^0,\, t,\, \boldsymbol{\rho}^r}{\text{minimize}} \quad F(\boldsymbol{\phi}) \triangleq \max_{i \in I_c} \left(-g_i(\boldsymbol{\phi})\right) \tag{4.24a}$$

subject to

$$\boldsymbol{\phi}^0 \geq \mathbf{0} \tag{4.24b}$$

$$t_i \geq 0\,, \quad i \in I_t \triangleq \{1, 2, \ldots, k\} \tag{4.24c}$$

$$-1 \leq \rho_i^r \leq 1, \quad i \in I_t, \quad r \in I_v\,, \tag{4.24d}$$

where

$$\phi_i = \phi_i^0 + \varepsilon_i \mu_i^r + t_i \rho_i^r, \quad i \in I_t, \; r \in I_v \tag{4.25}$$

or

$$\phi_i = \phi_i^0 + \varepsilon_i \mu_i^r\,, \quad i \notin I_t\,. \tag{4.26}$$

4.6.3 Tuning at the Design Stage with a Variable Set of Tunable Parameters

In this type of problem, the selection of tunable parameters at the design stage is involved. We would like to determine the minimum number of tunable parameters (and their ranges) such that all outcomes can be adjusted to meet the design specifications.

4.6.3.1 Mixed Programming Formulation

The following mixed programming formulation of the tuning problem ensures that the solution gives the minimum number of tunable parameters

$$\underset{t_i,\, \rho_i^r,\, s_i}{\text{minimize}} \quad \sum_{i=1}^{m} w_i s_i \tag{4.27a}$$

subject to

$$g_i(\boldsymbol{\phi}) \geq 0, \quad i \in I_c\,, \tag{4.27b}$$

$$-1 \leq \rho_i^r \leq 1\,, \quad i \in I_\phi, \quad r \in I_v\,, \tag{4.27c}$$

$$w_i,\ t_i \geq 0,\ i \in I_\phi, \tag{4.27d}$$

$$s_i = 0 \text{ or } s_i = 1,\ i \in I_\phi, \tag{4.27e}$$

where

$$\phi_i = \phi_i^0 + \varepsilon_i \mu_i^r + t_i \rho_i^r s_i . \tag{4.28}$$

4.6.3.2 ℓ_1 Objective Function in the Tuning Problem

In many cases, however, we are satisfied with a solution where a possibly small subset of all parameters is selected for tuning (without guarantee that it is the smallest) and in that case we formulate the problem in the way for which there exist efficient algorithms. The necessary conditions for optimality of the nonlinear ℓ_1 problem (Charalambous 1979, Bandler and Kellermann 1983) indicate that zeros of the nonlinear functions $f_j(\mathbf{x})$ play an important role in the characteristics of an ℓ_1 solution. This fact can be used in the tunable parameter selection problem at the design stage.

To select a small subset of tunable parameters at the design stage the following objective function can be used

$$C = \sum_{i=1}^{n} |c_i t_i| , \tag{4.29}$$

where the c_i's are positive weighting factors.

The corresponding optimization problem can be formulated as

$$\underset{\phi_i^0,\, t_i,\, \rho_i^r}{\text{minimize}} \quad \sum_{i=1}^{n} |c_i t_i| \tag{4.30a}$$

subject to

$$g_i(\boldsymbol{\phi}) \geq 0,\ i = 1, 2, \ldots, m_c , \tag{4.30b}$$

$$-1 \leq \rho_i^r \leq 1\ ,\ i = 1, 2, \ldots, n\ , \tag{4.30c}$$

where

$$\Phi_i = \Phi_i^0 + \varepsilon_i \mu_i^r + t_i \rho_i^r\ . \tag{4.31}$$

The ℓ_1 objective function plays an important role in the tuning problem. Many of these problems are formulated as ℓ_1 optimization problems with nonlinear constraints. From the necessary conditions for optimality of the ℓ_1 problem with nonlinear constraints (derived in Section 4.4) it can be seen that the active constraints (more precisely, the gradients of the active constraints) in the ℓ_1 problem have some influence on the number of functions $f_j(\mathbf{x})$ that are zero at a solution $\mathbf{x}^*$. In the tuning problem those zeros correspond to parameters that are not tunable. To illustrate the fact we show a simple two-dimensional example in which the constraints defining the feasible region R_c are linear.

4.6.4 Two-dimensional Example of the Properties of the ℓ_1 Objective Function in the Tuning Problem

Assuming no overlapping of nonfeasible regions defined by different constraints inside the orthotope R_ε, i.e.,

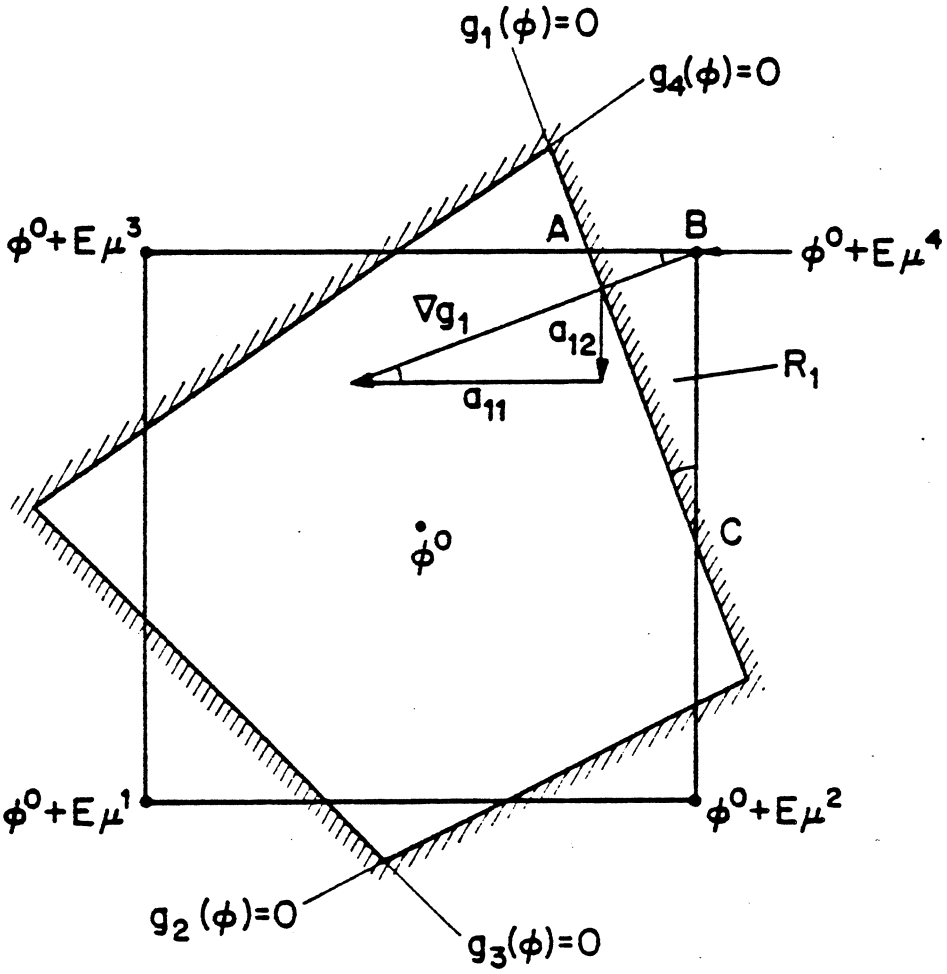
$$\bigcap_{\substack{i \neq j \\ i,j \in I_c}} R_i \cap R_j = \varnothing\ , \tag{4.32}$$

where

$$R_\ell \triangleq \{\Phi \in R_\varepsilon \mid g_\ell(\Phi) < 0\}\ ,\ \ell \in I_c, \tag{4.33}$$

we consider the objective function of the form $\Sigma |t_i|$.

A two-dimensional example of R_c and R_ε, assuming that the constraints defining R_c are linear, is shown in Fig. 4.2. The constraints $g_j(\Phi), j = 1,2,3,4$, are of



$$AB=|t_1\rho_1^4|$$

$$BC=|t_2\rho_2^4|$$

Fig. 4.2 Two-dimensional example of R_ε and R_c.

the form

$$g_j = a_{j1}\Phi_1 + a_{j2}\Phi_2 - b_j \geq 0, \quad j = 1,2,3,4, \tag{4.34}$$

where

$$\Phi_i = \Phi_i^0 + \varepsilon_i \mu_i^r + t_i \rho_i^r .$$

Using simple geometrical relations (see Fig. 4.2), we have

$$\frac{|a_{11}|}{|t_2 \rho_2^4|} = \frac{|a_{12}|}{|t_1 \rho_1^4|} \quad \text{or} \quad \frac{|t_1 \rho_1^4|}{|t_2 \rho_2^4|} = \frac{|a_{12}|}{|a_{11}|} . \tag{4.35}$$

Similarly for other vertices

$$\frac{|t_1 \rho_1^r|}{|t_2 \rho_2^r|} = \frac{|a_{j2}|}{|a_{j1}|} \quad \text{for } (r,j) = \{(2,2)\ (1,3)\ (3,4)\} . \tag{4.36}$$

Vector **t** is a vector of tuning ranges for both parameters. Tuning in the negative direction is allowed with $|\rho_i^r| \leq 1$.

The objective function

$$\sum_{i=1}^{2} |t_i| , \tag{4.37}$$

which minimizes the sum of tuning ranges, will select only one parameter for tuning, namely,

$$t_1 \quad \text{if} \quad \max_j |a_{j1}| > \max_j |a_{j2}| , \quad j = 1,2,3,4 , \tag{4.38}$$

or

$$t_2 \quad \text{if} \quad \max_j |a_{j1}| < \max_j |a_{j2}| , \quad j = 1,2,3,4 . \tag{4.39}$$

Introducing new variables

$$t_i^r = t_i \rho_i^r , \tag{4.40}$$

we can state that the objective (4.37) will select

$$t_1 \quad \text{if} \quad \max_r |t_1^r| < \max_r |t_2^r| , \tag{4.41}$$

or

$$t_2 \quad \text{if} \quad \max_r |t_1^r| > \max_r |t_2^r|. \tag{4.42}$$

In general, the objective function

$$\sum_{i=1}^{n} |t_i| \tag{4.43}$$

minimizes the sum of tuning ranges and is appropriate to minimize the cost of tuning associated with each element.

4.7 FAULT ISOLATION IN ANALOG CIRCUITS USING THE ℓ_1 NORM

4.7.1 Formulation of the Problem

This application of the ℓ_1 norm to circuit problems deals with fault isolation in linear analog circuits under an insufficient number of independent voltage measurements. The ℓ_1 norm is used to isolate the most likely faulty elements. Practically, the faulty components are very few and the relative change in their values is significantly larger than in the nonfaulty ones (Merrill 1973).

The method presented here is a modification of the method utilizing multiple test vectors to obtain the measurements (Bandler, Biernacki, Salama and Starzyk 1982).

For k different excitations applied to the faulty network we consider the following optimization problem

$$\underset{\Phi}{\text{minimize}} \quad \sum_{i=1}^{n} |\Delta\phi_i/\phi_i^0| \tag{4.44a}$$

subject to

$$\mathbf{V}_1^c - \mathbf{V}_1^m = \mathbf{0}\,,$$

$$\vdots \qquad (4.44b)$$

$$\mathbf{V}_k^c - \mathbf{V}_k^m = \mathbf{0}\,,$$

where $\boldsymbol{\Phi} \triangleq [\phi_1\ \phi_2\ \ldots\ \phi_n]^T$ is a vector of network parameters, $\boldsymbol{\Phi}^0$ represents the nominal parameter values, $\Delta\phi_i \triangleq \phi_i - \phi_i^0$, $i = 1, 2, \ldots, n$, represent the deviations in network parameters from nominal values, $\mathbf{V}_k^m$ is a p-dimensional vector of voltage measurements performed at the accessible nodes for the kth excitation and $\mathbf{V}_k^c$ is a p-dimensional vector of voltages at accessible nodes calculated using the vector $\boldsymbol{\Phi}$ as parameter values.

The corresponding nonlinear ℓ_1 problem can be formulated based on an exact penalty function (Charalambous 1979) as follows:

$$\underset{\boldsymbol{\Phi}}{\text{minimize}} \quad \sum_{j=1}^{n+kxp} |f_j(\boldsymbol{\Phi})|\,, \qquad (4.45)$$

where

$$f_i(\boldsymbol{\Phi}) \triangleq \Delta\phi_i/\phi_i^0\,,\ i = 1, 2, \ldots, n\,, \qquad (4.46)$$

$$f_{n+i}(\boldsymbol{\Phi}) \triangleq \beta_i\,(V_i^c - V_i^m)\,,\ i = 1, 2, \ldots, kxp\,, \qquad (4.47)$$

and β_i, $i = 1, 2, \ldots, kxp$, are appropriate multipliers (satisfying certain conditions stated in Charalambous 1979).

4.7.2 Mesh Network Example (Bandler, Biernacki, Salama and Starzyk 1982)

Consider the resistive network shown in Fig. 4.3 with the nominal values of elements $G_i = 1.0$ and tolerances $\varepsilon_i = \pm 0.05$, $i = 1, 2, \ldots, 20$. All outside nodes are

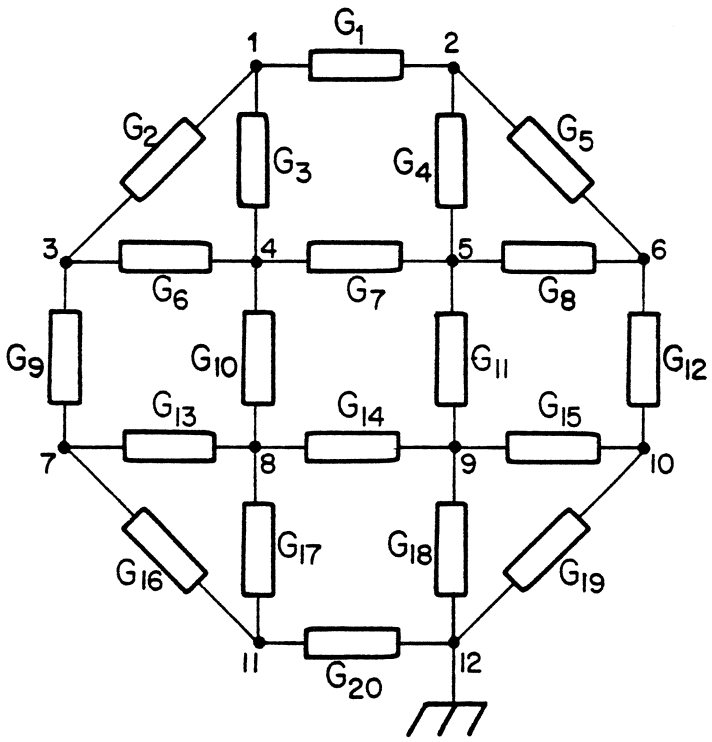


Fig. 4.3 The resistive mesh network (12 nodes).

assumed to be accessible with node 12 taken as the reference node. Nodes 4, 5, 8 and 9 are assumed internal, where no measurements can be performed. Two faults are assumed in the network in elements G_2 and G_{18}. In Case 1 we applied the Hald and Madsen ℓ_1 algorithm to optimization problem (4.45) with a single excitation at node 1. In Case 2 we considered two excitations applied at nodes 3 and 6 sequentially. The results of both optimization problems are summarized in Table 4.1. The nominal component values have been used as a starting point since just a few elements change significantly from nominal.

In both cases the actual faulty elements have been identified, but in Case 2, the estimated changes in the faulty elements are closer to their true values. Also some of the changes in the nonfaulty components approach better their true values in Case 2. The estimated changes in the faulty elements are much closer to the actual changes as compared to the results reported in Bandler, Biernacki, Salama and Starzyk (1982).

Consider the resistive network shown in Fig. 4.4 with 28 nodes, the nominal values of elements $G_i = 1.0$ and tolerances $\varepsilon_i = \pm 0.05$, $i = 1, 2, \ldots, 52$. All outside nodes are assumed to be accessible for measurements with node 28 taken as the reference node.

Two experiments have been performed using this network and a single excitation applied at node 21. First, six faults are assumed in the network in elements G_{37}, G_{38}, G_{39}, G_{49}, G_{51} and G_{52}. The optimization problem (4.45) has been solved and the results are shown in Table 4.2. All six faults have been identified successfully. In another experiment (see results in Table 4.3) four faults in the network in elements G_{41}, G_{44}, G_{45} and G_{48} have also been successfully detected with only one excitation. The results presented in Tables 4.2 and 4.3 have been obtained using a version of the

TABLE 4.1

RESULTS FOR THE MESH NETWORK EXAMPLE (12 NODES, 2 FAULTS)

Element	Nominal Value	Actual Value	Percentage Deviation		
			Actual	Case 1	Case 2
G_1	1.0	0.98	-2.0	0.00	0.13
G_2	1.0	0.50	-50.0*	-48.78	-49.44
G_3	1.0	1.04	4.0	0.00	3.60
G_4	1.0	0.97	-3.0	0.00	0.00
G_5	1.0	0.95	-5.0	-2.26	-1.71
G_6	1.0	0.99	-1.0	0.00	0.00
G_7	1.0	1.02	2.0	0.00	0.00
G_8	1.0	1.05	5.0	0.00	0.00
G_9	1.0	1.02	2.0	2.80	0.97
G_{10}	1.0	0.98	-2.0	0.00	0.00
G_{11}	1.0	1.04	4.0	0.00	0.00
G_{12}	1.0	1.01	1.0	3.45	2.08
G_{13}	1.0	0.99	-1.0	0.00	-0.44
G_{14}	1.0	0.98	-2.0	0.00	0.00
G_{15}	1.0	1.02	2.0	0.00	1.55
G_{16}	1.0	0.96	-4.0	-2.42	-5.71
G_{17}	1.0	1.02	2.0	0.00	2.67
G_{18}	1.0	0.50	-50.0*	-52.16	-48.94
G_{19}	1.0	0.98	-2.0	0.00	-1.95
G_{20}	1.0	0.96	-4.0	-3.67	-4.88
Number of Function Evaluations				8	8
Execution Time (secs) on Cyber 170/815				3.0	3.9

* Faults

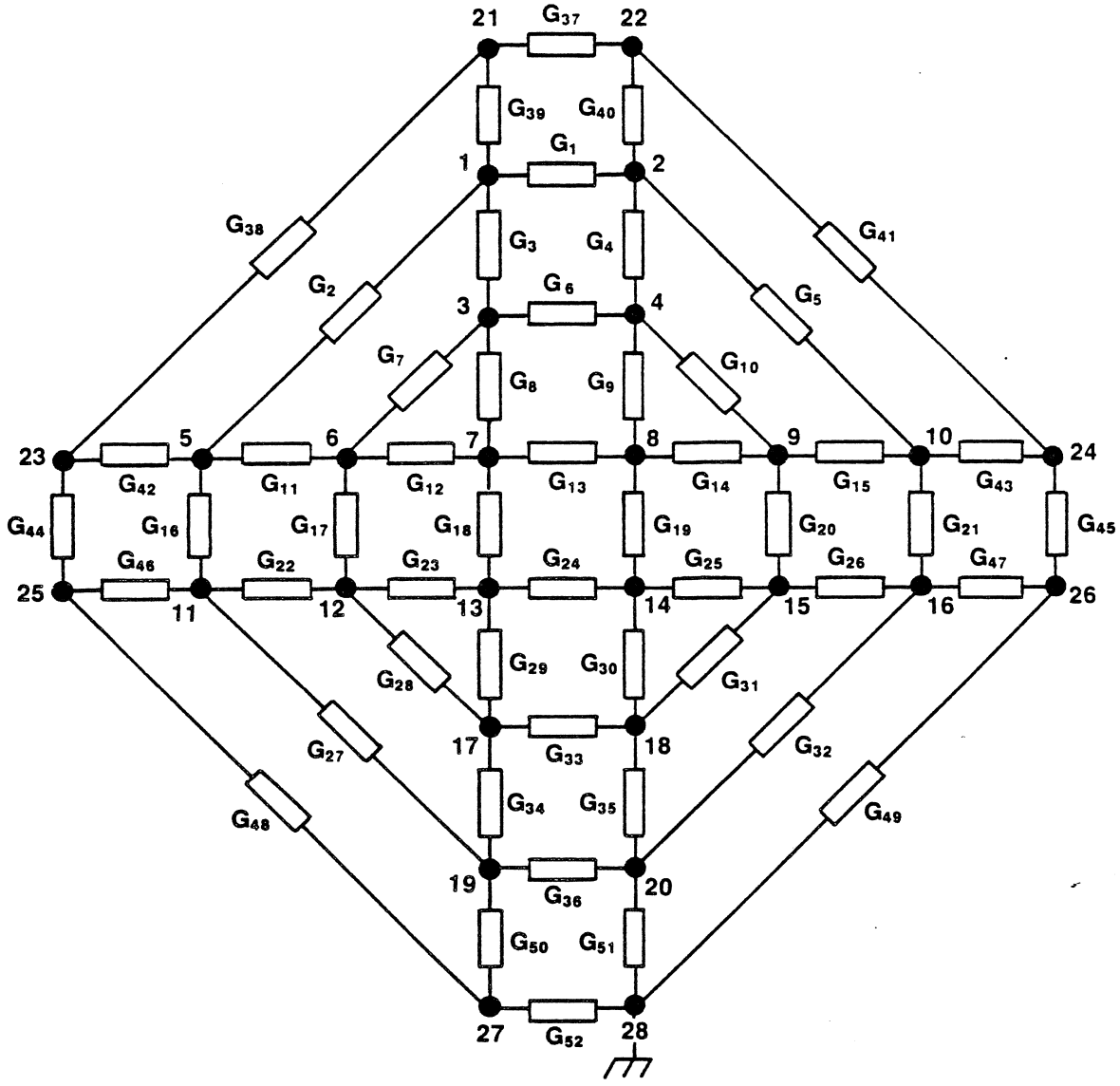


Fig. 4.4 The resistive mesh network (28 nodes).

TABLE 4.2

RESULTS FOR THE MESH NETWORK EXAMPLE (28 NODES, 6 FAULTS)

Element	Nominal Value	Actual Value	Percentage Deviation	
			Actual	Computed
G_1	1.00	0.98	-2.0	0.0
G_2	1.00	0.98	-2.0	0.0
G_3	1.00	1.04	4.0	0.0
G_4	1.00	1.01	1.0	0.0
G_5	1.00	0.99	-1.0	0.0
G_6	1.00	0.97	-3.0	0.0
G_y	1.00	0.95	-5.0	0.0
G_8	1.00	1.02	2.0	0.0
G_9	1.00	1.00	0.0	0.0
G_{10}	1.00	1.01	1.0	0.0
G_{11}	1.00	0.95	-5.0	0.0
G_{12}	1.00	0.99	-1.0	0.0
G_{13}	1.00	1.03	3.0	0.0
G_{14}	1.00	0.97	-3.0	0.0
G_{15}	1.00	0.95	-5.0	0.0
G_{16}	1.00	1.02	2.0	0.0
G_{17}	1.00	1.04	4.0	0.0
G_{18}	1.00	0.98	-2.0	0.0
G_{19}	1.00	0.99	-1.0	0.0
G_{20}	1.00	1.05	5.0	0.0
G_{21}	1.00	0.97	-3.0	0.0
G_{22}	1.00	1.04	4.0	0.0
G_{23}	1.00	0.99	-1.0	0.0
G_{24}	1.00	0.96	-4.0	0.0
G_{25}	1.00	1.05	5.0	0.0
G_{26}	1.00	0.96	-4.0	0.0
G_{27}	1.00	1.04	4.0	0.0
G_{28}	1.00	1.00	0.0	0.0

TABLE 4.2 (continued)

RESULTS FOR THE MESH NETWORK EXAMPLE (28 NODES, 6 FAULTS)

G_{29}	1.00	0.99	-1.0	0.0
G_{30}	1.00	0.95	-5.0	0.0
G_{31}	1.00	1.03	3.0	0.0
G_{32}	1.00	0.99	-1.0	0.0
G_{33}	1.00	1.00	0.0	0.0
G_{34}	1.00	0.98	-2.0	0.0
G_{35}	1.00	0.96	-4.0	0.0
G_{36}	1.00	0.98	-2.0	0.0
G_{37}	1.00	0.50	-50.0*	-49.0
G_{38}	1.00	0.50	-50.0*	-46.0
G_{39}	1.00	0.50	-50.0*	-56.0
G_{40}	1.00	1.01	1.0	0.0
G_{41}	1.00	1.03	3.0	0.0
G_{42}	1.00	0.98	-2.0	0.0
G_{43}	1.00	0.97	-3.0	0.0
G_{44}	1.00	0.95	-5.0	0.0
G_{45}	1.00	0.96	-4.0	-11.0
G_{46}	1.00	1.02	2.0	0.0
G_{47}	1.00	1.04	4.0	0.0
G_{48}	1.00	0.99	-1.0	0.0
G_{49}	1.00	0.50	-50.0*	-55.0
G_{50}	1.00	1.03	3.0	0.0
G_{51}	1.00	0.50	-50.0*	-40.0
G_{52}	1.00	0.50	-50.0*	-52.0
Number of Function Evaluations		12		

*Faults

TABLE 4.3

RESULTS FOR THE MESH NETWORK EXAMPLE (28 NODES, 4 FAULTS)

Element	Nominal Value	Actual Value	Percentage Deviation	
			Actual	Computed
G_1	1.00	0.98	-2.0	0.0
G_2	1.00	0.98	-2.0	0.0
G_3	1.00	1.04	4.0	0.0
G_4	1.00	1.01	1.0	0.0
G_5	1.00	0.99	-1.0	0.0
G_6	1.00	0.97	-3.0	0.0
G_7	1.00	0.95	-5.0	0.0
G_8	1.00	1.02	2.0	0.0
G_9	1.00	1.00	0.0	0.0
G_{10}	1.00	1.01	1.0	0.0
G_{11}	1.00	0.95	-5.0	0.0
G_{12}	1.00	0.99	-1.0	0.0
G_{13}	1.00	1.03	3.0	0.0
G_{14}	1.00	0.97	-3.0	0.0
G_{15}	1.00	0.95	-5.0	0.0
G_{16}	1.00	1.02	2.0	0.0
G_{17}	1.00	1.04	4.0	0.0
G_{18}	1.00	0.98	-2.0	0.0
G_{19}	1.00	0.99	-1.0	0.0
G_{20}	1.00	1.05	5.0	0.0
G_{21}	1.00	0.97	-3.0	0.0
G_{22}	1.00	1.04	4.0	0.0
G_{23}	1.00	0.99	-1.0	0.0
G_{24}	1.00	0.96	-4.0	0.0
G_{25}	1.00	1.05	5.0	0.0
G_{26}	1.00	0.96	-4.0	0.0
G_{27}	1.00	1.04	4.0	0.0

TABLE 4.3 (continued)

RESULTS FOR THE MESH NETWORK EXAMPLE (28 NODES, 4 FAULTS)

G_{28}	1.00	1.00	0.0	0.0
G_{29}	1.00	0.99	–1.0	0.0
G_{30}	1.00	0.95	–5.0	0.0
G_{31}	1.00	1.03	3.0	0.0
G_{32}	1.00	0.99	–1.0	0.0
G_{33}	1.00	1.00	0.0	0.0
G_{34}	1.00	0.98	–2.0	0.0
G_{35}	1.00	0.96	–4.0	0.0
G_{36}	1.00	0.98	–2.0	0.0
G_{37}	1.00	0.99	–1.0	2.0
G_{38}	1.00	1.05	5.0	0.0
G_{39}	1.00	0.97	–3.0	0.0
G_{40}	1.00	1.01	1.0	0.0
G_{41}	1.00	0.50	–50.0*	–46.0
G_{42}	1.00	0.98	–2.0	0.0
G_{43}	1.00	0.97	–3.0	0.0
G_{44}	1.00	0.50	–50.0*	–54.0
G_{45}	1.00	0.50	–50.0*	–45.0
G_{46}	1.00	1.02	2.0	0.0
G_{47}	1.00	1.04	4.0	0.0
G_{48}	1.00	0.50	–50.0*	–53.0
G_{49}	1.00	0.95	–5.0	0.0
G_{50}	1.00	1.03	3.0	0.0
G_{51}	1.00	0.98	–2.0	–6.0
G_{52}	1.00	1.03	3.0	0.0
Number of Function Evaluations		7		

*Faults

Hald and Madsen ℓ_1 algorithm developed on the Texas Instruments Professional Computer. Due to the size of the problem the execution times were of the order of hours on a TI/PC.

4.8 MODEL PARAMETER IDENTIFICATION USING THE ℓ_1 NORM

4.8.1 Formulation of the Problem

This application of the ℓ_1 norm to circuit problems deals with model parameter identification from measurements. The problem of approximating a measured response by a network or system response can be formulated as an optimization problem.

Let

$$\mathbf{F}^m(\omega) \triangleq [F_1^m \;\; F_2^m \;\; \ldots \;\; F_k^m]^T \tag{4.48}$$

be a measured response corresponding to measurements at data (frequency) points ω_i, $i = 1, 2, \ldots, k$, where

$$F_i^m \triangleq F^m(\omega_i), \; i = 1, 2, \ldots, k. \tag{4.49}$$

Let

$$\mathbf{F}^c(\boldsymbol{\phi}, \omega) \triangleq [F_1^c(\boldsymbol{\phi}) \;\; F_2^c(\boldsymbol{\phi}) \;\; \ldots \;\; F_k^c(\boldsymbol{\phi})]^T \tag{4.50}$$

be the response of an appropriate model which depends nonlinearly on a vector of parameters $\boldsymbol{\phi} \triangleq [\phi_1 \;\; \phi_2 \;\; \ldots \;\; \phi_n]^T$, where

$$F_i^c(\boldsymbol{\phi}) \triangleq F^c(\boldsymbol{\phi}, \omega_i), \; i = 1, 2, \ldots, k. \tag{4.51}$$

The identification problem may be stated as follows:

$$\underset{\boldsymbol{\phi}}{\text{minimize}} \; \|\mathbf{f}\|, \tag{4.52}$$

where

$$\mathbf{f} \triangleq [f_1 \;\; f_2 \;\; \ldots \;\; f_k]^T, \tag{4.53}$$

$$f_i(\boldsymbol{\phi}) \triangleq F_i^c(\boldsymbol{\phi}) - F_i^m, \; i = 1, 2, \ldots, k. \tag{4.54}$$

It is usually assumed that the expected values of components of $\mathbf{f}$ are zero, but due to the presence of measurement errors in observing $\mathbf{F}^m$, this cannot be realized in practice. The particular norm to be used depends on the distribution of these errors, represented by the components of $\mathbf{f}$.

It is commonly supposed that the values of the f_i's are independent and normally distributed, when the maximum likelihood estimate of the data is given by choosing the norm to be the least squares norm (Watson 1984). The measurements, however, might contain some isolated large errors, and in this case, minimization of the ℓ_1 norm residual is recommended due to its "filtering" properties w.r.t. large errors.

Using the ℓ_1 norm the identification problem becomes

$$\underset{\boldsymbol{\Phi}}{\text{minimize}} \quad \sum_{i=1}^{k} |f_i(\boldsymbol{\Phi})| \tag{4.55}$$

where $f_i(\boldsymbol{\Phi})$, $i = 1, 2, \ldots, k$, are defined in (4.54).

4.8.2 6th Order Multi-Coupled Cavity Filter Example

In this example we deal with multi-coupled cavity narrow band-pass filters used in microwave communication systems (see Fig. 4.5).

A narrow-band lumped model of an unterminated multi-cavity filter has been given by Atia and Williams (1972) as

$$\mathbf{Z}\,\mathbf{I} = \mathbf{V}\,, \tag{4.56}$$

where

$$\mathbf{Z} = j\,(s\mathbf{1} + \mathbf{M})\,, \tag{4.57}$$

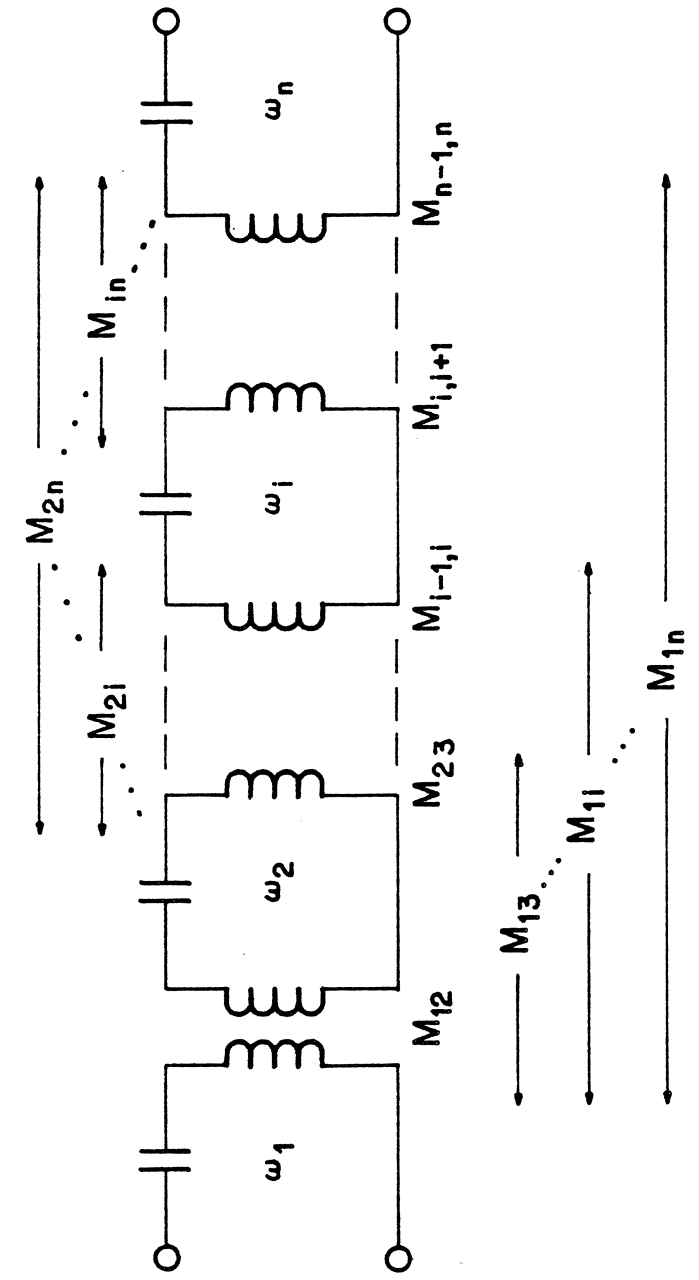


Fig. 4.5 Unterminated coupled-cavity filter illustrating the coupling coefficients.

$$s = \frac{\omega_0}{\Delta\omega}\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right), \tag{4.58}$$

1 denotes an nxn identity matrix, **M** is an nxn coupling matrix whose (i,j) element represents the normalized coupling between the ith and jth cavities, ω_0 is the center frequency and $\Delta\omega$ is the bandwidth parameter. The diagonal entries M_{ii} represent the deviations from the synchronous tuning.

In practice it is often desired to determine the actual filter couplings based on response (return loss or insertion loss) measurements. The problem can be formulated as an optimization problem (4.55) with the ℓ_1 objective function.

In this example reflection coefficient has been used as the filter response. A 6th order filter centered at 4000 MHz with 40 MHz bandwidth is considered. Optimally designed filter parameters have been perturbed and the filter has been simulated. Reflection coefficient at 26 frequency points is used as the specification (measured response). The optimization problem (4.55) has been solved using quite arbitrary couplings as starting values. The results of parameter identification are summarized in Table 4.4.

To demonstrate the properties of the ℓ_1 norm in the identification problem we deliberately introduce large errors to data representing the measurements. Table 4.5 contains data (frequency point, reflection coefficient) used in the previous example and data with two large errors. Measurement 0.23 at 3986.0 MHz has been replaced by 0.75 and measurement 0.14 at 3990.0 by 0.40.

TABLE 4.4

RESULTS OF IDENTIFICATION FOR THE 6TH ORDER FILTER EXAMPLE

Coupling	Optimal Design Value	Perturbed Design Value	Identified
M_{12}	0.819006	0.859956 (+5%)	0.860020
M_{23}	0.511264	0.526602 (+3%)	0.527018
M_{34}	0.824890	0.791894 (–4%)	0.791897
M_{45}	0.511264	0.526602 (+3%)	0.526183
M_{56}	0.819006	0.859956 (+5%)	0.859893
M_{16}	0.093863	0.087293 (–7%)	0.087287
M_{25}	–0.357895	–0.393685 (+10%)	–0.393684
Number of Function Evaluations		12	
ℓ_1 Objective Function at the Solution		4.3155×10^{-5}	
Execution Time (secs) on VAX 11/780		8.4	

TABLE 4.5

DATA USED AS MEASUREMENTS IN THE 6TH ORDER FILTER EXAMPLE

Frequency (MHz)	Reflection Coefficient
3975.0	0.99
3977.0	1.00
3979.0	0.89
3980.0	0.58
3982.0	0.26
3984.0	0.23
3986.0	0.23
3986.0	$0.23 \rightarrow 0.75$
3988.0	0.20
3990.0	$0.14 \rightarrow 0.40$
3992.0	0.08
3994.0	0.01
3996.0	0.05
3998.0	0.08
4000.0	0.09
4002.0	0.08
4004.0	0.05
4006.0	0.01
4008.0	0.08
4010.0	0.14
4012.0	0.20
4014.0	0.23
4016.0	0.23
4018.0	0.25
4020.0	0.55
4022.0	0.99
4024.0	0.99

The optimization problem (4.55) has been solved with the data containing two large errors. The results of identification are summarized in Table 4.6. All the couplings have been identified successfully in the presence of large errors. The ℓ_1 objective function at the solution is equal to the sum of the absolute values of the errors introduced in the measurements.

4.9 CONCLUSIONS

In this chapter we have investigated the ℓ_1 optimization techniques in computer-aided engineering. The area of nonlinear ℓ_1 optimization is reviewed with the emphasis on the Hald and Madsen algorithm. The necessary conditions for optimality of the nonlinear ℓ_1 problem subject to nonlinear constraints indicate that zeros of the nonlinear functions and active constraints play an important role in the characteristics of the ℓ_1 problem. This fact has been used in fault isolation techniques for linear analog circuits and we have demonstrated that the Hald and Madsen ℓ_1 algorithm is very successful in methods for fault isolation in linear circuits under an insufficient number of independent voltage measurements. Singular and regular ℓ_1 problems have been defined and a criterion for determining a singularity present in the ℓ_1 problem has been formulated.

We have also discussed in detail tunable parameter selection in the optimal DCTT problem. A new mixed programming formulation of the problem is given which ensures that the solution gives the minimum number of tunable parameters to tune all outcomes. The ℓ_1 type of objective function in the tuning problem has been examined and illustrated by a two-dimensional example. We have also presented a formulation using the ℓ_1 norm for model parameter identification problems and illustrated it with a 6th order multi-coupled cavity narrow bandpass filter.

TABLE 4.6

RESULTS OF IDENTIFICATION FOR THE 6TH ORDER FILTER PROBLEM WITH DATA CONTAINING LARGE ERRORS

Coupling	Actual Value	Identified Value
M_{12}	0.859956	0.860312
M_{23}	0.526602	0.527016
M_{34}	0.791894	0.791897
M_{45}	0.526602	0.526184
M_{56}	0.859956	0.859608
M_{16}	0.087293	0.087273
M_{25}	-0.393685	-0.393680
Number of Function Evaluations		19
ℓ_1 Objective Function at the Solution		0.77628
Execution Time (secs) on VAX 11/780		12.5

5

A MINIMAX APPROACH TO THE BEST ALIGNMENT OF MECHANICAL SYSTEMS

5.1 INTRODUCTION

An important practical extension of the problems discussed in the previous chapters is the best alignment problem (Bandler, El-Kady, Kellermann and Zuberek 1983b, 1984). Generally, the optimal design problem is to ensure that a design, when manufactured, will satisfy specifications. In many practical situations, however, due to manufacturing errors, a product may not meet the specifications (Petersen and Johnson 1979, 1980, 1982). There are two principal ways of tackling this problem: complete rejection and replacement of the manufactured part, or alignment or reworking (if possible) of the part. In the case of very expensive materials, the latter may be justified. The problem we address in this chapter is how to efficiently perform the part alignment process and, if reworking is needed, how to choose the best way to do it. We provide an attempt to formulate and to solve this problem using minimax optimization (Hald and Madsen 1981, Bandler and Zuberek 1982).

First, basic definitions and concepts are given and the problem is formulated in terms of minimax optimization. Tolerance regions, error functions and their derivatives are described together with examples of tolerance regions. We show the test results obtained by running the program (Bandler, El-Kady, Kellermann and Zuberek 1983a) for several samples (Woodward Governor Company 1982). Conclusions and suggestions for further development are also given.

5.2 PRELIMINARY CONCEPTS

Suppose we have a set of points P in a two-dimensional space

$$P \triangleq \{p_1, p_2, \ldots, p_m\}, \qquad m \geq 1, \tag{5.1}$$

and a system of coordinates $\overline{Y}\overline{O}\overline{X}$ associated with this set. Let

$$I \triangleq \{1, 2, \ldots, m\} \tag{5.2}$$

be the index set for these points.

The coordinates of a point $p_i \in P$, $i \in I$, may be given either w.r.t. the main origin of the YOX system of coordinates or w.r.t. another point of the set P. Let

$$I^0 \triangleq \{1, 2, \ldots, n_0\}, \qquad 1 \leq n_0 \leq m, \tag{5.3}$$

be the index set for points which are referenced to the main origin of the $\overline{Y}\overline{O}\overline{X}$ system of coordinates. With each point $p_i \in P$, $1 \leq i \leq n_0$, we associate a set of indices I^i such that elements of I^i are indices of points referenced to p_i. The set I^i, $1 \leq i \leq n_0$, may be an empty set or a subset of the set I.

Let $1 \leq \ell \leq n_0$. For $1 \leq i < \ell$ we have

$$I^i \triangleq \varnothing, \tag{5.4}$$

which means that no points are reference to $p_i \in P$, $1 \leq i < \ell$. For $\ell \leq i \leq n_0$, we define the following index sets

$$\begin{aligned}
&I^{\ell} \triangleq \{n_0+1, \ldots, n_0+n_{\ell}\},\\
&I^{\ell+1} \triangleq \{n_0+n_{\ell}+1, \ldots, n_0+n_{\ell}+n_{\ell+1}\},\\
&I^{\ell+2} \triangleq \{n_0+n_{\ell}+n_{\ell+1}+1, \ldots, n_0+n_{\ell}+n_{\ell+1}+n_{\ell+2}\},\\
&\vdots\\
&I^{i} \triangleq \{n_0+n_{\ell}+n_{\ell+1}+\ldots+n_{i-1}+1, \ldots, n_0+n_{\ell}+n_{\ell+1}+\ldots+n_i\}, \quad \ell \leq i \leq n_0,\\
&\vdots\\
&I^{n_0} \triangleq \{n_0+n_{n_0-1}+1, \ldots, n_0+n_{n_0}\}.
\end{aligned} \tag{5.5}$$

For each point $p_i \in P$, $i \in I$, we introduce a superscript indicating its reference point. For example, $p_i^0(\overline{x}_i^0, \overline{y}_i^0)$, $1 \leq i \leq n_0$, is the ith point of the set P with coordinates $\overline{x}_i^0$, $\overline{y}_i^0$ referenced to the main origin, $p_i^j(\overline{x}_j^0 + \overline{x}_i^j, \overline{y}_j^0 + \overline{y}_i^j)$, $n_0 < i \leq m$, $\ell \leq j \leq n_0$, is the ith point of the set P with coordinates $\overline{x}_i^j$, $\overline{y}_i^j$ referenced to the p_j^0.

5.2.1 Definitions of Subsets of Points

Three disjoint subsets of points can be distinguished in the set P:

- regular points, P_{reg},
- reference points, P_{ref},
- referenced points, P_{refd}.

For each of these subsets there is an associated index set.

Definition 1

A point $p_i^j \in P$ is a regular point if its coordinates are given w.r.t. the main origin of the $\overline{Y}\overline{O}\overline{X}$ sytem of coordinates and if it is not a reference point for other points. Formally,

$$j = 0\,,\ i \in I_{reg} \triangleq \{1\,, 2\,, \ldots\,, \ell - 1\} \ \Rightarrow\ p_i^j \in P_{reg}. \tag{5.6}$$

Definition 2

A point $p_i^j \in P$ is a reference point if its coordinates are given w.r.t. the main origin of the $\overline{Y}\overline{O}\overline{X}$ system of coordinates and if it is treated as an origin for other points. Formally,

$$j = 0\,,\ i \in I_{ref} \triangleq \{\ell\,, \ell+1\,, \ldots\,, n_0\} \ \Rightarrow\ p_i^j \in P_{ref}. \tag{5.7}$$

Definition 3

A point $p_i \in P$ is a referenced point if its coordinates are given w.r.t. another point of the subset P_{ref} and if it is not a reference point for other points. Formally,

$$j \in I_{ref}\,,\ i \in I_{refd} \triangleq \{n_0+1\,,\ldots,\,m\} \ \Rightarrow\ p_i^j \in P_{refd}. \tag{5.8}$$

The concepts and definitions introduced are illustrated in Fig. 5.1.

5.2.2 Example

Let $P \triangleq \{p_1,p_2,p_3,p_4,p_5,p_6,p_7\}$. From Fig. 5.1 we can define the following index sets: $I = \{1,2,3,4,5,6,7\}$, the index set for the set P; $I^0 = \{1,2,3\}$, the index set for points referenced to the main origin of the $\overline{Y}\overline{O}\overline{X}$ system of coordinates; $I^1 = \varnothing$, the index set for points referenced to $p_1{}^0$; $I^2 = \{4,5\}$, the index set for points referenced to $p_2{}^0$; $I^3 = \{6,7\}$, the index set for points referenced to $p_3{}^0$. We can also define the index sets for regular points $I_{reg} = \{1\}$, reference points $I_{ref} = \{2,3\}$ and referenced points $I_{refd} = \{4,5,6,7\}$.

5.2.3 Tolerance Regions

Suppose we have a set R of tolerance regions R_i, $i \in I \triangleq \{1,2,\ldots,m\}$, in the 2-dimensional space,

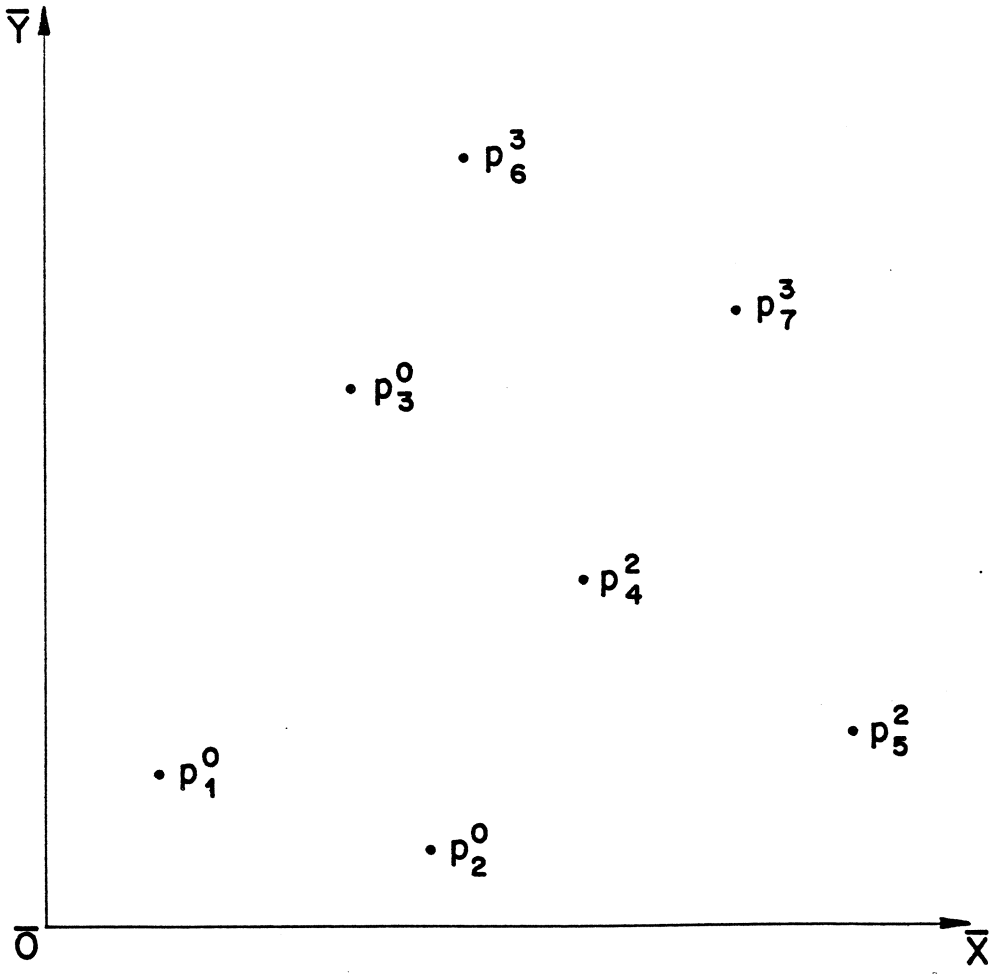
$$R \triangleq \{R_1,R_2,\ldots,R_m\} \tag{5.9}$$

and a system of coordinates YOX associated with this set. We can define a one- to-one mapping g which assigns elements $R_i \in R$ to elements $p_i{}^j \in P$,

$$\{g:P \rightarrow R\}\,. \tag{5.10}$$

The sets P, R and the mapping g are shown in Fig. 5.2.

The regions $R_i \in R$, $i \in I$, may have different shapes (e.g., circular, rectangular), they may be defined using polar coordinates, rectangular coordinates or



$p_1^0(\bar{x}_1^0, \bar{y}_1^0)$ $\quad$ $p_4^2(\bar{x}_4^2 + \bar{x}_2^0, \bar{y}_4^2 + \bar{y}_2^0)$

$p_2^0(\bar{x}_2^0, \bar{y}_2^0)$ $\quad$ $p_5^2(\bar{x}_5^2 + \bar{x}_2^0, \bar{y}_5^2 + \bar{y}_2^0)$

$p_3^0(\bar{x}_3^0, \bar{y}_3^0)$ $\quad$ $p_6^3(\bar{x}_6^3 + \bar{x}_3^0, \bar{y}_6^3 + \bar{y}_3^0)$

$p_7^3(\bar{x}_7^3 + \bar{x}_3^0, \bar{y}_7^3 + \bar{y}_3^0)$

Fig. 5.1 The set of points P and the $\overline{YOX}$ system of coordinates associated with it.

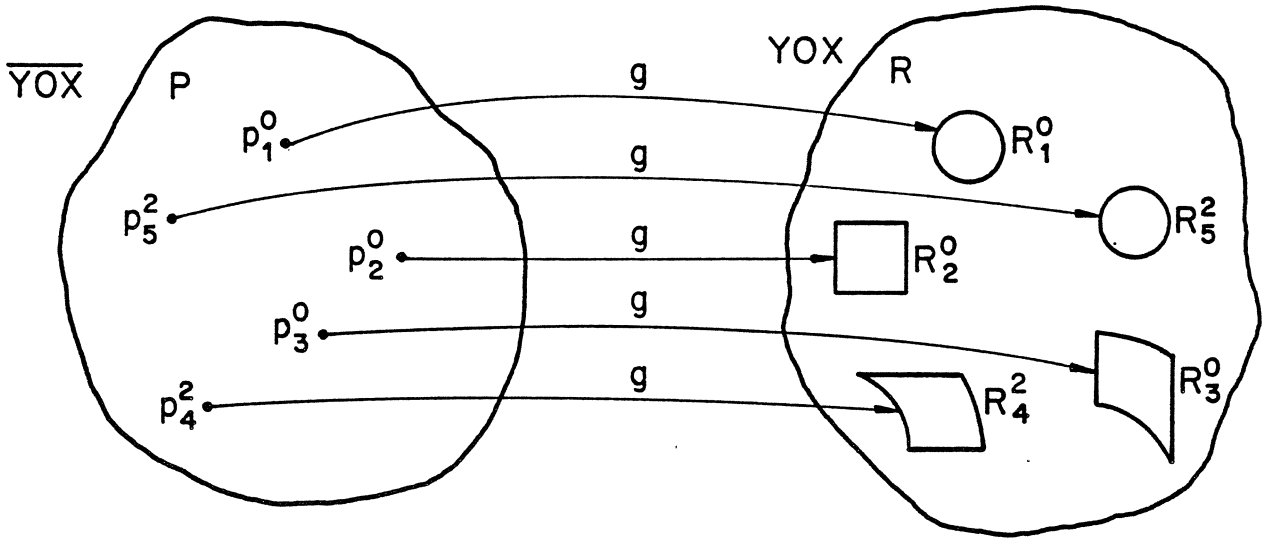


Fig. 5.2 The mapping g: P → R.

combined polar and rectangular coordinates. Dimensions of tolerance regions may be given either w.r.t. the main origin of the YOX system of coordinates (for $R_i = g(p_i)$, $i \in I^0 = I_{reg} \cup I_{ref}$) or w.r.t. the reference point (for $R_i = g(p_i)$, $i \in I_{refd}$).

We can use the same notation indicating the reference points for tolerance regions as for points, e.g., R_i^0, $1 \leq i \leq n_0$, is the ith tolerance region of the set R with dimensions given w.r.t. the main origin of the YOX system of coordinates and R_i^j, $n_0 < i \leq m$, $\ell \leq j \leq n_0$, is the ith tolerance region of the set R with dimensions given w.r.t. the transformed coordinates of p_j^0 from the $\bar{Y}\bar{O}\bar{X}$ to the YOX system of coordinates.

5.2.4 Transformation of Coordinates

The two systems of coordinates, YOX and $\bar{Y}\bar{O}\bar{X}$ are related by the following transformation of coordinates

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \cos\phi_3 & -\sin\phi_3 \\ \sin\phi_3 & \cos\phi_3 \end{bmatrix} \begin{bmatrix} \bar{x}_i \\ \bar{y}_i \end{bmatrix} + \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}, \tag{5.11}$$

where

$$\boldsymbol{\phi}_0 \triangleq [\phi_1 \ \phi_2 \ \phi_3]^T, \tag{5.12}$$

is a set of variables relating the two systems of coordinates (Fig. 5.3).

5.3 FORMULATION OF THE PROBLEM

The first step in the solution of the best alignment problem is to find $\boldsymbol{\phi}_0$ such that the maximum number of points $p_i^j \in P$, $i \in I$, $j \in I_{ref}$ or $j = 0$, are inside or on the boundary of the corresponding $R_i^j \in R$, $R_i^j = g(p_i^j)$. However, the solution to the problem stated above may not be unique and may not be equal to the number of points m.

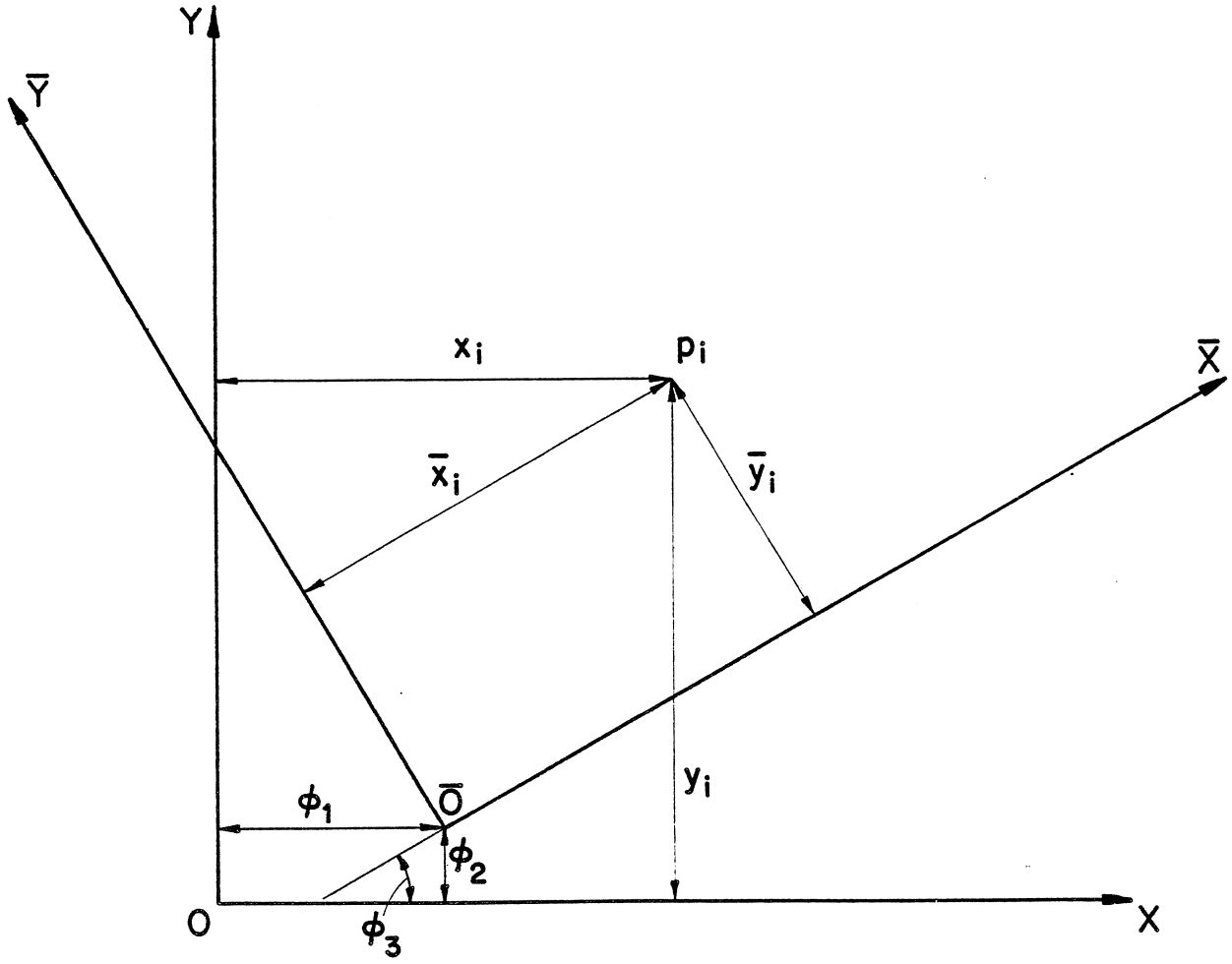


Fig. 5.3 Transformation of coordinates relating the two systems of coordinates.

If it is not possible to find $\boldsymbol{\phi}_0 = [\phi_1\ \phi_2\ \phi_3]^T$ such that all m points are inside or on the boundary of the corresponding tolerance region then it is necessary to delete one or more points in the set P to ensure that all other points satisfy this condition.

In general, the number of variables for the best alignment problem depends on the type of the point (regular, reference or referenced) being a candidate for deletion. The vector of variables $\boldsymbol{\phi}$ may be extended with new variables, which are the coordinates of reference points if these are the candidates for deleting.

Introducing new variables is necessary when deleting a reference point, because we have to determine the locations of all tolerance regions referenced to it. The general form of the vector of variables for the best alignment problem is

$$\boldsymbol{\phi}_k^T = [\phi_1\ \phi_2\ \phi_3\ x_{i_1}\ y_{i_1}\ x_{i_2}\ y_{i_2} \dots\ x_{i_k}\ y_{i_k}], \quad i_1, i_2, \dots, i_k \in I_{delref}, \tag{5.13}$$

where

$$I_{delref} \triangleq I_{ref} \cap I_{del}, \tag{5.14}$$

and k is the cardinality of I_{delref}. The index set I_{del} represents deleted points. For example, if the ith point of the set P ($i \in I_{ref}$) is a candidate for deleting then $\boldsymbol{\phi}_k$ reduces to the form $\boldsymbol{\phi}_1^T = [\phi_1\ \phi_2\ \phi_3\ x_i\ y_i]$, $i \in I_{delref}$. If the ith and jth points ($i,j \in I_{ref}$) are candidates for deleting then $\boldsymbol{\phi}_2^T = [\phi_1\ \phi_2\ \phi_3\ x_i\ y_i\ x_j\ y_j]$, $i,j \in I_{delref}$. If the candidates for deleting are not reference points then $\boldsymbol{\phi}_k = \boldsymbol{\phi}_0$.

The best alignment problem can be formulated as

$$\begin{aligned} &\text{minimize} \quad n_{del} \triangleq \text{card}(I_{del}) \\ &I_{del} \in 2^I \end{aligned}, \tag{5.15}$$

subject to the constraint

$$\min_{\boldsymbol{\phi}_k} \max_i f_i(\boldsymbol{\phi}_k) \le 0, \quad i \in I' \triangleq (I - I_{del}) \cup I_{delref}, \tag{5.16}$$

where I is the index set for points p_i which are to be aligned, I_{del} is the index set for points which should be deleted, 2^I is the family of all subsets of the set I, n_{del} is the

cardinality of I_{del} and Φ_k is the vector of optimization variables corresponding to the set I_{del}. Variables ϕ_1, ϕ_2 and ϕ_3 relate the $\overline{Y}\overline{O}\overline{X}$, and the YOX systems of coordinates and x_{i_k}, y_{i_k} are transformed coordinates (from $\overline{Y}\overline{O}\overline{X}$ to YOX) of a reference point actually being deleted. The error function $f_i(\Phi_k)$ is associated with the point p_i^j to indicate whether the point p_i^j is in ($f_i(\Phi_k) \leq 0$) or out ($f_i(\Phi_k) > 0$) of the tolerance region $R_i^j = g(p_i^j)$.

The index set I′ in (5.16) also contains the indices of deleted reference points in order to ensure that the error function corresponding to the new location of a reference point determined by the optimization (coordinates of the reference point considered are additional variables of the problem) will satisfy the constraint (5.16). This is not required for the deleted regular or referenced points since no other points and tolerance regions are referenced to them.

5.4 ALGORITHM FOR SOLVING THE PROBLEM

The solution to the best alignment problem consists of two stages. The first stage corresponds to a discrete (or combinatorial) minimization of the number of points which should be deleted from the original set of points, and the second stage is an unconstrained minimax optimization of a set of error functions f_i, $i \in I'$, determined by the first stage. The discrete minimization of the first stage is usually implemented as a systematic search of the solution in the family 2^I of all subsets of the set I. It is convenient to represent this search in the form of a multilevel tree in which the root (level 0) corresponds to the set $I_{del} = \varnothing$ ($\varnothing$ denotes the empty set), the level 1 contains all the single element subsets $I^1_{del,1} = \{1\}$, $I^1_{del,2} = \{2\}$, ..., the level 2 all the subsets of I which contain two elements, and so on. The first stage minimization traverses the tree level after level until the solution is found, i.e., until such a subset I_{del} is

encountered for which the constraint (5.16) is satisfied. It can be observed, however, that the minimax optimization of the second stage, which is performed for each step of the first stage search, can be used to eliminate those nodes (and their subtrees) of the search tree which cannot influence the solution. In fact, if the minimax constraint corresponding to the subset I_{del} at a particular level of the search tree is not satisfied then the next level subsets should be derived from the I_{del} of the previous level by adding only the indices of those points which correspond to the active error functions at the solution $\boldsymbol{\phi}_k^*$ of the minimax optimization since the remaining, nonactive error functions do not affect the solution. This observation is the basis of the implemented combinatorial search algorithm which dynamically creates and traverses the reduced search tree.

5.4.1 Details of the Algorithm

The algorithm always starts with the set $I_{del} = \varnothing$ (the root of the tree) and $\boldsymbol{\phi}_k = \boldsymbol{\phi}_0 = \mathbf{0}$. If the minimax objective function

$$F(\boldsymbol{\phi}_0) = \max_{i \in I} f_i(\boldsymbol{\phi}_0) \ , \tag{5.17}$$

at the solution $\boldsymbol{\phi}_0^*$ is nonpositive, $F(\boldsymbol{\phi}_0^*) \leq 0$, then $\boldsymbol{\phi}_0^*$ corresponds to the best alignment solution, and the solution is optimally centered. If $F(\boldsymbol{\phi}_0^*) > 0$, there is no possible alignment of all the points p_i, $i \in I$, and at least one of the points has to be deleted to allow the alignment of the remaining points. The candidates for deletion are the points for which the corresponding error functions are active at the solution $\boldsymbol{\phi}_k^*$, and their indices are attached to the root I_0 of the search tree, creating the level 1 nodes. The search is continued node after node of the created level and the minimax optimization with one less function (except the case of deleting a reference point) is repeated at each node. During the traversal of the level 1 nodes, the new nodes are

attached to the search tree creating the next level, and so on, until a subset I_{del} is found for which the minimax constraint is satisfied, $F(\Phi_k^*) \leq 0$. It should be noted that corresponding to each node of the search tree there is a unique associated index, and the set I_{del} corresponding to the node j is determined as the set of indices of the path from the node j to the root of the tree.

5.5 TOLERANCE REGIONS, ERROR FUNCTIONS AND THEIR DERIVATIVES

To form the error functions for the best alignment problem we have to decide in which system of coordinates these functions will be expressed. It is convenient to choose the system of coordinates associated with the regions, first of all because it is easier to transform points than tolerance regions to the new system of coordinates, and second, because the derivatives of the error functions w.r.t. optimization variables can be easily obtained using transformed coordinates of points and the Jacobian of the transformation.

5.5.1 Preliminary Considerations of Derivatives

For $\Phi_k = \Phi_0$ (no deletions of any points or deletions only of regular or referenced points) the error function is of the form

$$f_i(\Phi_0) = f_i(x_i(\Phi_0), y_i(\Phi_0)), \quad i \in I - I_{del}. \tag{5.18}$$

The derivative of f_i w.r.t. Φ_0 can be written as

$$\frac{\partial f_i}{\partial \Phi_0} = \frac{\partial f_i}{\partial x_i}\frac{\partial x_i}{\partial \Phi_0} + \frac{\partial f_i}{\partial y_i}\frac{\partial y_i}{\partial \Phi_0}, \tag{5.19}$$

where

$$\frac{\partial f_i}{\partial \boldsymbol{\Phi}_0} \triangleq \begin{bmatrix} \frac{\partial f_i}{\partial \Phi_1} \\ \frac{\partial f_i}{\partial \Phi_2} \\ \frac{\partial f_i}{\partial \Phi_3} \end{bmatrix}, \quad \frac{\partial x_i}{\partial \boldsymbol{\Phi}_0} \triangleq \begin{bmatrix} \frac{\partial x_i}{\partial \Phi_1} \\ \frac{\partial x_i}{\partial \Phi_2} \\ \frac{\partial x_i}{\partial \Phi_3} \end{bmatrix}, \quad \frac{\partial y_i}{\partial \boldsymbol{\Phi}_0} \triangleq \begin{bmatrix} \frac{\partial y_i}{\partial \Phi_1} \\ \frac{\partial y_i}{\partial \Phi_2} \\ \frac{\partial y_i}{\partial \Phi_3} \end{bmatrix}. \tag{5.20}$$

The terms of the form $(\partial f_i)/(\partial x_i)$, $(\partial f_i)/(\partial y_i)$ depend on the shape of the tolerance region and usually are not very complicated because the function f_i and the coordinates x_i, y_i are expressed in the same system. The terms of the form $(\partial x_i)/(\partial \boldsymbol{\Phi}_0)$, $(\partial y_i)/(\partial \boldsymbol{\Phi}_0)$ depend only on the transformation formula and are the same for the derivatives of all minimax functions. They can be calculated once for the actual point $\boldsymbol{\Phi}_0$ and used for all functions.

Partial derivatives $(\partial x_i)/(\partial \boldsymbol{\Phi}_0)$ and $(\partial y_i)/(\partial \boldsymbol{\Phi}_0)$ can be arranged in a matrix called the Jacobian of the transformation

$$\mathbf{J}_0^i \triangleq \begin{bmatrix} \frac{\partial x_i}{\partial \Phi_1} & \frac{\partial x_i}{\partial \Phi_2} & \frac{\partial x_i}{\partial \Phi_3} \\ \frac{\partial y_i}{\partial \Phi_1} & \frac{\partial y_i}{\partial \Phi_2} & \frac{\partial y_i}{\partial \Phi_3} \end{bmatrix}, \tag{5.21}$$

which, for the transformation (5.11) takes the form

$$\mathbf{J}_0^i = \begin{bmatrix} 1 & 0 & (-\bar{x}_i \sin\Phi_3 - \bar{y}_i \cos\Phi_3) \\ 0 & 1 & (\bar{x}_i \cos\Phi_3 - \bar{y}_i \sin\Phi_3) \end{bmatrix}. \tag{5.22}$$

5.5.2 General Formulation of Derivatives

For $\boldsymbol{\Phi}_k \neq \boldsymbol{\Phi}_0$ (deletion of reference points) depending on the type of a point for which we form the error function we have three cases:

$i \in I - I_{delref} - I^{i_1} - I^{i_2} - \ldots - I^{i_k}$ (regular point or reference point not deleted)

The error function is of the form

$$f_i(\Phi_k) = f_i(x_i(\Phi_0)), y_i(\Phi_0)). \tag{5.23}$$

The derivatives w.r.t. optimization variables Φ_0 are given by (5.19) and the derivatives w.r.t. additional variables are

$$\begin{aligned}
&\frac{\partial f_i}{\partial x_{i_1}} = 0, \\
&\quad \vdots \\
&\frac{\partial f_i}{\partial x_{i_k}} = 0, \\
&\frac{\partial f_i}{\partial y_{i_1}} = 0, \\
&\quad \vdots \\
&\frac{\partial f_i}{\partial y_{i_k}} = 0,
\end{aligned} \tag{5.24}$$

where $i_1,\ldots,i_k \in I_{delref}$.

$i \in I_{delref}$ (reference point deleted)

The error function is of the form

$$f_i(\Phi_k) = f_i(x_{i_1}, y_{i_1}, \ldots, x_{i_j}, y_{i_j}, \ldots, x_{i_k}, y_{i_k}) \tag{5.25}$$

and the derivatives are

$$\frac{\partial f_i}{\partial \Phi_0} = 0\,,$$

$$\frac{\partial f_i}{\partial x_{i_j}} = \begin{cases} (\ldots) & \text{for } i = i_j \\ 0 & \text{for } i \neq i_j \end{cases} \qquad (5.26)$$

$$i_j \in I_{\text{delref}}\,,\ j = 1, \ldots, k\,.$$

$$\frac{\partial f_i}{\partial y_{i_j}} = \begin{cases} (\ldots) & \text{for } i = i_j \\ 0 & \text{for } i \neq i_j \end{cases}$$

$i \in I^{i_1} \cup I^{i_2} \cup \ldots \cup I^{i_j} \cup \ldots \cup I^{i_k}$ (referenced point)

The error function is of the form

$$f_i(\Phi_k) = f_i(x_i(\Phi_0),\ y_i(\Phi_0),\ x_{i_1}, y_{i_1}, \ldots, x_{i_j}, y_{i_j}, \ldots, x_{i_k}, y_{i_k}) \qquad (5.27)$$

and the derivatives are given by (5.19) for Φ_0 and w.r.t. additional variables by

$$\frac{\partial f_i}{\partial x_{i_j}} = \begin{cases} (\ldots) & \text{for } i \in I^{i_j}, \\ 0 & \text{for } i \notin I^{i_j}, \end{cases}$$

(5.28)

$$\frac{\partial f_i}{\partial y_{i_j}} = \begin{cases} (\ldots) & \text{for } i \in I^{i_j}, \\ 0 & \text{for } i \notin I^{i_j}. \end{cases}$$

5.5.3 Tables of Error Functions and Derivatives

For the general form of the vector of variables, given by (5.13), we form error functions and derivatives for three cases:

1) regular point or reference point not deleted;

2) reference point deleted;

3) referenced point.

The general form of derivatives of f_i w.r.t. $\boldsymbol{\Phi}_0$ is given by (5.19), where the terms of the form $(\partial x_i)/(\partial \boldsymbol{\Phi}_0)$, $(\partial y_i)/(\partial \boldsymbol{\Phi}_0)$ may be calculated as in (5.22), and the terms of the form $(\partial f_i)/(\partial x_i)$, $(\partial f_i)(\partial y_i)$ are tabulated in Tables 5.1 - 5.3 for each type of tolerance region. The coordinates x_i, y_i are transformed coordinates of points using the transformation (5.11). The derivatives of error functions w.r.t. additional variables are also given in these tables.

For the circular tolerance region error functions and derivatives are given in Table 5.1. As an example, consider three points with circular tolerance regions shown in Fig. 5.4. Assume that x_2^0, y_2^0 are additional variables, so

$$\boldsymbol{\Phi}_k = \boldsymbol{\Phi}_1 = [\Phi_1 \quad \Phi_2 \quad \Phi_3 \quad x_2^0 \quad y_2^0]^T.$$

The error functions and derivatives for p_1^0 (regular point), p_2^0 (reference point deleted), and p_3^2 (referenced point) can be calculated using formulas given in Table 5.1.

For other types of tolerance region the location of a point w.r.t. corresponding tolerance region can be characterized by a system of four linear or nonlinear functions. For a regular point and rectangular tolerance region (Fig. 5.5), these functions result from the inequalities

$$x_{iL}^0 \le x_i^0 \le x_{iU}^0, \tag{5.29}$$

$$y_{iL}^0 \le y_i^0 \le y_{iU}^0, \tag{5.30}$$

and have the form

$$f_i^1 = x_{iL}^0 - x_i^0, \tag{5.31}$$

$$f_i^2 = x_i^0 - x_{iU}^0, \tag{5.32}$$

$$f_i^3 = y_{iL}^0 - y_i^0, \tag{5.33}$$

$$f_i^4 = y_i^0 - y_{iU}^0. \tag{5.34}$$

TABLE 5.1

DERIVATIVES OF ERROR FUNCTIONS FOR CIRCULAR TOLERANCE REGION

	Regular Point or Reference Point Not Deleted	Reference Point Deleted	Referenced Point
	$f_i = D - r_i$	$f_i = D - r_i$	$f_i = D_1 - r_i$
$\partial f_i/\partial x_i$	$-A/D$	*	$-A_1/D_1$
$\partial f_i/\partial y_i$	$-B/D$	*	$-B_1/D_1$
$\partial f_i/\partial x_{i_j}$	0	$-A/D,\ i = i_j$ $0,\quad i \neq i_j$	$A_1/D_1,\ i \in I^{i_j}$ $0,\quad i \notin I^{i_j}$
$\partial f_i/\partial y_{i_j}$	0	$-B/D,\ i = i_j$ $0,\quad i \neq i_j$	$B_1/D_1,\ i \in I^{i_j}$ $0,\quad i \notin I^{i_j}$

$$A = x^0_{n_i} - x^0_i \qquad A_1 = x^{i_j}_{n_i} + x^0_{i_j} - x_i$$

$$B = y^0_{n_i} - y^0_i \qquad B_1 = y^{i_j}_{n_i} + y^0_{i_j} - y_i$$

$$D = (A^2 + B^2)^{1/2} \qquad D_1 = (A_1^2 + B_1^2)^{1/2}$$

* For this case, terms $\dfrac{\partial x_i}{\partial \Phi_0}$, $\dfrac{\partial y_i}{\partial \Phi_0}$ are equal to zero, and consequently $\dfrac{\partial f_i}{\partial \Phi_0} = 0$.

TABLE 5.2

DERIVATIVES OF ERROR FUNCTIONS FOR RECTANGULAR TOLERANCE REGION

Regular Point or Reference Point not Deleted

f_i^s	$f_i^1 = x_{iL}^0 - x_i^0$	$f_i^2 = x_i^0 - x_{iU}^0$	$f_i^3 = y_{iL}^0 - y_i^0$	$f_i^4 = y_i^0 - y_{iU}^0$
$\partial f_i^s / \partial x_i^0$	-1	1	0	0
$\partial f_i^s / \partial y_i^0$	0	0	-1	1
$\partial f_i^s / \partial x_{i_j}$	0	0	0	0
$\partial f_i^s / \partial y_{i_j}$	0	0	0	0

Reference Point Deleted

f_i^s	$f_i^1 = x_{iL}^0 - x_i^0$	$f_i^2 = x_i^0 - x_{iU}^0$	$f_i^3 = y_{iL}^0 - y_i^0$	$f_i^4 = y_i^0 - y_{iU}^0$
$\partial f_i^s / \partial x_{i_j}$	$-1,\ i=i_j$ $0,\ i\neq i_j$	$1,\ i=i_j$ $0,\ i\neq i_j$	0	0
$\partial f_i^s / \partial y_{i_j}$	0	0	$-1,\ i=i_j$ $0,\ i\neq i_j$	$1,\ i=i_j$ $0,\ i\neq i_j$

Referenced Point

f_i^s	$f_i^1 = x_{i_j}^0 + x_{iL}^{i_j} - x_i$	$f_i^2 = x_i - x_{iU}^{i_j} + x_{i_j}^0$	$f_i^3 = y_{iL}^{i_j} + y_i^0 - y_i$	$f_i^4 = y_i^0 - (y_{iU}^{i_j} + y_{i_j}^0)$
$\partial f_i^s / \partial x_i$	-1	1	0	0
$\partial f_i^s / \partial y_i$	0	0	-1	1
$\partial f_i^s / \partial x_{i_j}$	$1,\ i \in I^{i_j}$ $0,\ i \notin I^{i_j}$	$-1,\ i \in I^{i_j}$ $0,\ i \notin I^{i_j}$	0	0
$\partial f_i^s / \partial y_{i_j}$	0	0	$1,\ i \in I^{i_j}$ $0,\ i \notin I^{i_j}$	$-1,\ i \in I^{i_j}$ $0,\ i \notin I^{i_j}$

TABLE 5.3

DERIVATIVES OF ERROR FUNCTIONS FOR X-R TOLERANCE REGION

Regular Point or Reference Point not Deleted

f_i^s	$f_i^1 = x_{iL}^0 - x_i^0$	$f_i^2 = x_i^0 - x_{iU}^0$	$f_i^3 = R_{iL}^0 - E$	$f_i^4 = E - R_{iU}$
$\partial f_i^s/\partial x_i^0$	-1	1	$-x_i^0/E$	x_i^0/E
$\partial f_i^s/\partial y_i^0$	0	0	$-y_i^0/E$	y_i^0/E
$\partial f_i^s/\partial x_{i_j}$	0	0	0	0
$\partial f_i^s/\partial y_{i_j}$	0	0	0	0

Reference Point Deleted

f_i^s	$f_i^1 = x_{iL}^0 - x_i^0$	$f_i^2 = x_i^0 - x_{iU}^0$	$f_i^3 = R_{iL} - E$	$f_i^4 = E - R_{iU}$
$\partial f_i^s/\partial x_{i_j}$	$-1,\ i=i_j$ $0,\ i\neq i_j$	$1,\ i=i_j$ $0,\ i\neq i_j$	$-x_i^0/E,\ i=i_j$ $0,\quad i\neq i_j$	$x_i^0/E,\ i=i_j$ $0,\quad i\neq i_j$
$\partial f_i^s/\partial y_{i_j}$	0	0	$-y_i^0/E,\ i=i_j$ $0,\quad i\neq i_j$	$y_i^0/E,\ i=i_j$ $0,\quad i\neq i_j$

Referenced Point

f_i^s	$f_i^1 = x_{i_j}^0 + x_{iL}^{i_j} - x_i$	$f_i^2 = x_i - (x_{iU}^{i_j} + x_{i_j}^0)$	$f_i^3 = R_{iL} - D$	$f_i^4 = D - R_{iU}$
$\partial f_i^s/\partial x_i$	-1	1	$-A/D$	A/D
$\partial f_i^s/\partial y_i$	0	0	$-B/D$	B/D
$\partial f_i^s/\partial x_{i_j}$	$1,\ i\in I^{i_j}$ $0,\ i\notin I^{i_j}$	$-1,\ i\in I^{i_j}$ $0,\ i\notin I^{i_j}$	$A/D,\ i\in I^{i_j}$ $0,\quad i\notin I^{i_j}$	$-A/D,\ i\in I^{i_j}$ $0,\quad i\notin I^{i_j}$
$\partial f_i^s/\partial y_{i_j}$	0	0	$B/D,\ i\in I^{i_j}$ $0,\quad i\notin I^{i_j}$	$-B/D,\ i\in I^{i_j}$ $0,\quad i\notin I^{i_j}$

$$A = x_i - x_{i_j}^0,\quad B = y_i - y_{i_j}^0,\quad D = (A^2 + B^2)^{1/2},\quad E = ((x_i^0)^2 + (y_i^0)^2)^{1/2}$$

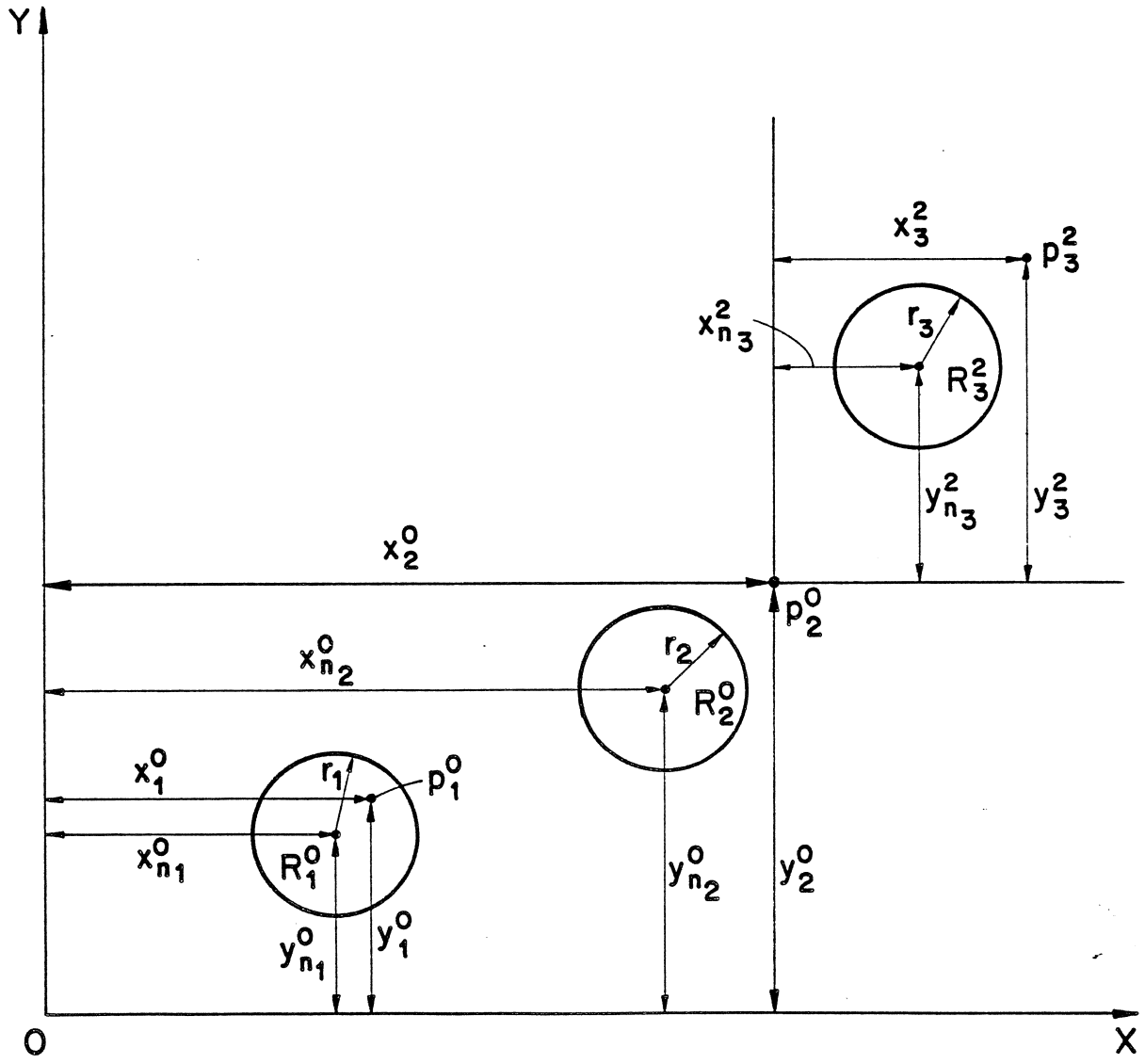


Fig. 5.4 Points with circular tolerance regions.

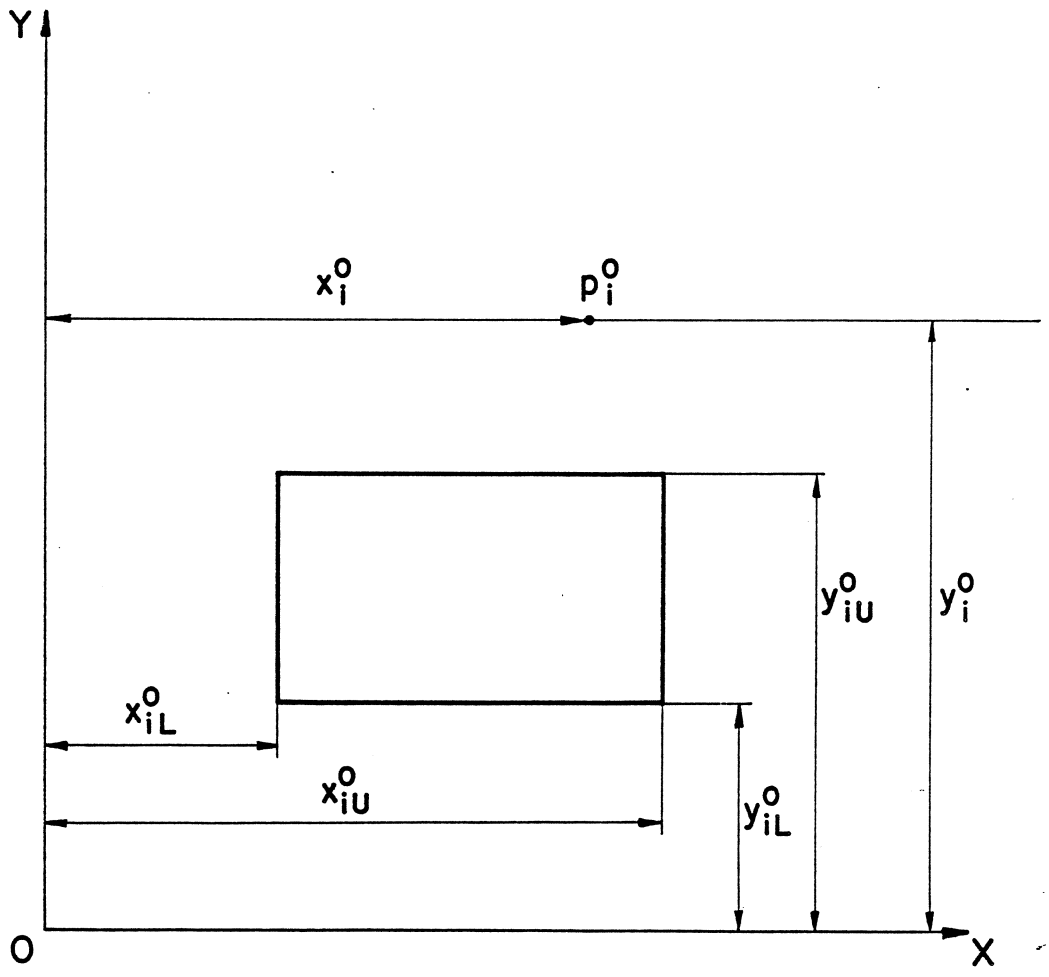


Fig. 5.5 Regular point with the rectangular tolerance region.

For a regular point and the X-R tolerance region (Fig. 5.6), the error functions result from (5.29) and from

$$(R_{iL}^0)^2 \le (x_i^0)^2 + (y_i^0)^2 \le (R_{iU}^0)^2, \tag{5.35}$$

and f_i^1, f_i^2 have the form of (5.31), (5.32), respectively, while f_i^3 and f_i^4 can be expressed as

$$f_i^3 = R_{iL}^0 - \sqrt{(x_i^0)^2 + (y_i^0)^2}\ , \tag{5.36}$$

$$f_i^4 = \sqrt{(x_i^0)^2 + (y_i^0)^2} - R_{iU}^0 . \tag{5.37}$$

Finally, for a regular point and the Y-R tolerance region (Fig. 5.7), the error functions result from (5.30) and (5.35), and f_i^1, f_i^2, f_i^3, f_i^4 are given by (5.33), (5.34), (5.36) and (5.37), respectively.

For each of these tolerance regions, we represent a point using only one error function. The four error functions may be combined into one using the following functions (Bandler and Charalambous 1972a, 1972b)

$$f_i = \begin{cases} M \left[\sum_{s \in S} \left[\dfrac{f_i^s(\Phi)}{M} \right]^q \right]^{1/q}, & \text{for } M \neq 0, \\ 0, & \text{for } M = 0, \end{cases} \tag{5.38}$$

where

$$M \triangleq \max_{s \in S_1} (f_i^s)\ ,\ S_1 \triangleq \{1, 2, 3, 4\}, \tag{5.39}$$

$$q \triangleq \frac{M}{|M|} p \begin{cases} 1 < p < \infty, & \text{for } M > 0, \\ 1 \le p < \infty, & \text{for } M < 0, \end{cases} \tag{5.40}$$

$$S \triangleq \begin{cases} S_1 & \text{if } M < 0, \\ S_2 & \text{if } M > 0, \end{cases} \tag{5.41}$$

$$S_2 \triangleq \{s \mid f_i^s \ge 0\ , s \in S_1\}. \tag{5.42}$$

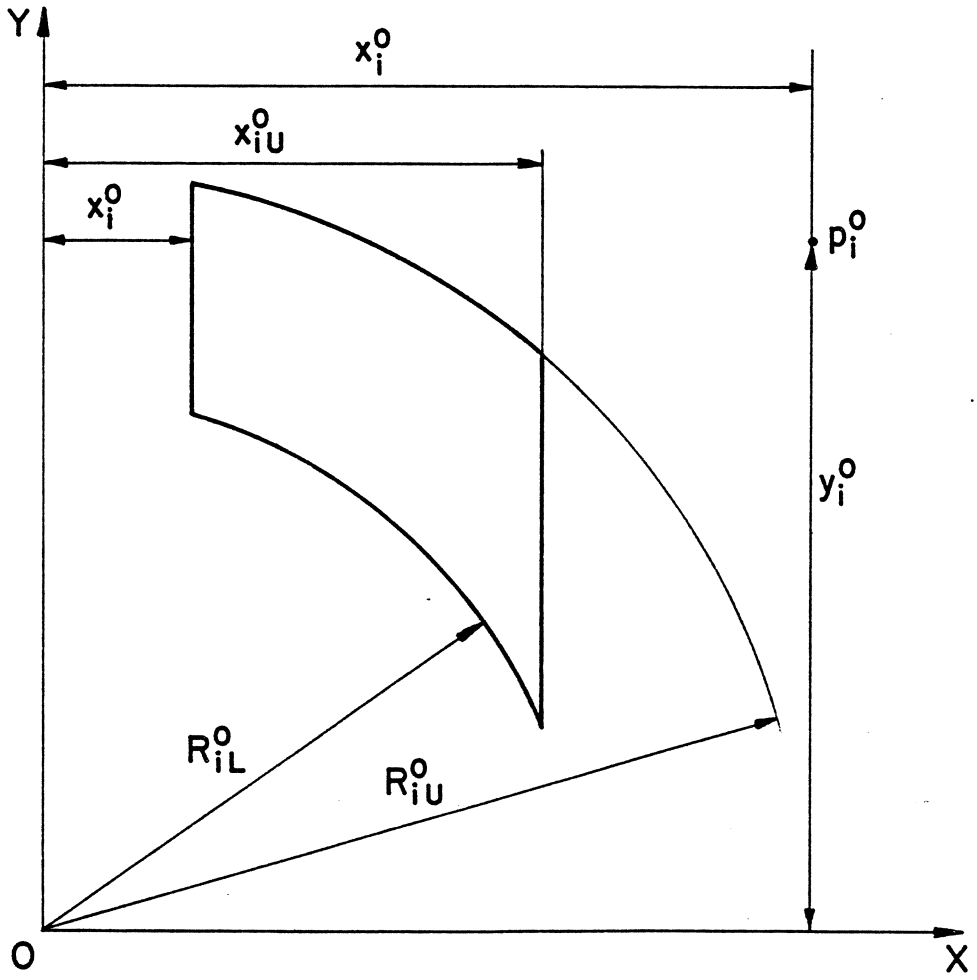


Fig. 5.6 Regular point and the X-R tolerance region.

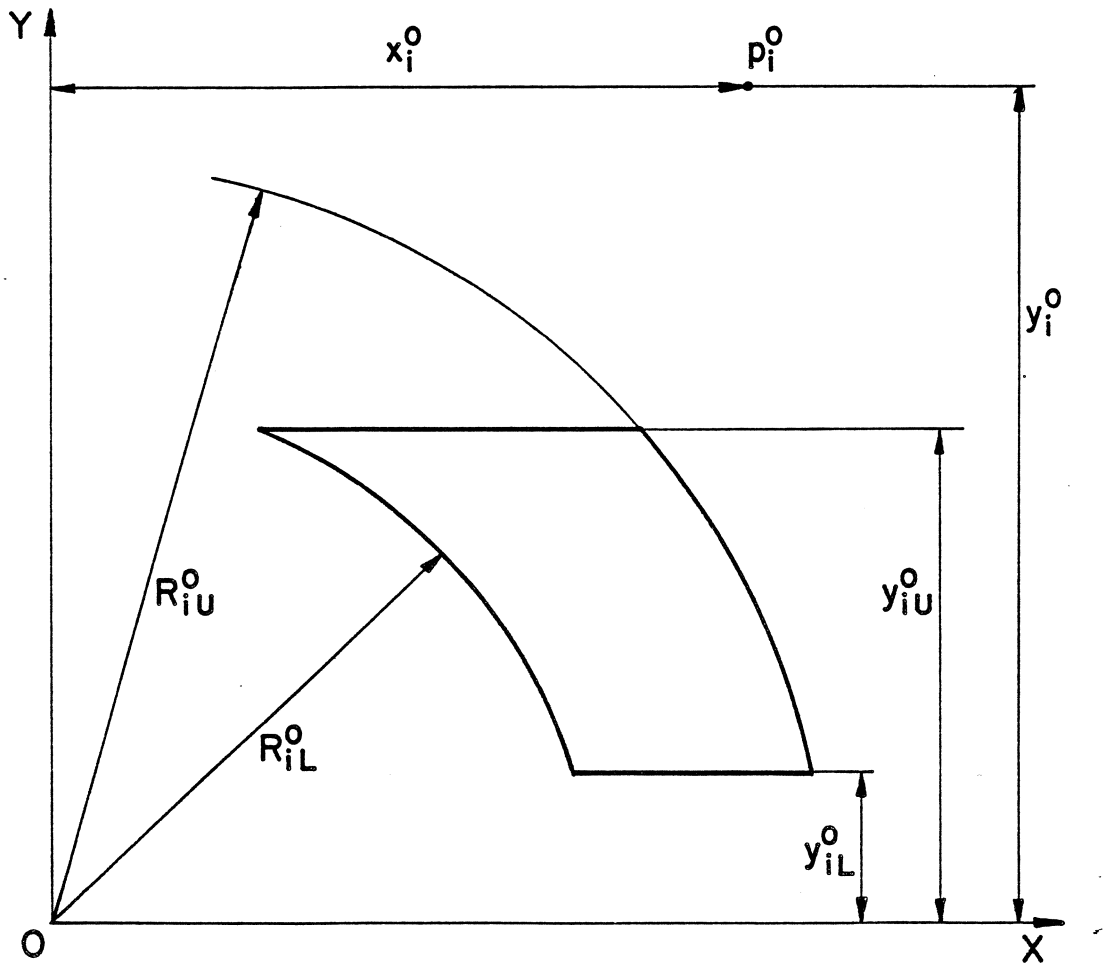


Fig. 5.7 Regular point and the Y-R tolerance region.

The gradient vector of the combined error function is given by

$$\mathbf{f}_i'(\boldsymbol{\Phi}) = \left[\sum_{s \in S} \left[\frac{f_i^s(\boldsymbol{\Phi})}{M} \right]^q \right]^{(1/q)-1} \cdot \sum_{s \in S} \left[\frac{f_i^s(\boldsymbol{\Phi})}{M} \right]^{q-1} \mathbf{f}_i'^s(\boldsymbol{\Phi}) \text{, for } M \neq 0 . \tag{5.43}$$

From (5.38) and (5.43), it can be seen that if $f_i^s(\boldsymbol{\Phi})$, s = 1,2,3,4, are continuous with continuous first partial derivatives, then, under the stated conditions, the function f_i is continuous everywhere with continuous first partial derivatives (except possibly when both M = 0 and two or more maxima are equal). For $p \rightarrow \infty$, practically

$$f_i = \max_{s \in S_1} (f_i^s)$$

The elements of the gradient vector $\mathbf{f}_i'^s$, $s \in S_1$, for the rectangular and the X-R tolerance regions are given in Tables 5.2 and 5.3, respectively. For the Y-R tolerance region error functions and their derivatives are the corresponding entries of the tables for the rectangular and the X-R tolerance regions.

5.6 COMPUTER IMPLEMENTATION OF THE ALGORITHM

(Bandler, El-Kady, Kellermann and Zuberek 1983a)

In this section, a Fortran program for solving the best alignment problem is briefly described. It has some limitations, resulting from the fact that it was designed for solving particular practical problems (e.g., the number of different shapes of tolerance regions is limited to 4). The program employs a package for linearly constrained minimax optimization (Hald 1981) available in the form of a library of subroutines.

The structure of the program is shown in Fig. 5.8. The main segment is BSTALN. It reads the data from the input file SAMPLE, prints the data, calls subroutine FDF at the starting point, calls subroutine PRSRCH and prints the final results. The subroutine PRSRCH organizes the workspace memory for SEARCH and

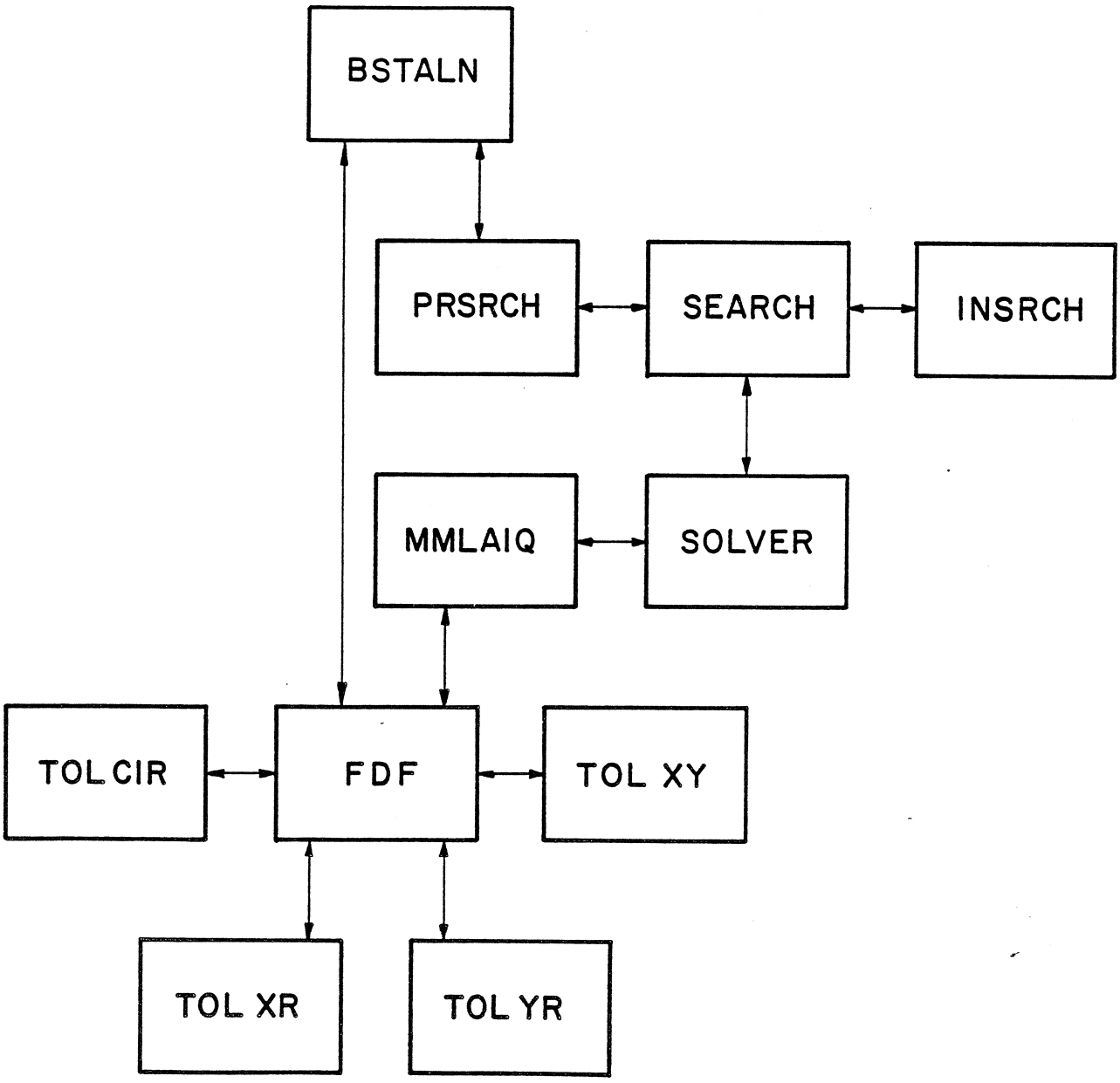


Fig. 5.8 Structure of the program for the best alignment problem.

calls SEARCH. The subroutine SEARCH implements the decision- tree structure described earlier. It calls SOLVER and INSRCH. The subroutine SOLVER prepares parameters and calls the minimax optimization routine MMLA1Q. The subroutine INSRCH eliminates identical entries in the decision-tree structure. The subroutine FDF performs the transformation of coordinates, evaluates error functions and calculates final derivatives. It calls TOLCIR, TOLXY, TOLXR and TOLYR. Subroutines TOLCIR, TOLXY, TOLXR and TOLYR calculate the error function and its derivatives for the circular, rectangular, X-R and Y-R tolerance regions, respectively, using $p = \infty$.

For the purpose of illustration an artificial simple example has been constructed.

5.6.1 Example (Bandler, El-Kady, Kellermann and Zuberek 1983b)

Suppose we have a set of points $P \triangleq \{p_1, p_2, p_3, p_4, p_5\}$ and a set of tolerance regions $R \triangleq \{R_1, R_2, R_3, R_4, R_5\}$. Fig. 5.9 illustrates the situation before the alignment. Error functions at the starting point $\phi_0^T = [0.0 \;\; 0.0 \;\; 0.0]$ are the following

$$f_1 = 2.071 \times 10^{-1},$$

$$f_2 = -5.000 \times 10^{-1},$$

$$f_3 = 5.000 \times 10^{-1},$$

$$f_4 = -5.000 \times 10^{-1},$$

$$f_5 = 5.000 \times 10^{-1}.$$

Fig. 5.10 shows the situation after running the alignment program. The best alignment was found at $\phi_0^T = [-2.316 \times 10^{-1} \;\; -2.792 \times 10^{-1} \;\; 4.758 \times 10^{-2}]$ with point 5 deleted. Remaining error functions at the solution are

$$f_1 = -1.540 \times 10^{-1},$$

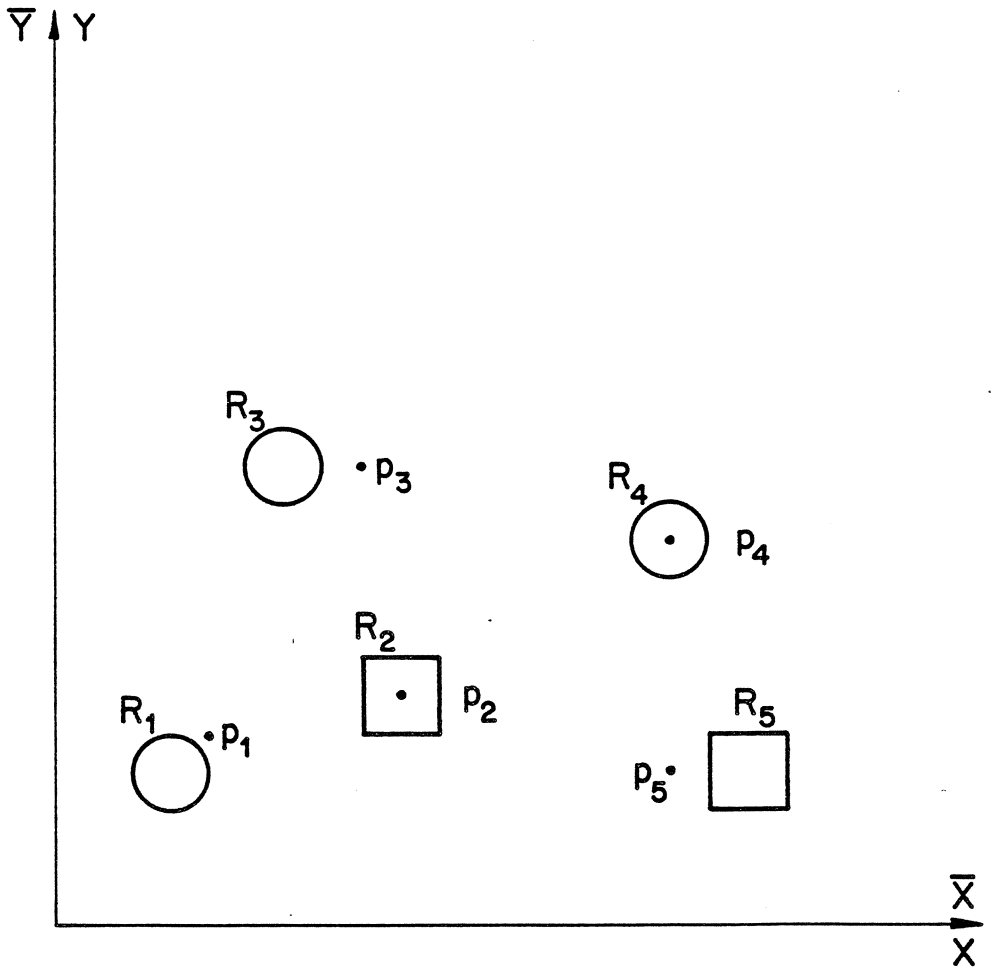


Fig. 5.9 Points and tolerance regions before alignment.

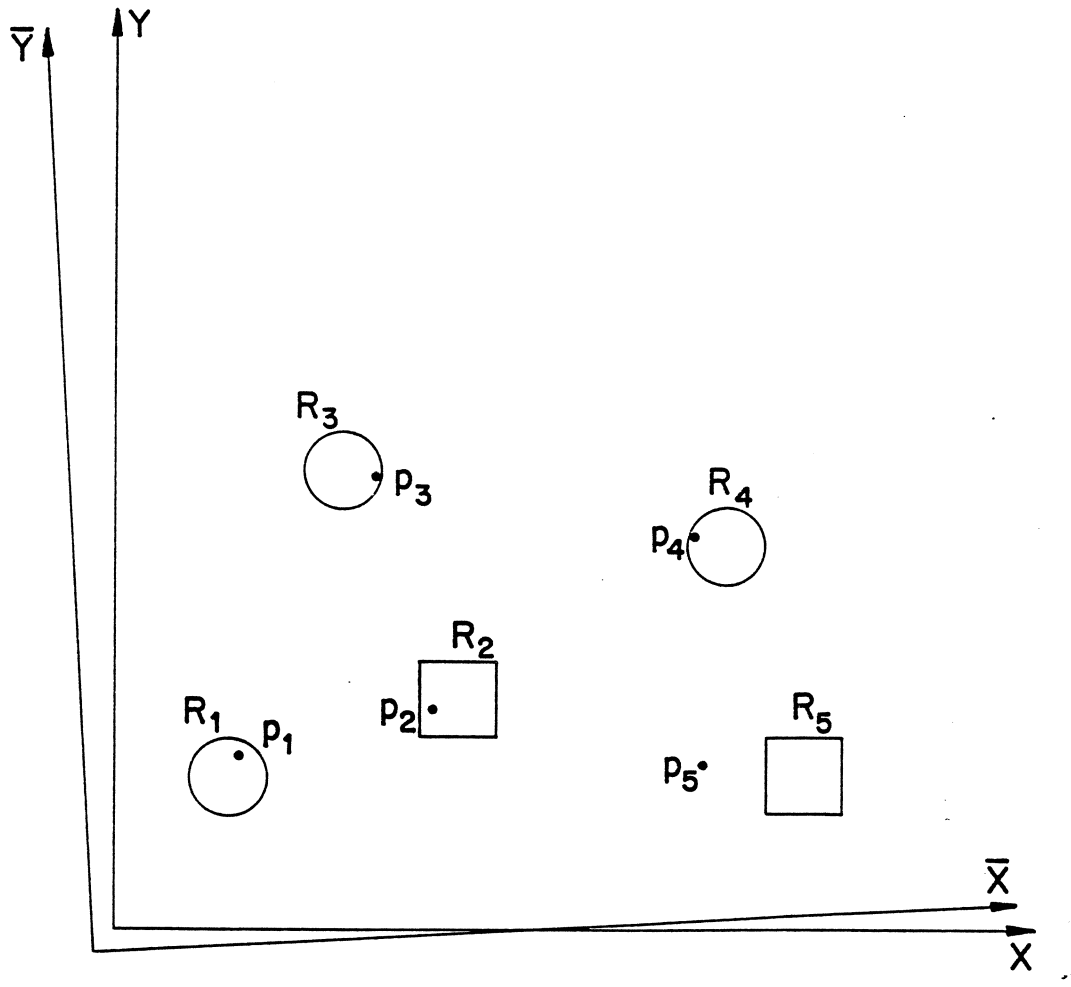


Fig. 5.10 Results of running the alignment program.

$$f_2 = -1.206 \times 10^{-1},$$

$$f_3 = -1.204 \times 10^{-2},$$

$$f_4 = -1.204 \times 10^{-2}.$$

5.7 TEST RESULTS ON PRACTICAL PROBLEMS

The program described in the previous section has been extensively tested. It has been run for seven sets of data (Woodward Governor Company 1982). The data resulted from practical problems of part alignment in manufactured mechanical systems and have been collected from inspecting actual parts, so the order of error function values represents the real life situation. The points represent holes in one part which have to meet certain specifications when coupled together with another part. Test samples have different numbers of points, varying from 5 to 13 and specified tolerance regions of different shapes. To give an idea of what the samples are like, we describe briefly two simple samples and one interesting sample in more detail.

Sample 1 (Table 5.4)

This sample has 5 points, 1 with circular and 4 with the rectangular tolerance regions. It has no reference points. Originally, the number of points out-of-tolerance was 4. After 12 iterations of stage 0, the minimum value of the maximum error function was 3.6078×10^{-4}. Three points (1, 3 and 4) have been selected as potential candidates for deleting. It turns out that deleting point number 1 gives the solution for which the remaining error functions are negative and the maximum error at the solution was -6.45668×10^{-4} (after 25 additional minimax iterations):

TABLE 5.4

DATA FOR SAMPLE 1 (Woodward Governor Company 1982)

Point	Tolerance Code +	Origin Code *	Actual $\bar{x}$	Actual $\bar{y}$	Tolerances			
					x_N	y_N	r_N	
1	0	0	0.0000	0.0000	0.0000	0.0000	0.0010	
					x_L	x_U	y_L	y_U
2	12	0	-0.8800	1.3682	-0.8780	-0.8750	1.3690	1.3720
3	12	0	0.6589	0.7499	0.6610	0.6630	0.7500	0.7520
4	12	0	0.8990	-0.4414	0.8990	0.9010	-0.4410	-0.4380
5	12	0	-0.5635	-1.5254	-0.5650	-0.5620	-1.5250	-1.5520

+ The tolerance code is one of four (0, 12, 13, 23), where

0 - the code for the circular tolerance region,
12 - the code for the rectangular tolerance region,
13 - the code for the X-R tolerance region,
23 - the code for the Y-R tolerance region.

* Any point with an origin code of 0 is referenced to the main origin of $\bar{x} = 0.0$, $\bar{y} = 0.0$. Any other origin code refers to the point by that number on the same sample. For instance, for an origin code of 4, the actual $\bar{x}$ and $\bar{y}$ dimensions are measured from the actual $\bar{x}$ and $\bar{y}$ dimensions of point number 4.

Sample 2 (Table 5.5)

This sample has 7 points, all with the circular tolerance regions and all referenced to the main origin. Originally, the number of points out-of-tolerance was 5. After 15 iterations, the solution was found with no deletions and the maximum error at the solution was -7.73563×10^{-4}.

Sample 6 (Table 5.6)

This sample is very interesting: it has 11 points, 4 with circular, 4 with rectangular, 1 with the X-R and 2 with the Y-R tolerance regions. Five points are referenced to points other than the main origin. Previous work on the best alignment problem (Peterson and Johnson 1980) does not permit a reference point to be deleted (translated). In our approach, any point can be deleted. Originally, there were 2 points out-of-tolerance, and one of them is a reference point. When a point which is an origin for one or more points is found to be out-of-tolerance, there is a good chance that any point referenced to it will also appear to be off location. In this sample, points 7 and 8 are referenced to point 1. Points 1 and 8 were both found to be out-of-tolerance. However, if point 1 was shifted by the amount specified (in other words, if hole number 1 was plugged and re-drilled in the proper location), point 8 would be in-tolerance without any rework needed. Thus, in a practical mechanical sense, there is only one point out-of-tolerance, that being point 1 (Peterson and Johnson 1980). Results of running the program for Sample 6 show that indeed deleting reference point 1 (plugging and redrilling hole) implies that all other points will be in-tolerance and the maximum error at the solution is -1.9911×10^{-4}.

We can observe how point 1 was selected for deleting from the details of the solution, given in Table 5.7. From the results of minimax optimization at stage 0,

TABLE 5.5

DATA FOR SAMPLE 2 (Woodward Governor Company 1982)

Point	Tolerance Code +	Origin Code *	Actual $\bar{x}$	Actual $\bar{y}$	Tolerances		
					x_N	y_N	r_N
1	0	0	0.0000	-0.0001	0.0000	0.0000	0.0050
2	0	0	-0.6412	1.1080	-0.6405	1.1094	0.0025
3	0	0	-1.2778	-0.0052	-1.2810	0.0000	0.0025
4	0	0	-0.6295	-1.1101	-0.6405	-1.1094	0.0025
5	0	0	0.6499	-1.1055	0.6405	-1.1094	0.0025
6	0	0	1.2846	0.0083	1.2810	0.0000	0.0025
7	0	0	0.6393	1.1126	0.6405	1.1094	0.0025

+ see Table 5.4 for code explanations.
* see Table 5.4 for code explanations.

TABLE 5.6

DATA FOR SAMPLE 6 (Woodward Governor Company 1982)

Point	Tolerance Code +	Origin Code *	Actual $\bar{x}$	Actual $\bar{y}$	Tolerances			
					x_N	y_N	r_N	
1	0	0	2.3970	-0.9508	2.3950	-0.9500	0.0010	
2	0	0	-1.6955	-1.9621	-1.6960	-1.9620	0.0010	
					x_L	x_U	y_L	y_U
3	12	0	0.6620	0.7507	0.6610	0.6630	0.7500	0.7520
4	12	0	0.8998	-0.4393	0.8990	0.9010	-0.4410	-0.4380
					y_L	y_U	R_L	R_U
5	23	0	-0.5629	-1.5231	-1.5260	-1.5210	1.6225	1.6260
					x_L	x_U	y_L	y_U
6	12	0	-0.8773	1.3700	-0.8780	-0.8750	1.3690	1.3720
					x_N	x_N	r_N	
7	0	1	-2.8646	3.5015	-2.8640	3.5010	0.0010	
					x_L	x_U	y_L	y_U
8	12	1	-0.8764	2.3274	-0.8750	-0.8710	2.3250	2.3290
					x_N	x_N	r_N	
9	0	4	0.6653	-0.7855	0.6650	-0.7860	0.0010	
					y_L	y_U	R_L	R_U
10	23	5	-0.9642	1.0227	1.0210	1.0260	1.4053	1.4073
					x_L	x_U	y_L	y_U
11	13	6	-0.0641	-1.1348	-0.0660	-0.0640	1.1358	1.1378

+,* see Table 5.4 for code explanations.

TABLE 5.7

RESULTS OF BEST MINIMAX ALIGNMENT FOR SAMPLE 6

(Woodward Governor Company 1982)

Error Function	Values of Error Functions +		
	Starting Point	Stage 0 Optimization (no deletions)	Optimization with Point 1 Deleted (translated)
1	1.1540659×10^{-3}	$\underline{7.8766877 \times 10^{-4}}$	$-6.0836163 \times 10^{-4}$*
2	$-4.9009805 \times 10^{-4}$	7.8054088×10^{-4}	$-3.1859860 \times 10^{-4}$
3	$-7.0000000 \times 10^{-4}$	$-6.7451522 \times 10^{-4}$	$-6.0366698 \times 10^{-4}$
4	$-8.0000000 \times 10^{-4}$	$-5.1145712 \times 10^{-4}$	$-6.0460585 \times 10^{-4}$
5	$-1.2887855 \times 10^{-3}$	$-4.1859431 \times 10^{-4}$	$-1.3816043 \times 10^{-3}$
6	$-7.0000000 \times 10^{-4}$	$-6.1087476 \times 10^{-4}$	$\underline{-1.9911453 \times 10^{-4}}$
7	$-2.1897503 \times 10^{-4}$	$\underline{7.8766877 \times 10^{-4}}$	$\underline{-1.9911453 \times 10^{-4}}$
8	1.4000000×10^{-3}	$\underline{7.8766877 \times 10^{-4}}$	$\underline{-1.9911453 \times 10^{-4}}$
9	$-4.1690481 \times 10^{-4}$	$-2.2387620 \times 10^{-4}$	$\underline{-1.9911453 \times 10^{-4}}$
10	$-2.5929437 \times 10^{-4}$	$-2.7637365 \times 10^{-4}$	$\underline{-1.9911453 \times 10^{-4}}$
11	$-1.0000000 \times 10^{-4}$	6.1249301×10^{-4}	$-4.0926333 \times 10^{-4}$

+ Maximum error functions are underlined

* This error function value corresponds to the new location of point 1.

points 1, 7 and 8 are selected as candidates for deleting. Results of minimax optimization with point 1 deleted (translated) show that a solution can be obtained with only one point deleted.

The results of running the program for all test samples are summarized in Table 5.8.

5.8 CONCLUDING REMARKS

This chapter provides an attempt to formulate and to solve the best mechanical alignment problem using minimax optimization. Results of running the best alignment program for practical problems (Table 5.8) confirm the efficiency of our approach. The concepts introduced and the algorithm proposed are described in this chapter by tackling a particular mechanical engineering problem. However, this class of problem may come from different sources and further generalization is possible. One natural extension of this approach, which may be very useful from the practical point of view, is considering alignment problems in three dimensions. Another suggestion for further exploration is the investigation of the least pth formulation to reduce the number of minimax functions.

The problem which originated from aligning mechanical designs is here formulated as a general optimization problem and we feel that this approach should prove useful in many other areas where problems of a similar nature may exist.

TABLE 5.8

RESULTS OF RUNNING THE BEST ALIGNMENT PROGRAM

ON DATA SUPPLIED BY THE WOODWARD GOVERNOR COMPANY

Sample No.	Total No. of Points	No. of Points Originally Out of Tolerance	Results (Points Deleted)	Comments	CYBER 170/730 Execution Time in Seconds
1	5	4	1	Reg. Point Deleted	0.7
2	7	5	0	No Deletions	0.4
3	11	2	1	Ref. Point Deleted	0.9
4	11	3	2	Reference and Reg. Points Deleted	2.8
5	11	3	2	Reference and Reg. Points Deleted	1.5
6	11	2	1	Reference Point Deleted	1.2
7	13	3	3	Regular Points Deleted	3.6

6

LOAD SHEDDING AND GENERATION RESCHEDULING IN POWER SYSTEMS

6.1 INTRODUCTION

This chapter provides an attempt to formulate and to solve the load shedding and generation rescheduling problem in an emergency state using a nonlinear optimization algorithm with the ℓ_1 type objective function.

The problem arises in many practical situations when an operating power system is under emergency conditions and some security constraints imposed on the system are violated. The control action taken may involve load shedding and generation rescheduling. The problem can be formulated as an optimization problem with the objective of minimizing the amount of customer load curtailments. We are concerned with the problem of selecting a possibly small subset of loads to be shed from the set of all candidates for shedding.

From the security (in the static sense) point of view there are two types of contingencies (Talukdar and Wu 1981):

- a sudden change in the power injection to the network caused by the partial or complete loss of a generator, load or tie;
- a sudden change in the network's configuration.

If the system is to be normal and secure in the static sense it must satisfy a set of algebraic constraints that can be written in the form

$$\mathbf{h}(\mathbf{x}, \mathbf{u}) = \mathbf{0}\,, \tag{6.1}$$

$$\mathbf{g}(\mathbf{x}, \mathbf{u}) \geq \mathbf{0}\,, \tag{6.2}$$

where $\mathbf{u}$ is the set of control variables (decision variables) and $\mathbf{x}$ is the set of state variables (dependent variables). Equation (6.1) is a set of equality constraints called the power flow equations and equation (6.2) is a set of inequality constraints often called security constraints.

If the system is insecure or in an emergency condition the control action taken to remove the violations of security constraints (6.2) (assuming that the system survives the outage and equation (6.1) can be solved) may involve load shedding and generation rescheduling. In such situations we usually want to minimize the control action (generation rescheduling and load shedding), often considering load shedding only as a last resort. If the load shedding is absolutely necessary to remove the inequality constraints violations then the objective may be to select the minimum number of interruptible loads for partial or complete curtailment.

Several approaches have been investigated to find post emergency schedule with the minimum of load shedding, however, in none of them the problem of selecting the minimum number of loads for shedding has been addressed directly.

A simple linear model which takes into account only the real power injections and voltage angles has been used for generation and load rescheduling (Ghoneim, Askourah and Rahman 1977, Stott and Hobson 1977a, 1977b, Chan and Schweppe 1979). Special techniques were used to improve the performance of algorithms proposed such as the dual simplex method (Stott and Hobson 1977a, 1977b) and sparse linear programming (Chan and Yip 1979).

The real-power-voltage-angle model may not be adequate for certain applications such as real-time emergency state control or detail off-line studies

(planning) where the voltage profile and the scheduling of reactive power sources are important (Chan and Schweppe 1979).

A different approach, which does not use optimization techniques has been presented in Medicherla, Billinton and Sachdev (1979, 1981). It basically consists of two sets of equations. First, a set of equations for determining the desired increments in state variables to alleviate line overloads is developed and then the second set of equations is developed, which can be solved for determining the generation rescheduling and load curtailment pattern to satisfy the calculated changes in state variables.

Recently, a multi-stage formulation of the generation rescheduling and load shedding problem has been proposed (Krogh and Javid 1983). It includes a technique for modeling the short-term thermal capacity of a transmission line, a generation scheduling model and constraints and a linear load flow model. The multi-stage formulation is then cast into a linear programming framework.

The purpose of this chapter is to give a general formulation of the load shedding and generation rescheduling problem which uses the Han-Powell algorithm for the general nonlinear programming problem (Han 1976, Powell 1977) with the ℓ_1 type objective function suggested in Bandler, El-Kady, Kellermann and Zuberek (1983b) to select a possibly small subset of loads to be shed from the set of all candidates for shedding. The Tellegen theorem method for power system sensitivity calculations (Bandler and El-Kady 1980a, 1980b) and the fast decoupled load flow technique (Stott and Alsac 1974) are employed. Numerical results for a 6-bus and a 26-bus power systems illustrate the performance of the method proposed.

6.2 FORMULATION OF THE PROBLEM

6.2.1 Objective Function and Constraints

The load shedding and generation rescheduling problem can be formulated as a mathematical optimization problem which attempts to minimize the number of loads to be shed under emergency conditions which can take the form of line outages or generator outages (partial or complete). The objective function used in this approach is of the ℓ_1 type. Formally the problem is to

$$\underset{|V_g|,\, P_g,\, P_\ell}{\text{minimize}} \sum_{\ell=1}^{n_L} (P_\ell - P_\ell^0), \qquad \begin{matrix} g = n_L + 1, \ldots, n_L + n_G, \\ \ell = 1, \ldots, n_L, \end{matrix} \tag{6.3}$$

subject to (6.1) and (6.2) corresponding to the emergency conditions, where

$|V_g|$ is the voltage magnitude at node g,

P_g is the real power injected at generator node g,

P_ℓ is the real power injected at load node ℓ,

P_ℓ^0 is the nominal real load power at node ℓ.

We let subscripts $\ell = 1,2,\ldots,n_L$, correspond to PQ buses (load buses), subscripts $g = n_L+1, \ldots, n_L + n_G$, correspond to PV buses (generator buses), and subscript $n = n_L + n_G+1$ corresponds to the slack bus. The objective function (6.3) has the property of setting as many as possible terms $(P_\ell - P_\ell^0)$, $\ell = 1, \ldots, n_L$, to zero at the solution. This corresponds to selecting a possibly small subset of loads to be shed under emergency conditions from the set of all loads being candidates for shedding. The set of optimization variables usually restricted only to control variables in optimal power flow problems, $|V_g|$ and P_g = $g = n_L+1, \ldots, n_L+n_G$, is here extended with the real load powers P_ℓ, $\ell = 1, \ldots, n_L$. In emergencies the operator may augment the decision vector with certain loads whose values can then be adjusted downwards. The objective (6.3)

simulates that process, which is basically a process of designing the real load powers (and the generator voltage magnitudes together with the generator real powers) to satisfy the security constraints.

The real and reactive loads are not independent; one cannot shed real loads without curtailing reactive loads. It is assumed in this approach, as well as in Chan and Schweppe (1979), that the power factor remains fixed after a portion of the load is shedded.

The minimization of (6.3) is subjected to (6.1) and (6.2), where in (6.1) it is assumed that the power network is represented by a steady-state a.c. power flow model in which the power flow equations are obtained by equating the power injected into each node with the power removed from the node and recognizing that under steady-state conditions the network can be represented by an admittance matrix (Talukdar and Wu 1981). In our implementation of the method proposed the fast decoupled load flow technique was used.

Security constraints (6.2) in terms of the power network variables form six groups of constraints, both linear and nonlinear w.r.t. optimization variables chosen:

1) load real power constraints (linear)

$$P_\ell^0 \le P_\ell \le 0\,, \quad \ell = 1, 2, \ldots, n_L\,, \tag{6.4}$$

2) load voltage magnitude constraints (nonlinear)

$$|V_\ell|^{\min} \le |V_\ell| \le |V_\ell|^{\max}, \quad \ell = 1, 2, \ldots, n_L\,, \tag{6.5}$$

3) voltage angle constraints (nonlinear)

$$\delta_i^{\min} \le \delta_i \le \delta_i^{\max}, \quad i = 1, 2, \ldots, n_L + n_G\,, \tag{6.6}$$

4) generator real power constraints (linear)

$$P_g^{\min} \le P_g \le P_g^{\max}, \quad g = n_L + 1, \ldots, n_L + n_G\,, \tag{6.7a}$$

$$P_n^{\min} \le P_n \le P_n^{\max}\,, \tag{6.7b}$$

5) generator reactive power constraints (nonlinear)

$$Q_g^{min} \le Q_g \le Q_g^{max}, \quad g = n_L + 1, ..., n, \tag{6.8}$$

6) generator voltage magnitude constraints (linear)

$$|V_g|^{min} \le |V_g| \le |V_g|^{max}, \quad g = n_L + 1, ..., n_L + n_G. \tag{6.9}$$

Transmission line loading limits can be current-magnitude constraints due to thermal considerations or electrical-angle (difference in voltage angles across a line) constraints due to stability considerations. In this formulation by appropriately choosing δ_i^{min} and δ_i^{max} in (6.6) the stability limits in terms of θ_{ij} ($\theta_{ij} = \delta_i - \delta_j$) can be accommodated as well as the current-magnitude constraints which can be converted into the form

$$-\theta_{ij}^{max} \le \theta_{ij} \le \theta_{ij}^{max}, \tag{6.10}$$

as shown in Chan and Yip (1979).

6.2.2 Gradients of the Objective Function and Constraints

The optimization method implemented in the form of a Fortran package (Bandler and Zuberek 1982b) requires first-order derivatives of the objective function and the constraints to be available. Since the voltage magnitudes $|V_g|$, real generator powers P_g, $g = n_L+1, ..., n_L+n_G$, and real load powers P_ℓ, $\ell = 1, ..., n_L$, are assumed to be optimization variables the gradient of the objective function (6.3) and the gradients of the constraints (6.4), (6.7a) and (6.9) can be easily found. To find the gradients of the constraints (6.5), (6.6), (6.7b) and (6.8) the Tellegen theorem method was used (Bandler and El-Kady 1980a, 1980b) and an implementation in the form of the TTM1 package (Bandler, El-Kady and Wojciechowski 1983).

6.3 REMARKS ON THE IMPLEMENTATION OF THE FAST DECOUPLED METHOD

The implementation (Bandler and Zuberek 1982c) of the fast decoupled load flow technique, incorporated in the TTM1 package, basically follows the method of Stott and Alsac, and differs only in a more flexible iteration scheme in which the order of successive P-δ and Q-V iterations is not fixed but depends on the relationship between the accuracies of P-δ and Q-V iterations. Let ε_δ be the accuracy of the last P-δ iteration, and ε_V be the accuracy of the last Q-V iteration. If, after the kth iteration

$$\varepsilon_\delta \geq 2\,\varepsilon_v$$

the $(k+1)$th iteration is the P-δ one. If

$$\varepsilon_v \geq 2\,\varepsilon_\delta$$

the $(k+1)$th iteration is the Q-V one. Otherwise, the basic (1δ, 1V) is followed. The implemented scheme tends to avoid large differences between the two accuracies and over-converging of the one of the two systems of equations implied by the method which slows down the overall convergence rate (Stott and Alsac 1974).

6.4 NUMERICAL RESULTS: 6-BUS AND 26-BUS SYSTEMS

6 Bus System

The method presented has been tested on a 6-bus sample power system shown in Fig. 6.1. The required data is given in Table 6.1. Powers injected into buses are shown. Contingency in the form of the loss of 50% of generation capacity at bus 5 is considered. Upper and lower bounds on network variables in terms of the base case solution are given in Table 6.2. The results of optimization showing the variables at the starting point and at the solution are summarized in Table 6.3. It can be seen that the objective function (6.3) selected only one load to be shed (out of three), namely, P_1.

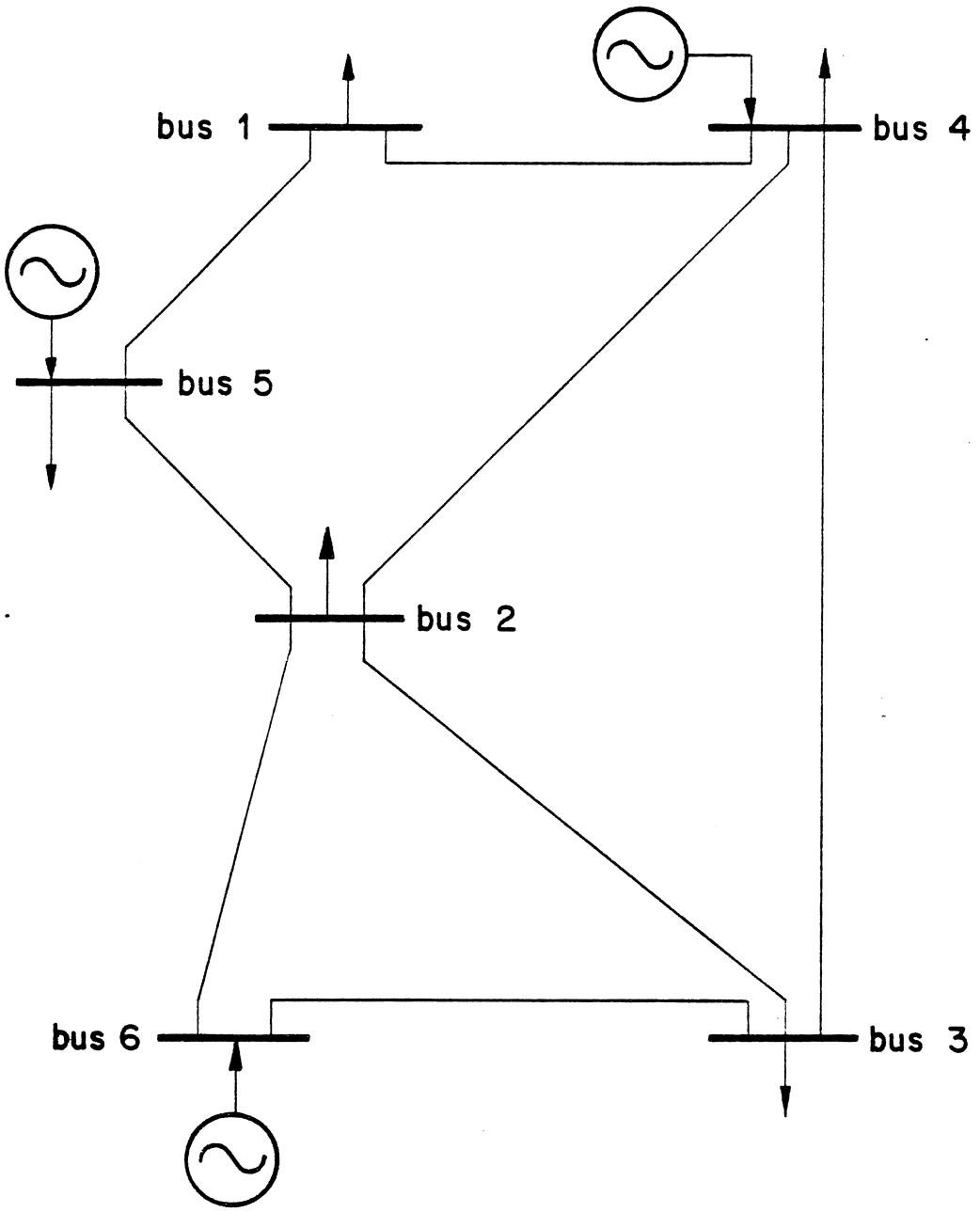


Fig. 6.1 6-bus power system.

TABLE 6.1a

BUS DATA FOR THE 6-BUS POWER SYSTEM

Bus Index, i	Bus Type	P_i (pu)	Q_i (pu)	$\lvert V_i \rvert \angle \delta_i$ (pu)
1	load	-2.40	0	- ∠ -
2	load	-2.40	0	- ∠ -
3	load	-1.60	-0.40	- ∠ -
4	generator	-0.30	-	1.02 ∠ -
5	generator	1.25	-	1.04 ∠ -
6	slack	-	-	1.04 ∠ 0.0

TABLE 6.1b

LINE DATA FOR THE 6-BUS POWER SYSTEM

Branch Index, t	Terminal Buses	Resistance R_t (pu)	Reactance X_t (pu)	Number of Lines
7	1,4	0.05	0.20	1
8	1,5	0.025	0.10	2
9	2,3	0.10	0.40	1
10	2,4	0.10	0.40	1
11	2,5	0.05	0.20	1
12	2,6	0.01875	0.075	4
13	3,4	0.15	0.60	1
14	3,6	0.0375	0.15	2

TABLE 6.2

UPPER AND LOWER BOUNDS ON NETWORK VARIABLES

FOR THE 6-BUS POWER SYSTEM

Variable	Lower Bound	Upper Bound
$\|V_\ell\|$	$0.93\,\|V_\ell\|^0$	$1.07\,\|V_\ell\|^0$
δ_i	$0.9\,\delta_i^0$	$1.1\,\delta_i^0$
P_g	$0.85\,P_g^0$	$1.3\,P_g^0$
Q_g	$0.85\,Q_g^0$	$1.3\,Q_g^0$
$\|V_g\|$	$0.85\,\|V_g\|^0$	$1.15\,\|V_g\|^0$

$|V_\ell|^0$, δ_i^0, P_g^0, Q_g^0 and $|V_g|^0$ denote nominal values (at the base-case solution) except for the power injected at mode 5, for which $P_5^0 = 0.625$ (taking into account the forced outage of 50% of generating capacity).

TABLE 6.3

RESULTS OF RUNNING THE LOAD SHEDDING PROGRAM

FOR THE 6-BUS POWER SYSTEM

	Starting Point	Solution
Variable		
P_1	-2.40	-2.07225
P_2	-2.40	-2.40000
P_3	-1.60	-1.60000
P_4	-0.30	-0.25500
P_5	0.625	0.704375
$\|V_4\|$	1.02	1.040629
$\|V_5\|$	1.04	1.037985
Objective Function	0.0	0.327745
No. of Function Evaluations	20	
Execution Time on Cyber 170/815 (in seconds)	4.5	

The other two, P_2 and P_3 are exactly the same as at the starting point. The objective function value at the solution is exactly equal to the amount of load no. 1 to be curtailed. The changes in values of P_4 and P_5 reflect the fact that we allowed up to 30% incrase in real generation at buses 4 and 5. It can be observed also that the 30% reserve has not been utilized in full. This can be explained by the fact that other constraints have limited the use of full reserve at buses 4 and 5.

26-Bus System

The method presented has also been tested on a 26-bus power system shown in Fig. 6.2. The required data for the system is given in Tables 6.4, 6.5 and 6.6. Contingency in the form of loss of generation capacity at bus 19 from 1.45 to 1.00 is considered. Upper and lower bounds on network variables in terms of the base case solution are the same as those given in Table 6.2 for the 6-bus system with the exception of P_{19} for which P_{19}^0 is taken as 1.00 (taking the contingency into account). The results of optimization showing all variables at the starting point and at the solution are summarized in Table 6.7. From the results it can be seen that again only one load (P_6) out of seventeen has been selected for shedding, assuming contingency at bus 19. All the other loads have not been affected. The objective function at the solution is exactly equal to the amount of load at bus 6 to be curtailed.

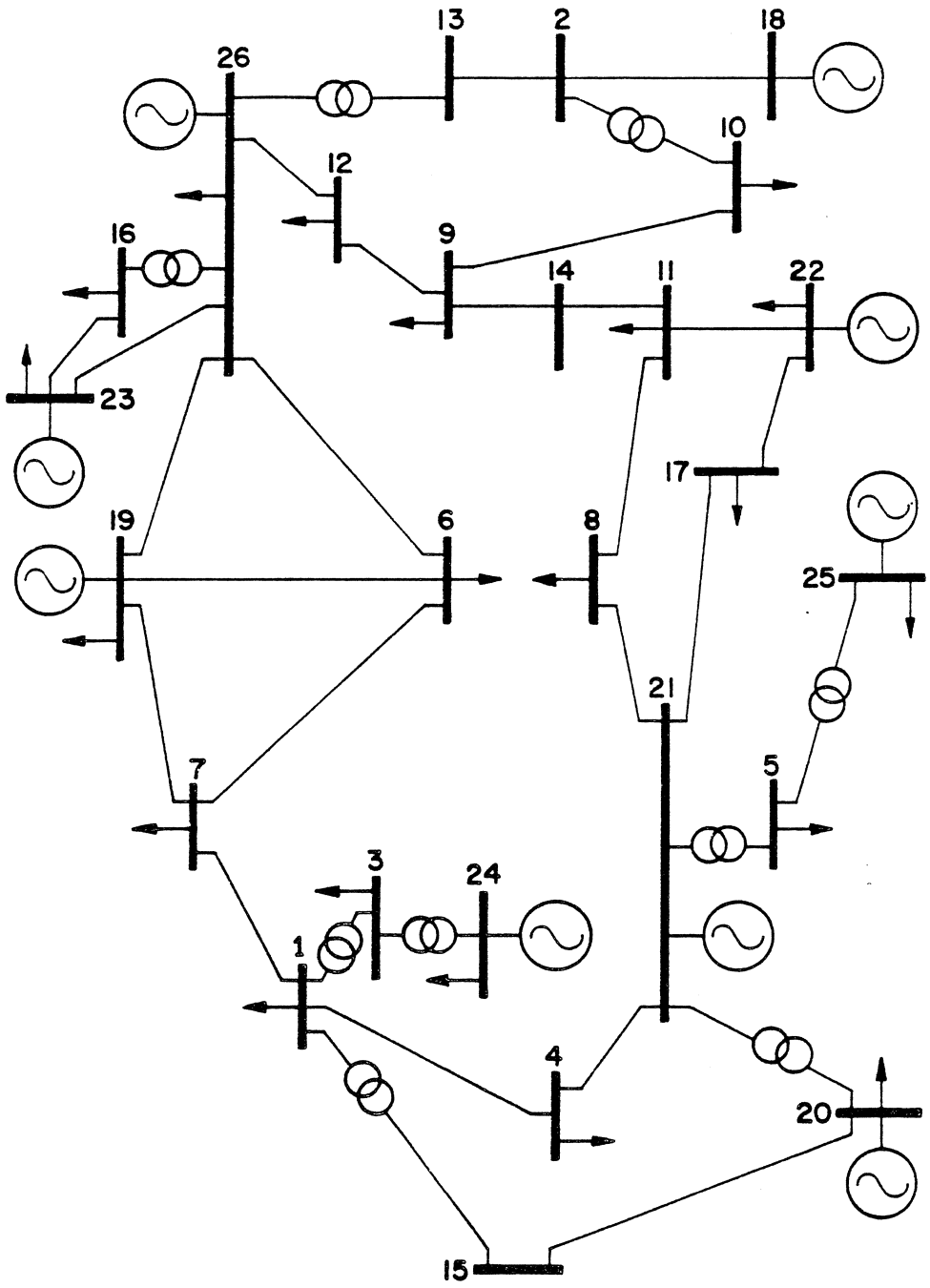


Fig. 6.2 26-bus power system.

TABLE 6.4

BUS DATA FOR THE 26-BUS POWER SYSTEM

Bus	Injected Power		Bus Voltage	
	P_m	Q_m	$\|V_m\|$	δ_m
1	-0.82	-0.21	-	-
2	0.01	0.0	-	-
3	-0.57	-0.17	-	-
4	-0.48	-0.21	-	-
5	-0.43	-0.11	-	-
6	-0.40	-0.10	-	-
7	-1.11	-0.27	-	-
8	-0.23	-0.06	-	-
9	-0.67	-0.21	-	-
10	-1.02	-0.27	-	-
11	-0.43	-0.14	-	-
12	-0.43	-0.12	-	-
13	0.01	0.0	-	-
14	0.01	0.0	-	-
15	0.01	0.0	-	-
16	-1.31	-0.30	-	-
17	-0.03	-0.01	-	-
18	2.80	-	1.07	-
19	1.45	-	1.05	-
20	2.80	-	1.00	-
21	1.10	-	1.02	-
22	-0.56	-	0.89	-
23	-0.04	-	1.00	-
24	-0.05	-	1.00	-
25	0.63	-	1.00	-
26	0.0	-	1.01	0.0

TABLE 6.5

LINE DATA FOR THE 26-BUS POWER SYSTEM

Line	Terminal Buses	Resistance R_t (pu)	Reactance X_t (pu)	1/2 Shunt Susceptance
1	13,26	0.0	0.0131	0.0
2	26,16	0.0	0.0392	0.0
3	16,23	0.0	0.4320	0.0
4	23,26	0.0	0.3140	0.0
5	2,10	0.0	0.0150	0.0
6	9,10	0.1494	0.3392	0.4120
7	9,12	0.0658	0.1494	0.0182
8	12,26	0.0533	0.1210	0.0147
9	9,14	0.0618	0.2397	0.0319
10	11,14	0.0676	0.2620	0.0349
11	19,26	0.0610	0.2521	0.0295
12	6,26	0.0513	0.1986	0.0265
13	6,19	0.0129	0.0532	0.0074
14	7,19	0.0906	0.3742	0.0437
15	6,7	0.0921	0.3569	0.0475
16	11,22	0.0513	0.2118	0.0248
17	8,11	0.0865	0.3355	0.0447
18	17,22	0.0281	0.1869	0.0237
19	8,21	0.0735	0.2847	0.0379
20	17,21	0.0459	0.3055	0.0387
21	1,4	0.0619	0.2401	0.0319
22	4,21	0.0610	0.2365	0.0315
23	20,21	0.0	0.0305	0.0
24	15,1	0.0	0.0147	0.0
25	2,13	0.0086	0.0707	0.3017
26	1,7	0.0199	0.0785	0.0404
27	15,20	0.0107	0.0617	0.4471
28	2,18	0.0074	0.0608	0.2593
29	1,3	0.0	0.0392	0.0
30	24,3	0.0	0.1450	0.0
31	5,21	0.0	0.1750	0.0
32	5,25	0.0	0.1540	0.0

TABLE 6.6

TRANSFORMER TAPS FOR THE 26-BUS POWER SYSTEM

No.	Terminal Buses	Real	Imaginary
1	13,26	1.03	0.0
2	20,21	0.97	0.0
3	24,3	0.98	0.0
4	26,16	0.96	0.0
5	15,1	0.89	0.0
6	5,21	0.99	0.0
7	2,10	1.03	0.0
8	1,3	0.98	0.0
9	5,25	1.03	0.0

TABLE 6.7

RESULTS OF RUNNING THE LOAD SHEDDING PROGRAM

FOR THE 26-BUS POWER SYSTEM

	Starting Point	Solution
Variable		
P_1	-0.82	-0.82000
P_2	-0.01	-0.01000
P_3	-0.57	-0.57000
P_4	-0.48	-0.48000
P_5	-0.43	-0.43000
P_6	-0.40	-0.155187
P_7	-1.11	-1.11
P_8	-0.23	-0.23000
P_9	-0.67	-0.67000
P_{10}	-1.02	-1.02000
P_{11}	-0.43	-0.43000
P_{12}	-0.43	-0.43000
P_{13}	-0.01	-0.01000
P_{14}	-0.01	-0.01000
P_{15}	-0.01	-0.01000
P_{16}	-1.31	-1.31000
P_{17}	-0.03	-0.03000
P_{18}	2.80	2.86677
P_{19}	1.00	1.15000
P_{20}	2.80	2.83164
P_{21}	1.10	1.12126
P_{22}	-0.56	-0.58739
P_{23}	-0.04	-0.04600
P_{24}	-0.05	-0.04436
P_{25}	0.63	0.64753
$\|V_{18}\|$	1.07	1.06814
$\|V_{19}\|$	1.05	1.05055
$\|V_{20}\|$	1.00	1.00684
$\|V_{21}\|$	1.02	1.02583
$\|V_{22}\|$	0.89	0.88974
$\|V_{23}\|$	1.00	0.99768
$\|V_{24}\|$	1.00	1.00721
$\|V_{25}\|$	1.00	1.00635
Objective Function	0.0	0.24481
Number of Function Evaluations	9	
Execution Time on Cyber 170/815 (in seconds)	79.5	

6.5 CONCLUSIONS

In this chapter the load shedding and generation rescheduling problem is formulated as a nonlinear optimization problem. The problem of selecting a possibly small subset of loads to be shed is addressed directly by proposing the ℓ_1 type of the objective function for the problem. The Han-Powell algorithm for general nonlinear programming problem is used. To preserve the nonlinearity of the problem (appearing in the security constraints considered) the exact first-order sensitivities based on the Tellegen theorem are employed. Numerical results for 6-bus and 26-bus systems show that the objective function proposed selects small number of loads to be shed (one out of three for a 6-bus system and one out of seventeen for a 26-bus system) from the set of all loads being candidates for shedding.

The Han-Powell algorithm has proved to be fast and robust for small optimum power flow problems (up to 100 buses) (Giras and Talukdar 1981). For large problems a class of decompositions has been developed (Talukdar, Giras and Kalyan 1983) which can be applied to the load shedding problem.

7

OPTIMAL DESIGN OF MICROWAVE MULTIPLEXING NETWORKS

7.1 INTRODUCTION

Practical design and manufacture of contiguous-band multiplexers consisting of multi-cavity filters distributed along a waveguide manifold has been a problem of significant interest over the past ten years (Atia 1974; Chen, Assal and Mahle 1976; Chen 1983). Recently, a general multiplexer design procedure using an extension of the normal least squares method has been described (Egri, Williams and Atia 1983).

In this chapter, we formulate the design of a contiguous-band multiplexer structure as a centering problem using a recently developed minimax algorithm of Hald and Madsen (Hald and Madsen 1981). All design parameters of interest, e.g., waveguide spacings, input-output and filter coupling parameters, can be directly optimized. A wide range of possible multiplexer optimization problems can be formulated and solved by appropriately defining specifications on common port return loss and individual channel insertion loss functions. The minimax error functions are created using those specifications, simulated exact multiplexer responses and weighting factors. A typical structure under consideration is shown in Fig. 7.1.

The multiplexer optimization procedure to be described in this chapter exploits exact network sensitivities. Evaluation of the exact sensitivities for the multiplexer structure is based on the exact sensitivity analysis of individual filters and a direct application of the method of forward and reverse analyses for cascaded structures developed by Bandler, Rizk and Abdel-Malek (1978). The details of the

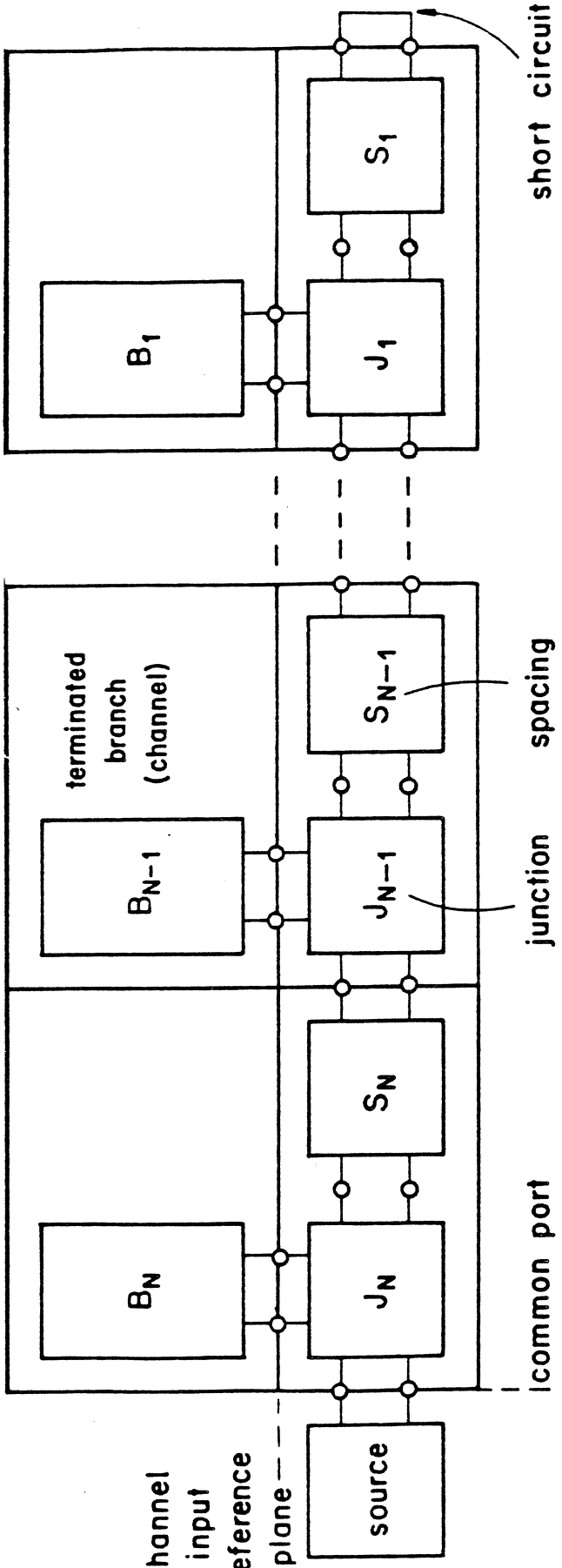


Fig. 7.1 The multiplexer configuration under consideration. $J_1, J_2, \ldots, J_N$ are arbitrarily defined 3-port junctions, $B_1, B_2, \ldots, B_N$ are terminated branches or channels which may each be represented in reduced cascade form and $S_1, S_2, \ldots, S_N$ are usually waveguide spacing elements.

simulation and sensitivity analysis aspect of the problem have been described in Bandler, Chen, Daijavad and Kellermann (1984), Bandler, Daijavad and Zhang (1985), and Bandler, Chen and Daijavad (1985).

7.2 FORMULATION OF THE PROBLEM

The objective function to be minimized is given by

$$F(\boldsymbol{\phi}) = \max_{j \in J} f_j(\boldsymbol{\phi}), \tag{7.1}$$

where $\boldsymbol{\phi} = \boldsymbol{\phi}^0$ is a vector of design parameters (e.g., section or spacing lengths, channel input and output couplings and filter coupling parameters) and $J \triangleq \{1, 2, \ldots, m\}$ is an index set. The minimax functions $f_j(\boldsymbol{\phi})$, $j \in J$, can be of the form

$$w^1_{Uk}(\omega_i)(F^1_k(\boldsymbol{\phi}, \omega_i) - S^1_{Uk}(\omega_i)), \tag{7.2}$$

$$-w^1_{Lk}(\omega_i)(F^1_k(\boldsymbol{\phi}, \omega_i) - S^1_{Lk}(\omega_i)), \tag{7.3}$$

$$w^2_U(\omega_i)(F^2(\boldsymbol{\phi}, \omega_i) - S^2_U(\omega_i)), \tag{7.4}$$

$$-w^2_L(\omega_i)(F^2(\boldsymbol{\phi}, \omega_i) - S^2_L(\omega_i)), \tag{7.5}$$

where $F_k{}^1(\boldsymbol{\phi}, \omega_i)$ is the insertion loss for the kth channel at the ith frequency, $F^2(\boldsymbol{\phi}, \omega_i)$ is the return loss at the common port at the ith frequency, $S_{Uk}{}^1(\omega_i)$ ($S_{Lk}{}^1(\omega_i)$) is the upper (lower) specification on insertion loss of the kth channel at the ith frequency, $S_U{}^2(\omega_i)$ ($S_L{}^2(\omega_i)$) is the upper (lower) specification on return loss at the ith frequency, and $w_{U_k}{}^1$, $w_{L_k}{}^1$, $w_U{}^2$, $w_L{}^2$ are the arbitrary user-chosen nonnegative weighting factors.

A typical example of specifications on return loss and insertion loss for a three-channel multiplexer is shown in Fig. 7.2.

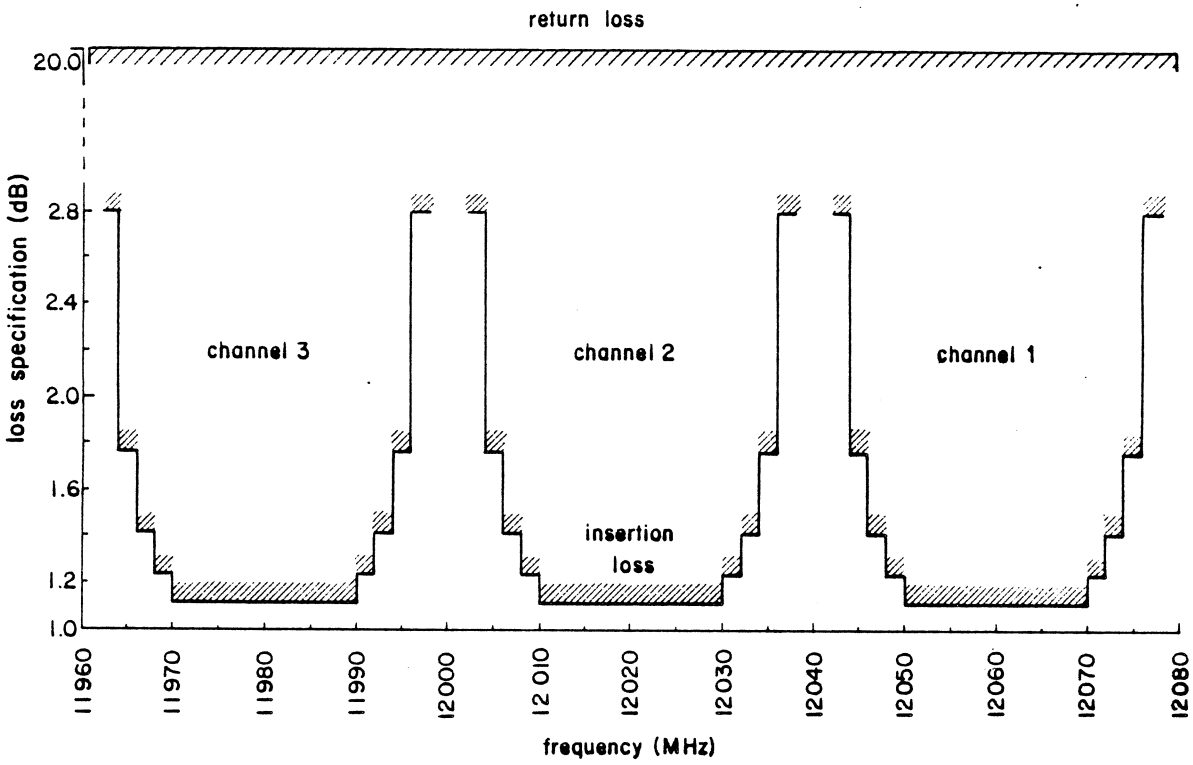


Fig. 7.2 Illustration of specifications for a three-channel multiplexer.

7.3 UNTERMINATED FILTER SIMULATION AND SENSITIVITY ANALYSIS

The model of an unterminated multi-cavity filter has been given by Atia and Williams (1972) (see Section 4.8.2, equations (4.55-4.57) and Fig. 4.4).

We reduce the system to a two-port (see Fig. 7.3) given by

$$\begin{bmatrix} I_1 \\ I_n \end{bmatrix} = \mathbf{y} \begin{bmatrix} V_1 \\ V_n \end{bmatrix}, \tag{7.6}$$

where $\mathbf{y}$ is the s.c. admittance matrix of the filter, including input $1{:}n_1$ and output $n_2{:}1$ ideal transformers (see Fig. 7.4). Matrix $\mathbf{y}$ and its sensitivities w.r.t. all variables, including frequency, can be obtained by solving the systems

$$\mathbf{Zp} = \mathbf{e}_1 \quad \text{and} \quad \mathbf{Zq} = \mathbf{e}_n, \tag{7.7}$$

where $\mathbf{e}_1 \triangleq [1\ 0\ \ldots\ 0]^T$ and $\mathbf{e}_n \triangleq [0\ \ldots\ 0\ 1]^T$ are n-dimensional unit excitation vectors. Note that $\mathbf{q}$ can be found with minimal extra effort after factorization of $\mathbf{Z}$ for the solution of $\mathbf{p}$. The following formulas for the evaluation of $\mathbf{y}$ and its sensitivities are readily derived (Bandler, Chen, Daijavad and Kellermann 1984)

$$\mathbf{y} = \begin{bmatrix} n_1^2 y_{11} & n_1 n_2 y_{12} \\ n_1 n_2 y_{12} & n_2^2 y_{22} \end{bmatrix} = \begin{bmatrix} n_1^2 p_1 & n_1 n_2 p_n \\ n_1 n_2 p_n & n_2^2 q_n \end{bmatrix}, \tag{7.8}$$

$$\frac{\partial \mathbf{y}}{\partial M_{\ell k}} = -jc \begin{bmatrix} 2 n_1^2 p_\ell p_k & n_1 n_2 (p_\ell q_k + p_k q_\ell) \\ n_1 n_2 (p_\ell q_k + p_k q_\ell) & 2 n_2^2 q_\ell q_k \end{bmatrix}, \tag{7.9}$$

where

$$c = \begin{cases} 1 & \text{if } \ell \neq k, \\ 0.5 & \text{if } \ell = k, \end{cases}$$

$$\frac{\partial \mathbf{y}}{\partial \omega} = -j \frac{\partial s}{\partial \omega} \begin{bmatrix} n_1^2 \mathbf{p}^T \mathbf{p} & n_1 n_2 \mathbf{p}^T \mathbf{q} \\ n_1 n_2 \mathbf{p}^T \mathbf{q} & n_2^2 \mathbf{q}^T \mathbf{q} \end{bmatrix}, \tag{7.10}$$

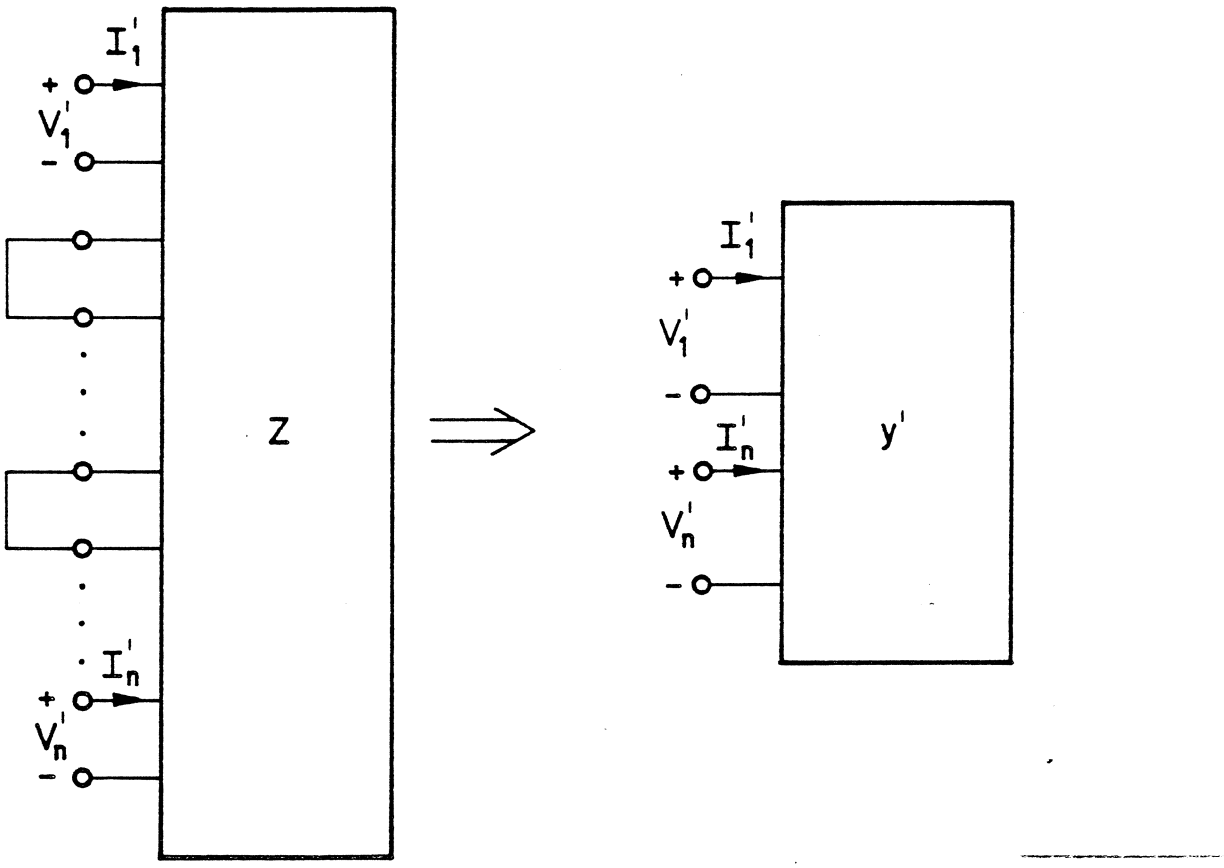


Fig 7.3 Reduction of the filter to a two-port.

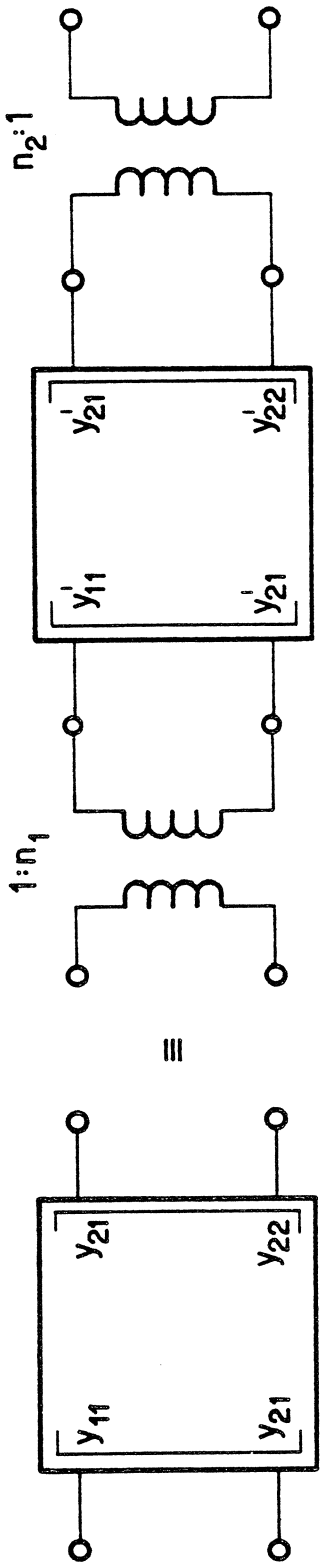


Fig. 7.4 Short circuit admittance matrix of the filter including input $1:n_1$ and output $n_2:1$ ideal transformers.

$$\frac{\partial \mathbf{y}}{\partial n_1} = \begin{bmatrix} 2n_1^2 p_1 & n_1 p_n \\ n_2 p_n & 0 \end{bmatrix} \quad \text{and} \quad \frac{\partial \mathbf{y}}{\partial n_2} = \begin{bmatrix} 0 & n_1 p_n \\ n_1 p_n & 2n_2 q_n \end{bmatrix}. \tag{7.11}$$

7.4 COMPUTER IMPLEMENTATION

A Fortran package has been developed for multiplexer simulation, sensitivity analysis and optimization. Functional blocks of the package are shown in Fig. 7.5. This package has been designed to reflect the requirements of ComDev Ltd. of Cambridge, Ontario, Canada. It has been tested in close cooperation with engineers directly involved in multiplexer design and postproduction tuning.

Options of the Package

The required mode of operation of the package is selected by the user by setting an indicator as follows:

1 – if only multiplexer simulation is required;

2 – if multiplexer sensitivity analysis is required (implies simulation);

3 – if multiplexer optimization is required (implies both simulation and sensitivity analysis).

Options of the Optimization Mode

If the multiplexer optimization option is selected three modes of optimization are allowed for, namely, only return loss optimization (suggested by Chen (1983)), only insertion loss optimization, return loss and insertion loss optimization, all at user-defined sets of frequency points. A suitable and sophisticated coding scheme has been developed which creates a consecutively numbered set of

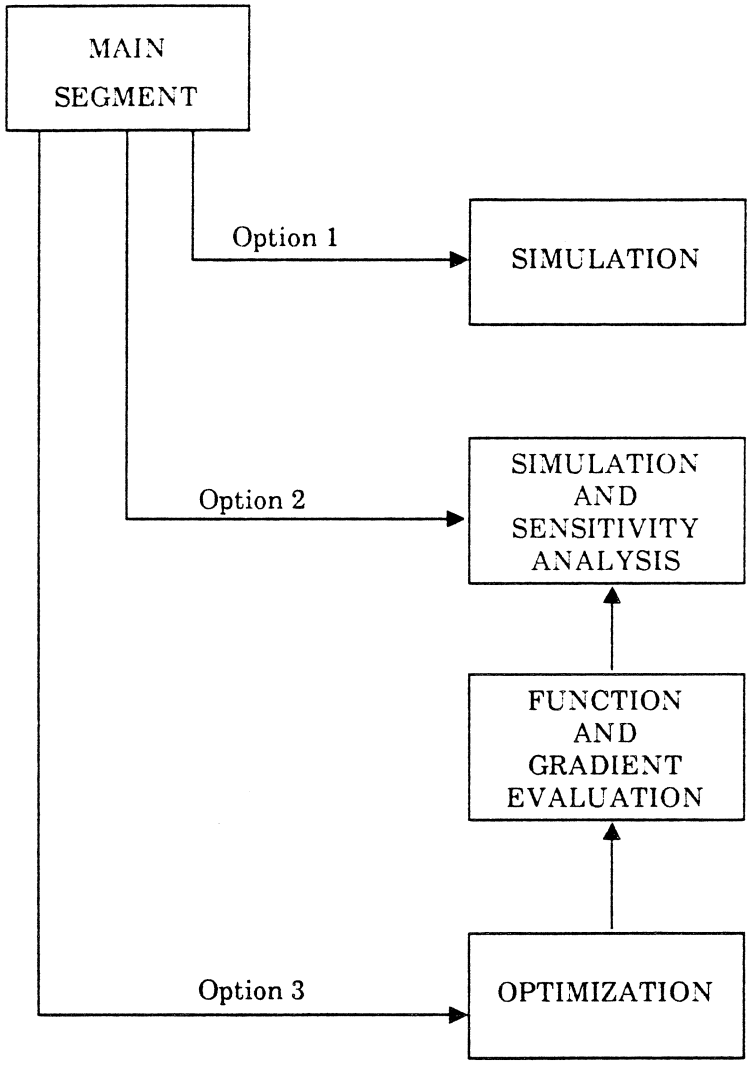


Fig. 7.5 Functional blocks of the computer package for multiplexer simulation, sensitivity analysis and optimization.

minimax functions depending on whether we have only lower (upper) specifications, both or no specifications on a function of interest at a certain frequency point.

Options Related to the Selection of Optimization Variables

The coding scheme developed and employed in the package allows also a very flexible choice of optimization variables. In general, all parameters are candidates for optimization variables, however, with very little effort, the user can declare any parameters to be optimization variables.

Options Related to the Microwave Model of a Multiplexer

The package can exploit three commonly used practical models of the multiplexer, depending on whether the junctions are ideal or nonideal (junction susceptance is included), whether the filters are lossless or lossy (dissipation is included), and whether the filters are modelled as dispersive or non-dispersive. (The waveguide manifold is always assumed dispersive.)

7.5 5-CHANNEL, 11 GHz MULTIPLEXER DESIGN USING MINIMAX OPTIMIZATION

The procedure is illustrated by designing an 11 GHz, 5-channel multiplexer having the center frequencies and bandwidths (similar to those in Egri, Williams and Atia 1983) given in Table 7.1.

TABLE 7.1

5-CHANNEL MULTIPLEXER CENTER FREQUENCIES AND BANDWIDTHS

Channel	Center Frequency (MHz)	Bandwidth (MHz)
1	10992.5	81
2	11075.0	76
3	11155.0	76
4	11495.0	76
5	11618.5	154

Suppose we want to design this multiplexer such that certain specifications on the common port return loss and individual channel insertion loss functions are satisfied. A lower specification of 20 dB on return loss over the passbands of all five channels should be satisfied. We want also to control return loss between channels 1 and 2, 2 and 3, 4 and 5 in a similar way. We impose also additional specifications on insertion loss for all channels, i.e., we want the insertion loss in the transition bands not to drop below 20 dB. We start the design process with five identical six pole, pseudo-elliptic function filters. Starting values of the coupling coefficients for the filters are given in the following matrix (Chen 1983):

$$\mathbf{M} = \begin{bmatrix} 0 & 0.62575 & 0 & 0 & 0 & 0 \\ 0.62575 & 0 & 0.57615 & 0 & 0 & 0 \\ 0 & 0.57615 & 0 & 0.32348 & 0 & -0.74957 \\ 0 & 0 & 0.32348 & 0 & 1.04102 & 0 \\ 0 & 0 & 0 & 1.04102 & 0 & 1.04239 \\ 0 & 0 & -0.74957 & 0 & 1.04239 & 0 \end{bmatrix} \quad (7.12)$$

The initial spacing lengths are set equal to $\lambda_{g_k}/2$ (half the wavelength corresponding to the kth center frequency). The filters are assumed lossy and dispersive. Waveguide junctions are assumed nonideal.

Fig. 7.6 shows the responses of the multiplexer at the start of the optimization process. As we see the specifications on the common port return loss are seriously violated.

The optimization process is performed in several steps. First we select only non-zero couplings, input/output transformer ratios and filter spacings as optimization variables. This gives a total of 45 optimization variables. The error

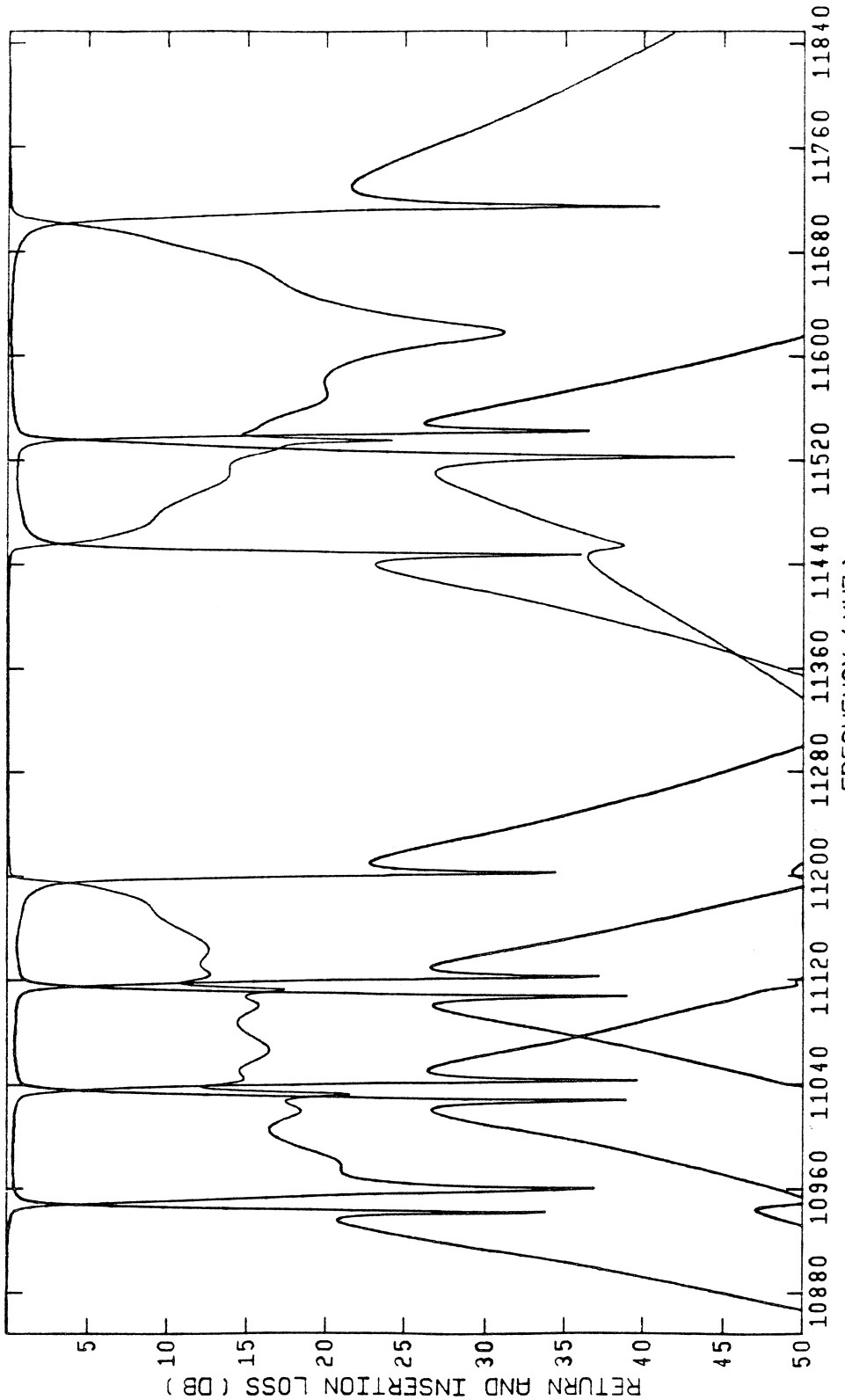


Fig. 7.6 Responses of the 5-channel, 11 GHz multiplexer at the starting point of the optimization process.

functions resulting from the multiplexer responses and specifications are created at 51 nonuniformly spaced frequency points. An improved design is obtained after 30 function evaluations (230s on the Cyber 170/815). The responses corresponding to the first step of the optimization process are shown in Fig. 7.7.

In order to completely satisfy the design specifications we perform a second step of optimization in which we release additional optimization variables, i.e., cavity resonances. This gives a total of 75 nonlinear optimization variables. Using the same frequency points as in step 1 and results of the first optimization as a starting point we continue the optimization process. After 30 additional function evaluations (and 470s of CPU time on the Cyber 170/815), the design specifications are satisfied and the optimized responses of the 5-channel multiplexer are shown in Fig. 7.8. To improve the return loss response of the multiplexer, the third step of optimization is performed in which a search technique for maxima of the response is employed. This gives 66 minimax functions and the same number of variables as previously. After 25 additional function evaluations (and 360s of CPU time on the Cyber 170/815) we obtain the final optimized responses as shown in Fig. 7.9.

In the approach presented the emphasis is on achieving a maximally effective set of early iterations of optimization using a subset of all possible optimization variables. This subset should correspond to "dominant" variables of the problem. Initial selection of the variables can be facilitated by the full knowledge and experience of the designer and by an initial sensitivity analysis at selected frequency points.

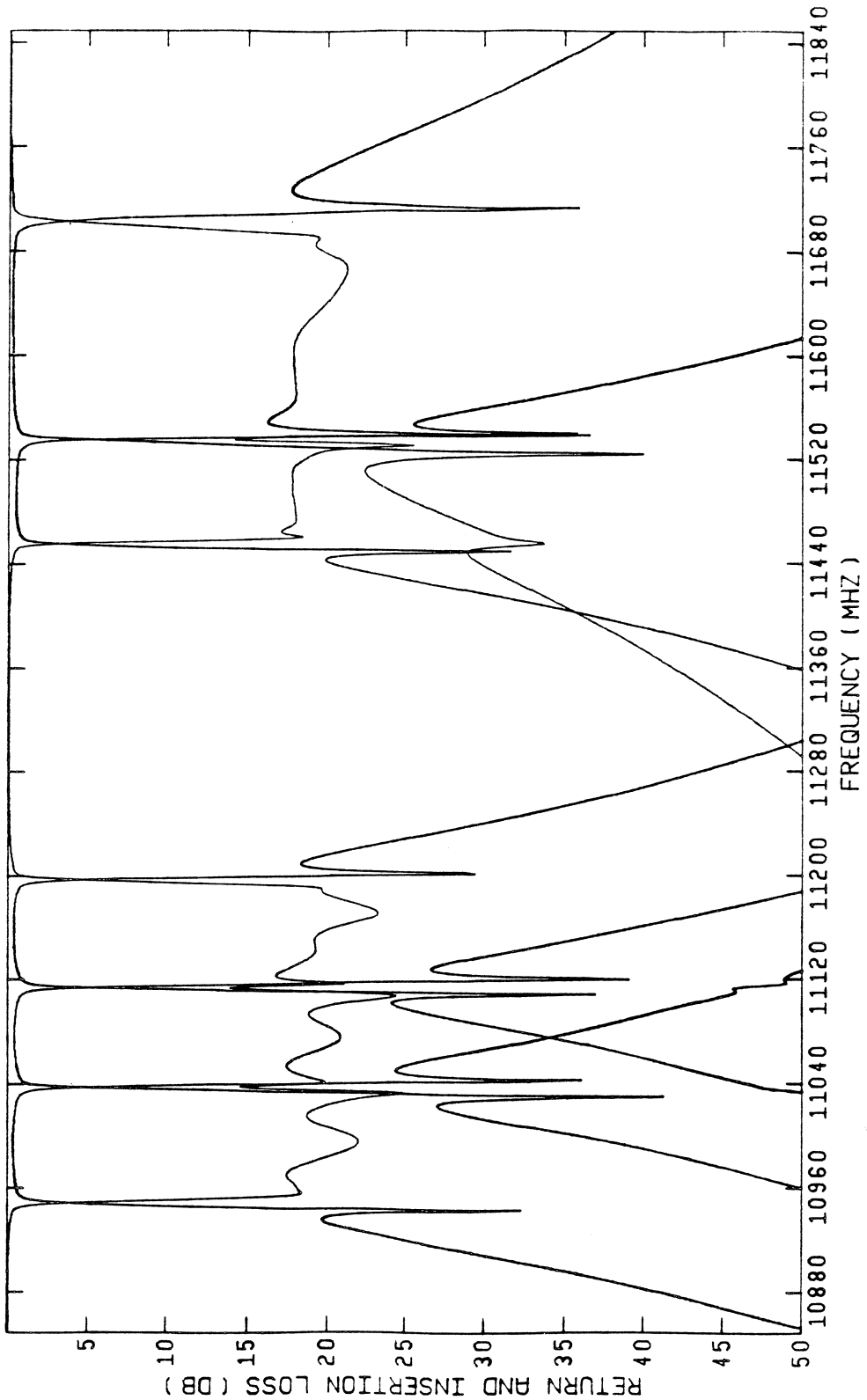


Fig. 7.7 Responses of the 5-channel, 11 GHz multiplexer after the first 30 iterations using 45 optimization variables and 51 nonuniformly spaced sample points.

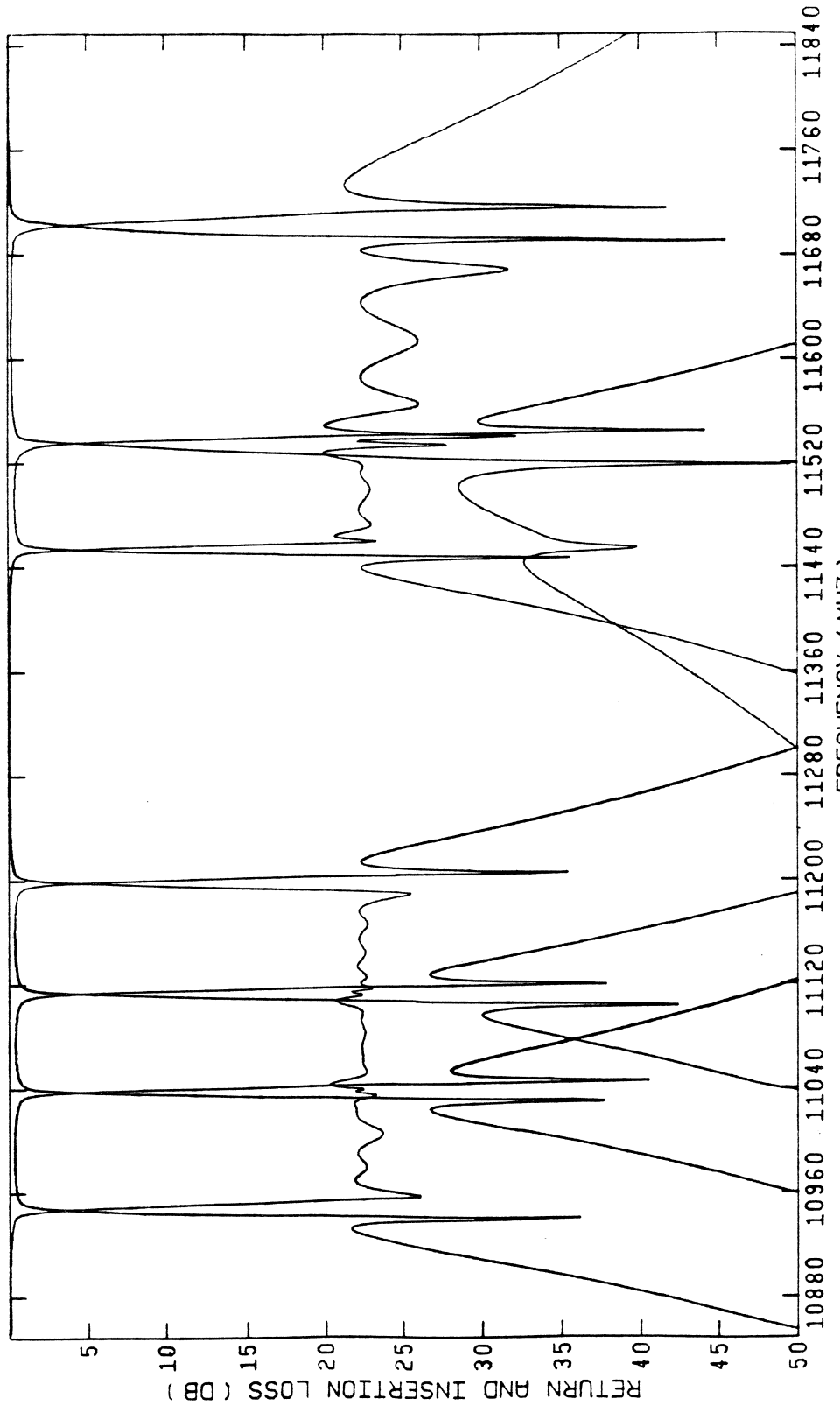


Fig. 7.8 Responses of the 5-channel, multiplexer after 30 additional iterations using 75 optimization variables and 51 minimax functions.

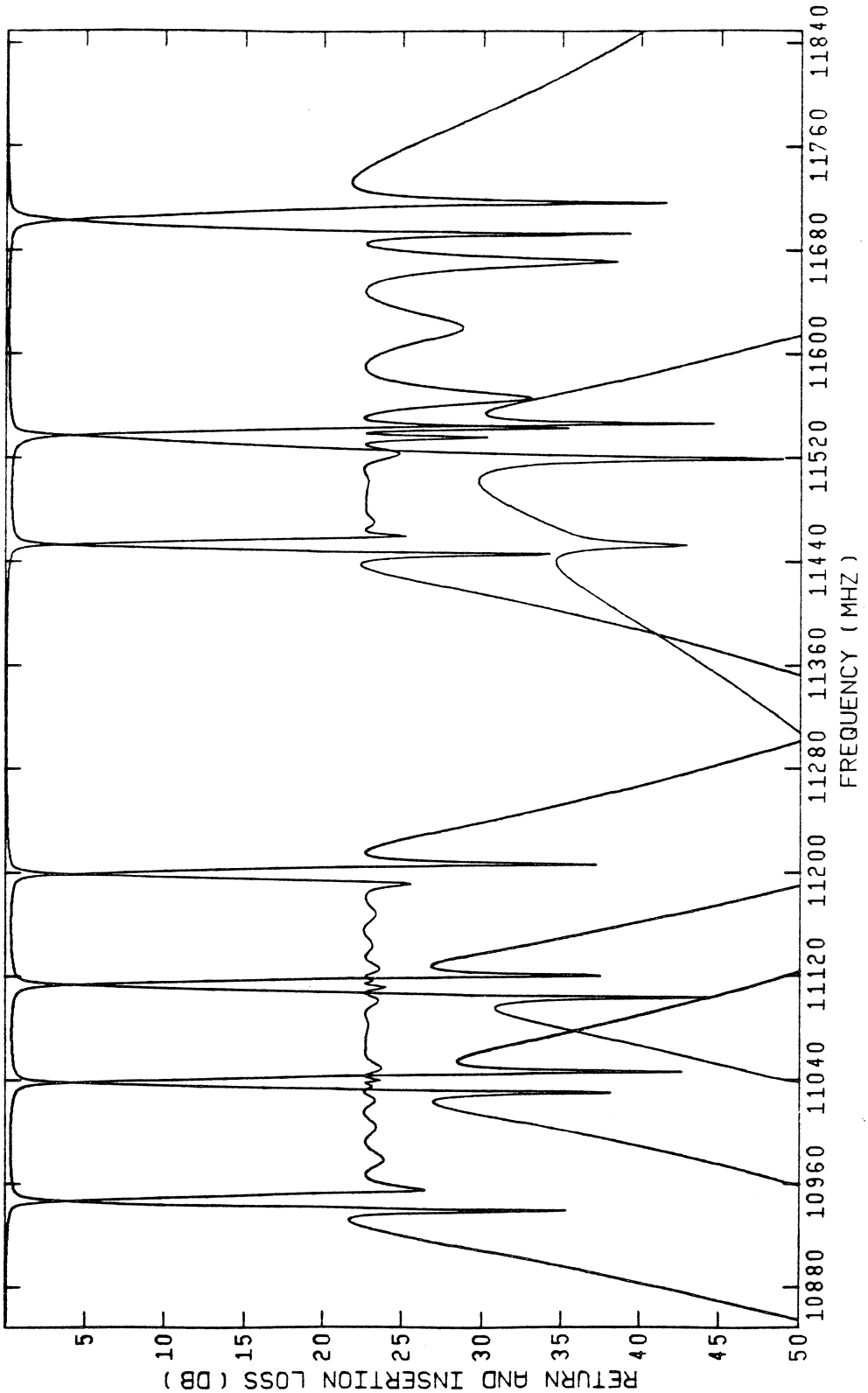


Fig. 7.9 The final optimized responses of the 5-channel multiplexer obtained with 75 variables and 66 minimax functions. A search technique for the peaks of the return loss response has been employed.

7.6 12-CHANNEL, 12 GHz MULTIPLEXER DESIGN USING ℓ_1 OPTIMIZATION

A major task in designing a multiplexer is to determine the location of the channel filters along the waveguide manifold (Chen 1983). This is very important for designs using the common port return loss as the only optimization criterion. A typical value of lower specification on return loss over the passbands of all multiplexer channels is 20 dB. The error functions $f_j(\boldsymbol{\Phi})$ for this type of problem are of the form (7.5).

If we perform a minimax optimization based on these error functions and at the solution the minimax objective function value is negative then the goal has been achieved. In many cases, however, using the filter spacings as the only optimization variables may not be sufficient to satisfy all specifications and minimax optimization gives results corresponding to the situation where the specification violations are distributed over all multiplexer channels. In that case the use of the one-sided ℓ_1 optimization of the same error functions may lead to more desirable results where the violations occur only over a few multiplexer channels. This process of identifying "bad channels" has two very important consequences. First, the results indicate in which channels the additional variables have to be released to improve locally (in the frequency domain) the performance of the multiplexer and second, it gives very good starting values of the waveguide spacings to be used in the subsequent minimax optimization. The idea presented is illustrated by designing a 12 GHz, 12-channel multiplexer without dummy channels. The 12-channel contiguous band multiplexer has a channel frequency separation of 40 MHz and a usable bandwidth of 39 MHz with the center frequency of channel 1 12 180.0 MHz.

Suppose we want to design this multiplexer such that a lower specification of 20 dB on the common port return loss over the passbands of all 12 channels should be satisfied. We start the design process with twelve identical 6th order filters with the coupling coefficients given in the following matrix (Tong and Smith 1984):

$$\mathbf{M} = \begin{bmatrix} 0 & 0.594 & 0 & 0 & 0 & 0 \\ 0.594 & 0 & 0.535 & 0 & 0 & 0 \\ 0 & 0.535 & 0 & 0.425 & 0 & -0.400 \\ 0 & 0 & 0.425 & 0 & 0.834 & 0 \\ 0 & 0 & 0 & 0.834 & 0 & 0.763 \\ 0 & 0 & -0.400 & 0 & 0.763 & 0 \end{bmatrix} \tag{7.13}$$

Initially we select the spacing lengths along the waveguide manifold as the only optimization variables with starting values set equal to $\lambda_{gk}/2$ (half the wavelength corresponding to the kth center frequency). For the kth channel the waveguide spacing is measured along the manifold from the adjacent (k–1)th channel. For the first channel the spacing is the distance from the short circuit. The filters are assumed lossy and dispersive. Waveguide junctions are assumed nonideal.

Fig. 7.10 shows the return loss response of the multiplexer at the start of the optimization process. The specification on the common port return loss is seriously violated, especially in the lower frequencies range (corresponding to channels 8-12).

The filter spacings are the dominant variables of the problem. This is based on the initial sensitivity analysis of the common port return loss function w.r.t. all variables at selected frequency points.

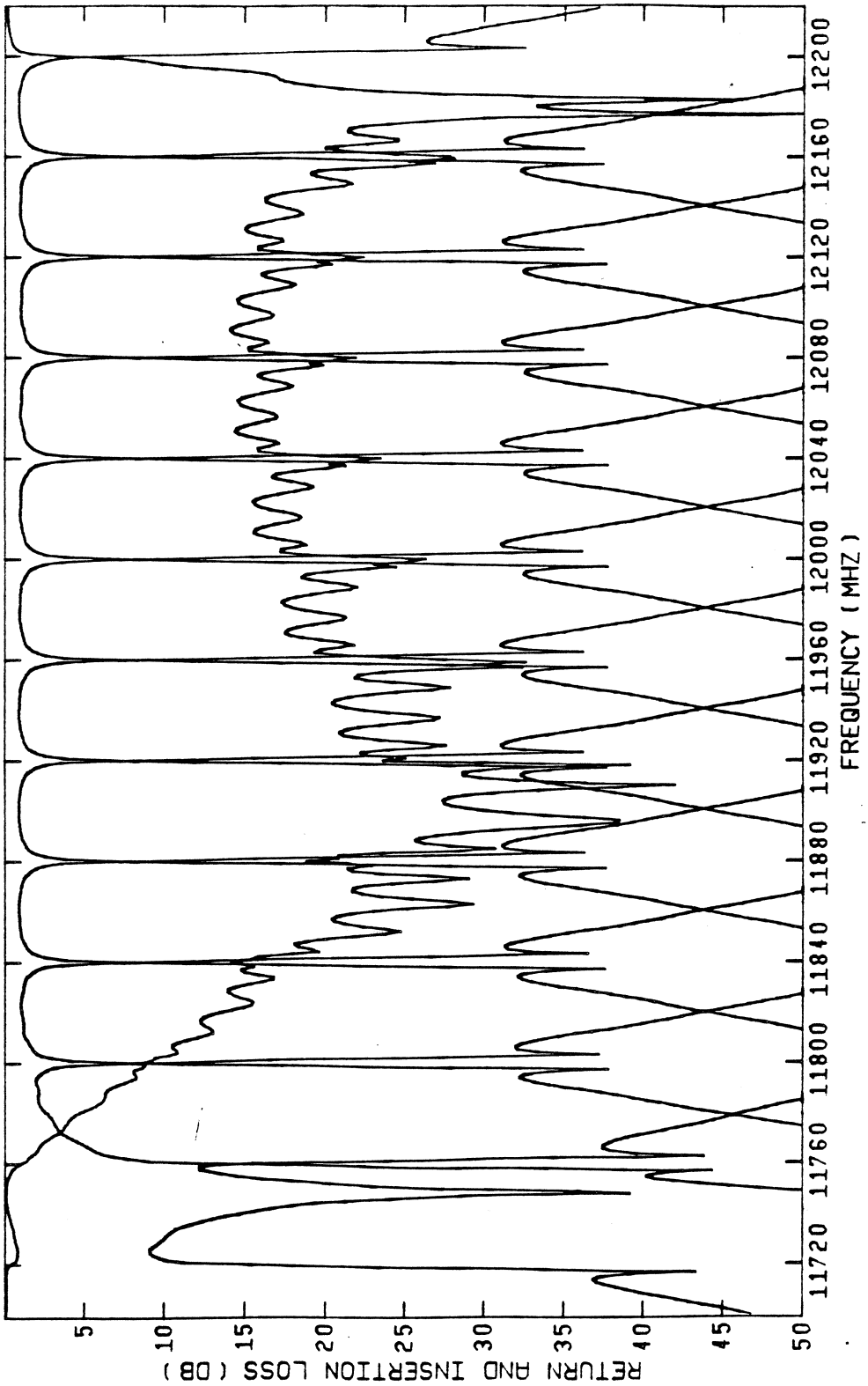


Fig. 7.10 Responses of the 12-channel multiplexer at the start of the optimization process.

We perform the one-sided ℓ_1 optimization which is defined in the following way

$$\underset{\boldsymbol{\Phi}}{\text{minimize}} \sum_{i=1}^{m} |f_i^*(\boldsymbol{\Phi})|, \tag{7.14}$$

where

$$f_i^* \triangleq \begin{cases} f_i(\boldsymbol{\Phi}) & \text{if } f_i(\boldsymbol{\Phi}) \geq 0 \\ 0 & \text{if } f_i(\boldsymbol{\Phi}) < 0 \end{cases} \tag{7.15}$$

The functions $f_i(\boldsymbol{\Phi})$ are the original error functions defined in (7.5).

We define also the gradients of the functions $f_i^*(\boldsymbol{\Phi})$ in the following way

$$\frac{\partial f_i^*(\boldsymbol{\Phi})}{\partial \boldsymbol{\Phi}} \triangleq \begin{cases} \dfrac{\partial f_i(\boldsymbol{\Phi})}{\partial \boldsymbol{\Phi}} & \text{if } f_i(\boldsymbol{\Phi}) \geq 0 \\ 0 & \text{if } f_i(\boldsymbol{\Phi}) < 0 \end{cases} \tag{7.16}$$

The results of the ℓ_1 optimization defined above are shown in Fig. 7.11. The violations of the 20 dB specification are most serious in the frequency range corresponding to channels 1-2 and 8-12. This motivates us to release additional optimization variables in the filters corresponding to these channels. As additional optimization variables we release the input-output transformer ratios, cavity resonant frequencies as well as intercavity couplings. From that point minimax optimization is employed using the ℓ_1 optimized spacings as the starting values for the spacings. The final optimized return loss of the 12 channel multiplexer is shown in Fig. 7.12. The problem involves 60 nonlinear design variables.

7.7 CONCLUDING REMARKS

A powerful and efficient optimization procedure for contiguous band multiplexers has been presented. It employs a fast and robust gradient-based minimax algorithm. The multiplexer responses and their first-order sensitivities are

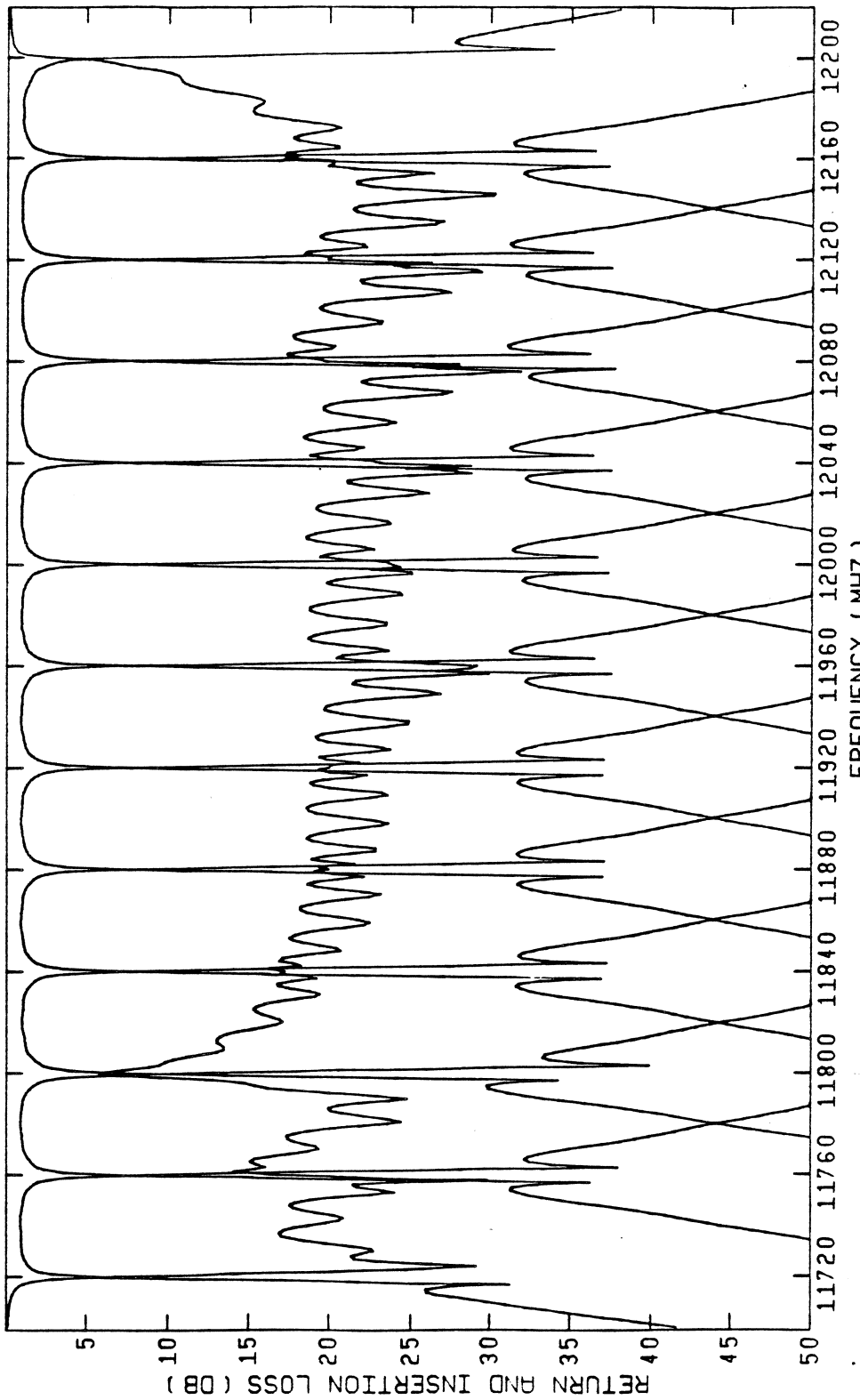


Fig. 7.11 Responses of the 12-channel multiplexer with optimized spacings only using ℓ_1 optimization.

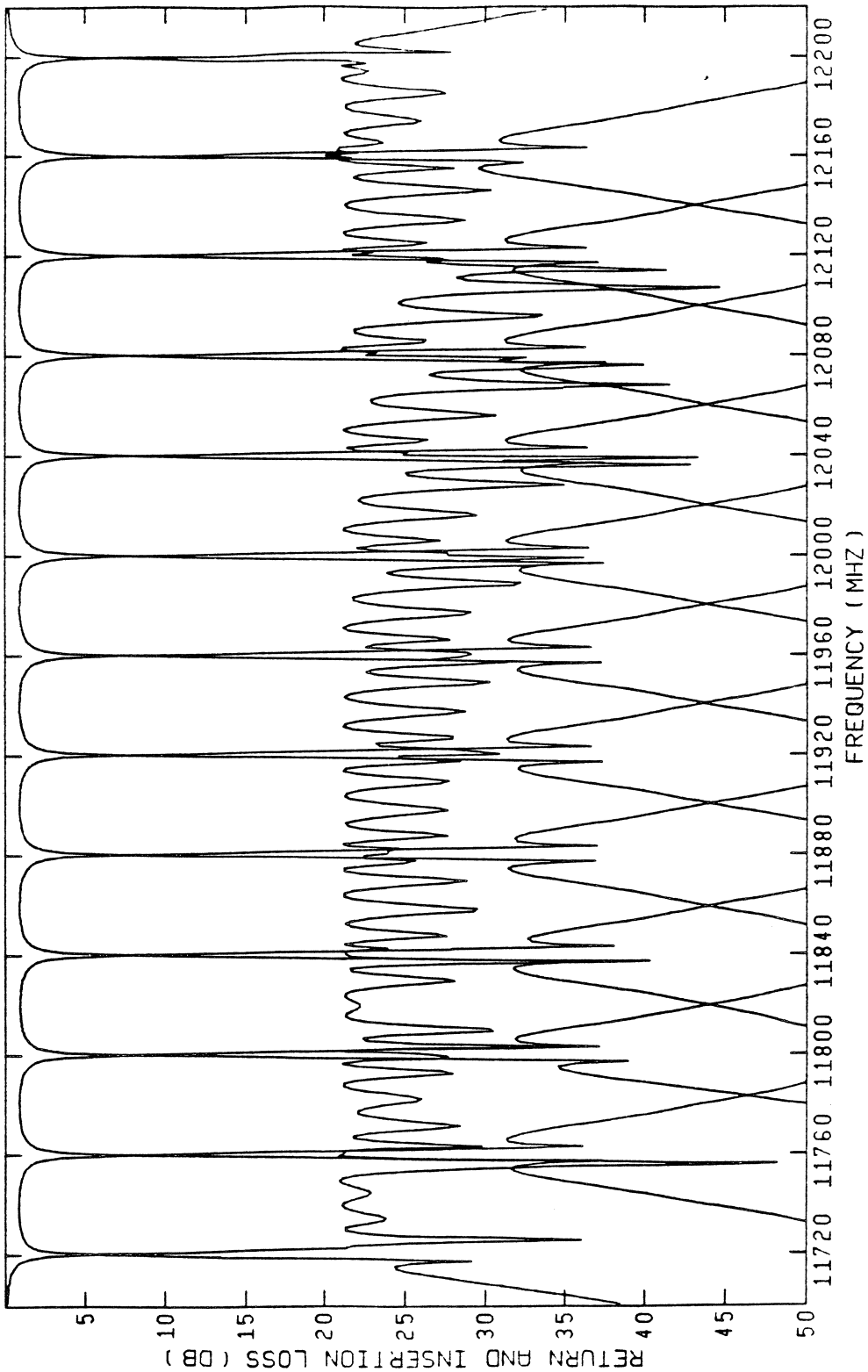


Fig. 7.12 Responses of the 12-channel multiplexer with optimized spacings, input-output transformer ratios, cavity resonances and coupling parameters using minimax optimization.

calculated efficiently and exactly. The procedure developed allows flexibility in selecting optimization variables and multiplexer models. The important feature is the possibility of including linear equality and inequality constraints on optimization variables. To our knowledge, the 5- and 12-channel multiplexer designs are the first successful attempts to use gradient-based optimization for multiplexer design as well as being the largest nonlinear optimization process ever demonstrated on microwave circuit design for a reasonable computational cost.

A formulation using the ℓ_1 norm for the initial stage of multiplexer design has been presented and illustrated by a 12-channel, 12 GHz multiplexer problem. The one-sided ℓ_1 optimization sets to zero as many error functions as possible, and this results in identifying channels of the multiplexer where the specification violations are most serious.

8

CONCLUSIONS

This thesis has considered a number of important problems associated with computer-aided design and computer-aided testing of engineering systems. These problems include design centering, tolerancing and tuning, worst-case selection, tunable parameter selection, fault location and model parameter identification from measurements.

Many of the engineering system problems discussed in this thesis have been formulated as optimization problems. The formulations exploit characteristic features of the minimax and ℓ_1 norms.

The minimax objective function has been well established in the design of circuits and systems, especially in design centering and tolerance optimization. Two new applications have been presented in this thesis. One is an algorithm for the fixed tolerance problem embodying worst-case search and selection of sample points. Linearly constrained optimization is used to determine actual worst cases during the optimization of the nominal design subject to fixed tolerances.

Another application is an algorithm for minimizing the cardinality of a set subject to nonlinear, nondifferentiable constraints. It is a combinatorial problem in which an unconstrained minimax optimization of a set of error functions is used to select candidates for deleting from the original set of points. The algorithm has been used for solving a practical mechanical engineering problem which originated from aligning mechanical designs. The approach presented in Chapter 5 should prove useful in many other areas where problems of a similar nature may exist.

Unlike the minimax norm, the ℓ_1 norm has not been given its proper place in computer-aided design and computer-aided testing of circuits and systems. This thesis, therefore, has examined some aspects of ℓ_1 optimization and its relevance to tunable parameter selection at the design stage, fault location in analog systems and model parameter identification from measurements.

Necessary conditions for optimality of the nonlinear ℓ_1 problem with nonlinear constraints indicate that zeros of the nonlinear functions play an important role in the characteristics of the ℓ_1 objective function. Active constraints play the same role as zero functions. This fact has been used in fault isolation techniques for linear analog circuits. The ℓ_1 norm is used to isolate the most likely faulty elements. In Chapter 4, a new formulation for fault isolation in analog circuits based on an exact penalty function has been presented.

Another important application of the ℓ_1 norm is the selection of tunable parameters in the functional approach to post-production tuning. Two new formulations for the tuning problem have been presented based on the classification of the tuning problems into tuning with the fixed set of tunable parameters and tuning with the variable set of tunable parameters. A mixed programming formulation given in Chapter 4 ensures that the solution gives the minimum number of tunable parameters. The importance of selecting tuning parameters in real life engineering systems has been demonstrated in Chapter 6 where the load shedding and generation rescheduling problem in power systems is formulated as an optimization problem with the ℓ_1 type of objective function. Although the number of applications of the ℓ_1 norm to circuits and systems problems is increasing, it is felt that the full potential of the ℓ_1 norm in solving engineering system problems has not been realized and requires further research.

All the formulations in this thesis are supported by fast and efficient algorithms for minimax and ℓ_1 optimization originated by Hald and Madsen. The results presented in Chapter 3 (comparison of minimax algorithms on a three- section transmission-line transformer) and in Hald and Madsen (1985) indicate that both algorithms may be the best of their class currently available. The robustness of the algorithms has been proved by solving practical engineering design problems of significant size. An optimization procedure for the design of contiguous and noncontiguous band microwave multiplexers, described in Chapter 7, allows us to solve problems involving up to 75 nonlinear design variables and as many as 100 nonlinear error functions. An important factor in the success of both algorithms in solving difficult engineering problems is the fact that the algorithms recognize singularities in the problems defined and treat them efficiently. As was shown in Chapter 3 on a two-section transmission-line transformer, those singularities are not only abstract concepts, but are inherent in many circuits and systems design problems.

Many of the subproblems associated with overall problems in different engineering disciplines are of a similar nature and sufficient complexity to motivate the development of a conceptual framework within which it would be possible to integrate and apply the results of computer-aided engineering system research. The aim would be to provide the means of integrating the design process (centering, tolerancing and tuning) with the post-production process (parameter identification, fault location, post-production tuning). Such a framework would be an extension of the concepts and definitions used in DCTT to include fault location, post-production tuning and model parameter identification.

In order to integrate the classical design problem and, for example, the model parameter identification problem, we could extend the concept of specification to include a measured response of a network or system. In the design problem, the specifications which a manufactured design must meet are known exactly. This is not the case with the specifications for the identification problem due to the presence of measurement errors or uncertainties in observing the particular response. The concept of performance function in both problems is exactly the same. This is an appropriate model which depends, in general, nonlinearly on a set of parameters. The difference between the two problems lies basically in the nature of the specifications. It should also be noted that for the identification problem, we usually consider the single point specification while in the design problem we usually have lower and/or upper specifications. Single point specifications can always be replaced by two specifications, lower and upper, of equal value. Thus we can think of a general design problem formulated as an optimization problem with the objective being the norm of the error functions resulting from specifications on performance functions of interest. The particular norm to be used depends on the nature of specifications.

A number of other problems are also worth further research and development.

(a) In the present version of the algorithm for the fixed tolerance problem, worst-case search is performed for each iteration of the optimization w.r.t. nominal values. It was observed, however, that close to a solution, worst cases selected do not change from iteration to iteration. It would be worthwhile to develop a criterion allowing us to test whether the worst-case corresponding to a particular function has changed or not. Such a criterion

would be based on the signs of gradients and would alleviate the worst-case search for each iteration.

(b) In Chapter 4, a mixed programming formulation for the selection of tunable parameters at the design stage is proposed. The formulation ensures that a minimum number of parameters for tuning is selected. It would be useful to investigate possible algorithms for solving this type of mixed programming problem.

(c) A natural extension of the approach to the best alignment is one which considers alignment problems in three dimensions. This would be very useful from the practical point of view.

(d) In the present implementation of the approach to load shedding and generation rescheduling in power systems, all arrays are assumed to be dense. Since the problem is structurally very sparse, implementations taking sparsity directly into account would prove very useful for solving problems of significant size.

APPENDIX A

DESCRIPTION OF THE HALD AND MADSEN MINIMAX ALGORITHM

The algorithm is a combination of two methods denoted Method 1 and Method 2. Method 1 is intended to be used far away from a solution whereas Method 2 is a local method. We first describe these two methods.

Method 1

This is essentially the algorithm of Madsen (1975). At the kth step a feasible approximation $\mathbf{x}_k$ of a solution of (3.9) and a local bound Λ_k are given. In order to find a better estimate of a solution the following linearized problem is solved:

$$\underset{\mathbf{h}}{\text{minimize}} \quad \overline{F}\ (\mathbf{x}_k, \mathbf{h}) \triangleq \max_j \{f_j(\mathbf{x}_k) + f_j'(\mathbf{x}_k)^T\mathbf{h}\}$$

subject to

$$\|\mathbf{h}\|_\infty \le \Lambda_k \tag{A.1}$$

$$\mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i = 0\,, \qquad i = 1, \ldots, \ell_{eq}\,,$$

$$\mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i \ge 0\,, \qquad i = (\ell_{eq}+1), \ldots, \ell\,.$$

The solution of (A.1), denoted $\mathbf{h}_k$, is found by linear programming. Notice that $\mathbf{x}_k + \mathbf{h}_k$ is feasible. The next iterate is $\mathbf{x}_k + \mathbf{h}_k$ provided this point is better than $\mathbf{x}_k$ in the sense of F, i.e., if $F(\mathbf{x}_k + \mathbf{h}_k) < F(\mathbf{x}_k)$. Otherwise $\mathbf{x}_{k+1} = \mathbf{x}_k$. In Fig. A.1, an example with one variable, two functions and no constraints ($\ell = 0$) is shown. $F(\mathbf{x})$ is the kinked bold-faced curve. At $\mathbf{x}_k$ linear approximations of the two functions f_1 and f_2 are made and the solution of (A.1) is $\mathbf{h}_k$ which is found at the intersection of the two linear approximations. We assume that the local bound Λ_k is so large that it has no

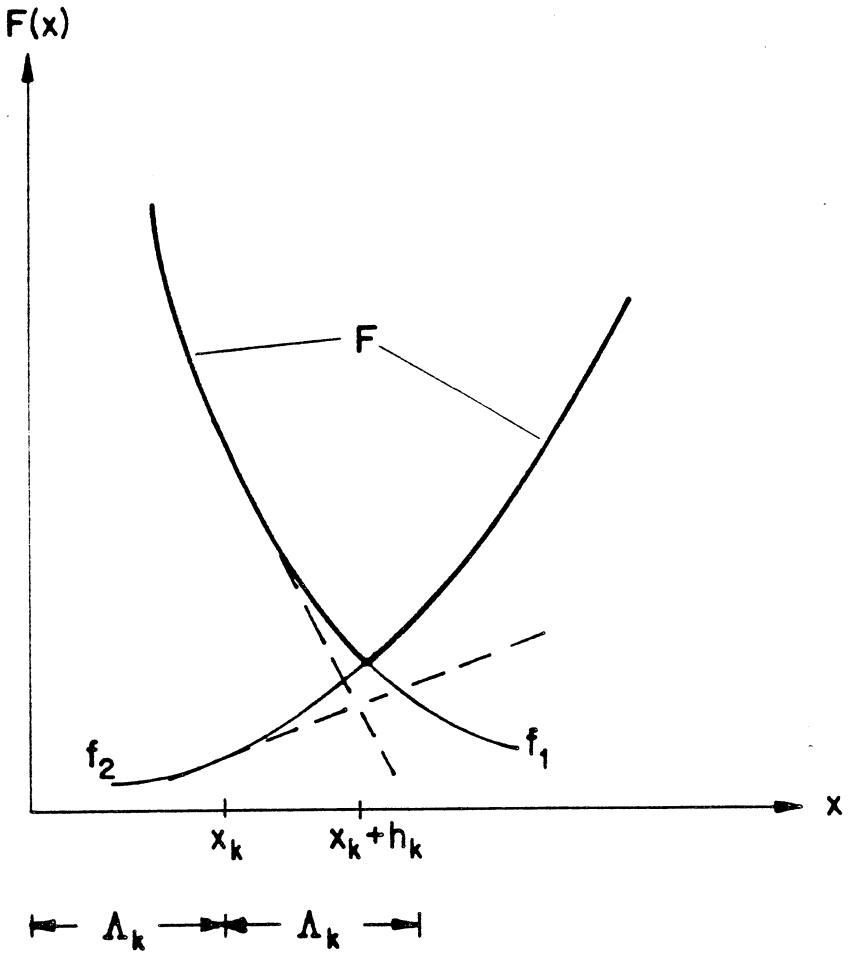


Fig. A.1 An example with one variable and two functions illustrating a Method 1 iteration of the algorithm.

influence. The new point is $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k$ which is seen to be close to a local minimum of F.

The local bound Λ_k is introduced because the linear model (A.1) is a good approximation of (3.9) only in some neighbourhood of $\mathbf{x}_k$. Therefore, it makes sense to consider only small values of $\| \mathbf{h} \|$ in connection with the linear model (A.1). The size of the bound is adjusted in every iteration based on a comparison between the decrease in the objective function F and the decrease predicted by the model (A.1). If the ratio between the two is small,

$$F(\mathbf{x}_k) - F(\mathbf{x}_k + \mathbf{h}_k) \leq 0.25\,[\overline{F}(\mathbf{x}_k, \mathbf{0}) - \overline{F}(\mathbf{x}_k, \mathbf{h}_k)] \tag{A.2}$$

then the bound is decreased, $\Lambda_{k+1} = \Lambda_k/4$. Otherwise, if

$$F(\mathbf{x}_k) - F(\mathbf{x}_k + \mathbf{h}_k) \geq 0.75\,[\overline{F}(\mathbf{x}_k, \mathbf{0}) - \overline{F}(\mathbf{x}_k, \mathbf{h}_k)] \tag{A.3}$$

then $\Lambda_{k+1} = 2\Lambda_k$. If neither (A.2) nor (A.3) hold then we leave the bound unchanged, $\Lambda_{k+1} = \Lambda_k$.

Experiments have shown that the algorithm is rather insensitive to small changes in the constants used in the updating of the bound. This method has safe global convergence properties (Madsen 1975) and if the solution is regular then the final rate of convergence is quadratic (Madsen and Schjaer-Jacobsen 1978).

Method 2

It is a local method. It is assumed that a point near a solution $\mathbf{z}$ is known, and that the active sets $A(\mathbf{z}) \triangleq \{ j \mid f_j(\mathbf{z}) = F(\mathbf{z}) \}$ and $C(\mathbf{z}) \triangleq \{ i \mid \mathbf{a}_i^T \mathbf{z} + b_i = 0 \}$ are known. At a local minimum $\mathbf{z}$ of (3.9) the following necessary conditions hold (see, e.g., Hettich 1976),

$$
\begin{aligned}
&\sum_{j\in A(z)} \lambda_j f_j'(z) - \sum_{i\in C(z)} \mu_i a_i = 0 ,\\
&\sum_{j\in A(z)} \lambda_j - 1 = 0,\\
&f_{j_0}(z) - f_j(z) = 0 , \qquad j\in A(z)\setminus\{j_0\},\\
&a_i^T z + b_i = 0 , \qquad i\in C(z),
\end{aligned}
\tag{A.4}
$$

where the multipliers λ_j and μ_i are nonnegative and $j_0\in A(\mathbf{z})$ is fixed. Method 2 is an approximate Newton method for solving the nonlinear system (A.4) (in the variables $(\mathbf{z},\boldsymbol{\lambda},\boldsymbol{\mu})$). Exact first derivatives are used but the matrix $\Sigma\ \lambda_j f_j''(\mathbf{z})$ is approximated using a modified BFGS update. In this way an approximate Jacobian $\mathbf{J}_k$ is obtained at the estimate $(\mathbf{x}_k,\boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)})$ of the solution of (A.4). The next estimate is found by

$$
\mathbf{J}_k \begin{bmatrix} \Delta\mathbf{x}_k \\ \Delta\boldsymbol{\lambda}^{(k)} \\ \Delta\boldsymbol{\mu}^{(k)} \end{bmatrix} = -\mathbf{R}(\mathbf{x}_k, \boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)})
\tag{A.5}
$$

$$
(\mathbf{x}_{k+1}, \boldsymbol{\lambda}^{(k+1)}), \boldsymbol{\mu}^{(k+1)}) = (\mathbf{x}_k, \boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)}) + (\Delta\mathbf{x}_k, \Delta\boldsymbol{\lambda}^{(k)}, \Delta\boldsymbol{\mu}^{(k)})
$$

where $\mathbf{R}(\mathbf{z},\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathbf{0}$ is the vector formulation of (A.4).

We consider one iteration of Method 2. For simplicity, we use the notation $\mathbf{x} = \mathbf{x}_k$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}_k$, $\boldsymbol{\mu} = \boldsymbol{\mu}_k$, $A = A(\mathbf{z})$ and $C = C(\mathbf{z})$. In a Newton iteration for solving (A.4), we should use the Jacobian

$$
\mathbf{R}'(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \begin{bmatrix} \sum_{j\in A} \lambda_j f_j'' & \mathbf{E} & -\mathbf{F} \\ 0\,0\ldots 0\,0 & 1\,1\ldots 1 & 0\,0\ldots 0\,0 \\ \mathbf{G}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{F}^T & \mathbf{0} & \mathbf{0} \end{bmatrix}
\tag{A.6}
$$

where $\mathbf{E}$ has the columns $f_j'(\mathbf{x})$, $j\in A$, $\mathbf{F}$ has the columns a_i, $i\in C$, and $\mathbf{G}$ has the columns $f_{j_0}'(\mathbf{x}) - f_j'(\mathbf{x})$, $j\in A\setminus\{j_0\}$. Only the upper left hand block involves more than

first derivatives. In Method 2, this block is approximated by an updating formula whereas the exact values are used in the other blocks of $\mathbf{R}'$.

The Lagrangian function corresponding to (3.9) is

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \sum_{j=1}^{m} \lambda_j f_j(\mathbf{x}) - \sum_{i=1}^{\ell} \mu_i [\mathbf{a}_i^T \mathbf{x} + b_i], \tag{A.7}$$

so the upper left hand block of (A.6) is $\mathbf{L_{xx}}''(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu})$ since $\lambda_j = 0$ for $j \notin A$.

This block is approximated by the BFGS formula with the modifications of Powell (1978) that keep the approximation positive definite. Thus the matrix $\mathbf{J}_k$ of (A.5) is

$$\mathbf{J}_k = \begin{bmatrix} \mathbf{B}_k & \mathbf{E} & -\mathbf{F} \\ 0\,0 \ldots 0\,0 & 1\,1 \ldots 1 & 0\,0 \ldots 0\,0 \\ \mathbf{G}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{F}^T & \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{A.8}$$

where $\mathbf{B}_k$ is updated through

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \mathbf{B}_k \mathbf{s}\,\mathbf{s}^T \mathbf{B}_k / [\mathbf{s}^T \mathbf{B}_k \mathbf{s}] + \mathbf{y}\,\mathbf{y}^T / [\mathbf{s}^T \mathbf{y}] \tag{A.9}$$

with

$$\mathbf{s} = \mathbf{x}_{k+1} - \mathbf{x}$$

$$\mathbf{y} = \mathbf{L}_x'(\mathbf{x}_{k+1}, \boldsymbol{\lambda}, \boldsymbol{\mu}) - \mathbf{L}_x'(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$$

An iteration of Method 2 is now given by (A.8), (A.5), and (A.9) with

$$\mathbf{R}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{L}_x'(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \sum \lambda_j - 1 \\ \mathbf{e} \\ \mathbf{f} \end{bmatrix}, \tag{A.10}$$

where $\mathbf{e}$ has the components $f_{j_0}(\mathbf{x}) - f_j(\mathbf{x})$, $j \in A \backslash \{j_0\}$ and $\mathbf{f}$ has the components $\mathbf{a}_i^T \mathbf{x} + b_i$, $i \in C$.

The Combined Method

The combined method is the algorithm which we use in this thesis. Initially, Method 1 is used and the active sets used in (A.4) are estimated. When a singular local minimum seems to be approached a switch to Method 2 is made. If the Method 2 iteration is unsuccessful Method 1 is used again. Several switches between the two methods may take place. When Method 1 is used we say that the iteration is in Stage 1, otherwise it is in Stage 2. A detailed description of the two stages follows.

The Stage 1 Iteration

We have a point $\mathbf{x}_k$, a local bound Λ_k and a matrix $\mathbf{J}_k$ which should approximate the Jacobian of (A.4).

1. $\mathbf{x}_{k+1}$ and Λ_{k+1} are found using Method 1, and approximations A_{k+1} and C_{k+1} of the active sets at $\mathbf{x}_{k+1}$ are found via the active sets at the solution $\mathbf{h}_k$ of the linear model problem (A.1).
2. An estimate $(\boldsymbol{\lambda}^{(k+1)},\boldsymbol{\mu}^{(k+1)})$ of the multipliers is found through a least squares solution of (A.4) with $(\mathbf{x}_{k+1},A_{k+1},C_{k+1})$ inserted for $(\mathbf{z},A(\mathbf{z}),C(\mathbf{z}))$. This estimate is used for finding a new Jacobian estimate $\mathbf{J}_{k+1}$ by the BFGS update.
3. A switch to Stage 2 is made if the following two conditions hold:
   (a) The active set estimates have been constant over v consecutive different Stage 1 iterates.
   (b) The components of $\boldsymbol{\lambda}^{(k+1)}$ and $\boldsymbol{\mu}^{(k+1)}$ are nonnegative.

The Stage 2 Iteration

$\mathbf{x}_k$, Λ_k, $\mathbf{J}_k$ and active set estimates A_k,C_k are given.

1. Find $(\mathbf{x}_{k+1},\Lambda^{(k+1)},\mu^{(k+1)})$ and $\mathbf{J}_{k+1}$ using Method 2 with (A_k,C_k) inserted for $(A(\mathbf{z}), C(\mathbf{z}))$.

2. Let $A_{k+1}=A_k$, $C_{k+1}=C_k$ and $\Lambda_{k+1}=\Lambda_k$.

3. Switch to Stage 1 if one of the following conditions hold:

   (a) A function or constraint outside of A_{k+1} or C_{k+1} is active at $\mathbf{x}_{k+1}$.

   (b) A component of $\Lambda^{(k+1)}$ or $\mu^{(k+1)}$ is negative.

   (c) $\| \mathbf{R}(\mathbf{x}_{k+1},\Lambda^{(k+1)},\mu^{(k+1)}) \| > 0.999 \| \mathbf{R}(\mathbf{x}_k,\Lambda^{(k)},\mu^{(k)}) \|$ (see (A.5) for the definition of $\mathbf{R}$).

This completes the description of the combined method.

It has been shown (Hald and Madsen 1981) that the combined method can only converge to stationary points and that the final rate of convergence is quadratic on regular problems and superlinear on singular problems (provided that the Jacobian of (A.4) is regular).

APPENDIX B

CUBIC INTERPOLATION FORMULA

As a well-known fact, a maximum of a continuous differentiable function $e(\omega)$ is characterized by $e' \triangleq \partial e/\partial\omega = 0$ and $\partial^2 e/\partial\omega^2 < 0$. This implies a change in the sign of $\partial e/\partial\omega$ and, in the neighbourhood of the maximum, $\partial e/\partial\omega$ decreases as frequency increases. It follows that if there exist two points $\omega_1 < \omega_2$ such that

$$e'_{\omega_1} > 0 \quad \text{and} \quad e'_{\omega_2} < 0$$

at least one maximum of $e(\omega)$ lies between ω_1 and ω_2. If ω_1 and ω_2 are close enough to exclude the existence of multiple maxima, the location of the detected maximum can be estimated by the cubic interpolation formula (Fletcher and Powell 1963)

$$\omega_{max} = \omega_2 - \frac{(\omega_2 - \omega_1)\left[x - y - e'_{\omega_2}\right]}{e'_{\omega_1} - e'_{\omega_2} + 2x}, \tag{B.1}$$

where

$$y = -e'_{\omega_1} - e'_{\omega_2} + 3\,\frac{e(\omega_2) - e(\omega_1)}{\omega_2 - \omega_1} \tag{B.2}$$

and

$$x = \left[y^2 - e'_{\omega_1}\, e'_{\omega_2}\right]^{1/2}. \tag{B.3}$$

APPENDIX C

DESCRIPTION OF THE HALD AND MADSEN ℓ_1 ALGORITHM

We now give a detailed description of the method which is a combination of Method 1 and Method 2. We first describe the two basic methods and next the combined method, including switching rules.

Method 1

This is a method providing global convergence. At the kth step a feasible approximation $\mathbf{x}_k$ to a solution of (4.1) and a local bound Λ_k are given. In order to find a better estimate the following linearized problem is solved:

$$\underset{\mathbf{h}}{\text{minimize}} \qquad \overline{F}(\mathbf{x}_k, \mathbf{h}) \triangleq \sum_{j=1}^{m} |f_j(\mathbf{x}_k) + \mathbf{f}_j'(\mathbf{x}_k)^T \mathbf{h}|$$

subject to

$$\|\mathbf{h}\|_\infty \leq \Lambda_k, \tag{C.1}$$

$$\mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i = 0, \qquad i = 1, \ldots, \ell_{eq},$$

$$\mathbf{a}_i^T(\mathbf{x}_k + \mathbf{h}) + b_i \geq 0, \qquad i = (\ell_{eq}+1), \ldots, \ell.$$

The solution of (C.1), $\mathbf{h}_k$, may be found by a standard linear programming routine. However, an implementation of the algorithm of Bartels, Conn and Sinclair (1978) is used, which is more efficient. Notice that $(\mathbf{x}_k + \mathbf{h}_k)$ is feasible.

The next iterate is $(\mathbf{x}_k + \mathbf{h}_k)$ provided that this point is better than $\mathbf{x}_k$ in the sense of F, i.e., if $F(\mathbf{x}_k + \mathbf{h}_k) < F(\mathbf{x}_k)$. Otherwise $\mathbf{x}_{k+1} = \mathbf{x}_k$.

The local bound Λ_k is adjusted in every iteration based on comparison between the decrease in the nonlinear objective function and the decrease predicted

by the model $\bar{F}$. If the ratio between the two is small,

$$F(\mathbf{x}_k) - F(\mathbf{x}_k + \mathbf{h}_k) \leq 0.25[\bar{F}(\mathbf{x}_k, \mathbf{0}) - \bar{F}(\mathbf{x}_k, \mathbf{h}_k)], \tag{C.2}$$

then the bound is decreased: $\Lambda_{k+1} = \Lambda_k/4$. Otherwise, if

$$F(\mathbf{x}_k) - F(\mathbf{x}_k + \mathbf{h}_k) \geq 0.75[\bar{F}(\mathbf{x}_k, \mathbf{0}) - \bar{F}(\mathbf{x}_k, \mathbf{h}_k)], \tag{C.3}$$

then $\Lambda_{k+1} = 2\,\Lambda_k$. If neither (C.2) nor (C.3) hold then we leave the bound unchanged, $\Lambda_{k+1} = \Lambda_k$.

Experiments have shown that the method is rather insensitive to small changes in the constants used in this updating procedure of the local bound. Notice that if the new point $(\mathbf{x}_k + \mathbf{h}_k)$ is not accepted then the bound is decreased.

Method 2

This is a local method. It is assumed that a point near a solution $\mathbf{x}^*$ is known and that the set of zero functions

$$Z(\mathbf{x}^*) \triangleq \{j \mid f_j(\mathbf{x}^*) = 0\} \tag{C.4}$$

and the set of active constraints,

$$A(\mathbf{x}^*) \triangleq \{i \mid \mathbf{a}_i^T \mathbf{x}^* + b_i = 0\} \tag{C.5}$$

are known.

Method 2 is an approximate Newton method for solving the nonlinear system (C.9) (in the variables $(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\mu})$). Exact first derivatives are used but the matrix

$$\mathbf{g}''(\mathbf{x}^*) + \sum_{j \in Z} \delta_j \mathbf{f}_j''(\mathbf{x}^*)$$

is approximated using a modified BFGS update. In this way an approximate Jacobian $\mathbf{J}_k$ is obtained at the estimate $(\mathbf{x}_k, \boldsymbol{\delta}^{(k)}, \boldsymbol{\mu}^{(k)})$ of the solution of (C.9). The next estimate is obtained by

$$\mathbf{J}_k \begin{bmatrix} \Delta \mathbf{x}_k \\ \Delta \delta^{(k)} \\ \Delta \mu^{(k)} \end{bmatrix} = -\mathbf{R}\left(\mathbf{x}_k, \delta^{(k)}, \mu^{(k)}\right) \tag{C.6}$$

$$\left(\mathbf{x}_{k+1}, \delta^{(k+1)}, \mu^{(k+1)}\right) = \left(\mathbf{x}_k, \delta^{(k)}, \mu^{(k)}\right) + \left(\Delta \mathbf{x}_k, \Delta \delta^{(k)}, \Delta \mu^{(k)}\right)$$

where **R** is defined by (C.10). Notice that no line search is involved.

The Combined Method

The combined method is the algorithm which is used in this thesis. Method 1 is intended to provide the global convergence and Method 2 is used to obtain fast local convergence.

Initially, Method 1 is used and the sets (C.4) and (C.5) are estimated. When a local minimum seems to be approached a switch to Method 2 is made. If the Method 2 iteration is unsuccessful then Method 1 is used again. Several switches between the two methods may take place. When Method 1 is used we say that the iteration is in Stage 1, otherwise it is in Stage 2. A detailed description of the two stages follows.

The Stage 1 Iteration

We have a point $\mathbf{x}_k$, a local bound Λ_k and a matrix $\mathbf{J}_k$ which should approximate the Jacobian of (C.9).

1. $\mathbf{x}_{k+1}$ and Λ_{k+1} are found using Method 1, and approximations Z_{k+1} and A_{k+1} of the sets (C.4) and (C.5) are found via the zero and active sets at the solution $\mathbf{h}_k$ of the linear model problem (C.1).
2. An estimate $(\delta^{(k+1)}, \mu^{(k+1)})$ of the multipliers is found through a least squares solution of (C.9) with $(\mathbf{x}_{k+1}, Z_{k+1}, A_{k+1})$ inserted for $(\mathbf{x}, Z(\mathbf{x}),$

$A(\mathbf{x})$). This estimate is used for finding a new Jacobian estimate $\mathbf{J}_{k+1}$ by the BFGS method as described later in this appendix.

3. A switch to Stage 2 is made if the following two conditions hold:

   (a) The estimates Z_{k+1} and A_{k+1} have been constant over v consecutive different Stage 1 iterates ($v \geq 3$).

   (b) The multiplier estimates are in the correct ranges,

$$|\delta_j^{(k+1)}| \leq 1 ,$$

$$\mu_j^{(k+1)} \geq 0 .$$

The Stage 2 Iteration

We have an estimate $(\mathbf{x}_k, \boldsymbol{\delta}^{(k)}, \boldsymbol{\mu}^{(k)})$, estimates Z_k and A_k of (C.4) and (C.5), and a matrix $\mathbf{J}_k$ which should approximate the Jacobian of (C.9).

1. Find $(\mathbf{x}_{k+1}, \boldsymbol{\delta}^{(k+1)}, \boldsymbol{\mu}^{(k+1)})$ and $\mathbf{J}_{k+1}$ using Method 2 with (Z_k, A_k) inserted for $(Z(\mathbf{x}^*), A(\mathbf{x}^*))$.

2. Let $A_{k+1} = A_k$, $Z_{k+1} = Z_k$ and $\Lambda_{k+1} = \Lambda_k$.

3. Switch to Stage 1 if one of the following conditions holds:

   (a) A function f_j with $j \notin Z_k$ has changed sign, or a constraint corresponding to an index i with $i \notin A_k$ has become violated.

   (b) A component of $\boldsymbol{\delta}^{(k+1)}$ or of $\boldsymbol{\mu}^{(k+1)}$ is outside its range:

$$|\delta_j^{(k+1)}| > 1 ,$$

   or

$$\mu_j^{(k+1)} < 0 .$$

   (c) $$\|\mathbf{R}(\mathbf{x}_{k+1}, \boldsymbol{\delta}^{(k+1)}, \boldsymbol{\mu}^{(k+1)}\| > 0.999\|\mathbf{R}(\mathbf{x}_k, \boldsymbol{\delta}^{(k)}, \boldsymbol{\mu}^{(k)})\|$$

This completes the description of the combined method.

It has been shown by Hald and Madsen (1985), that the method has safe global convergence properties: it can only converge to stationary points. Furthermore, the final rate of convergence is at least superlinear, i.e.,

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \varepsilon_k \|\mathbf{x}_k - \mathbf{x}^*\| , \tag{C.7}$$

where $\varepsilon_k \to 0$ for $k \to \infty$.

Necessary Conditions for a Solution

At a solution $\mathbf{x}^*$ of the linearly constrained ℓ_1 problem (4.1) the functions which are zero play a special role since they contribute to the kinks of F. The functions which are non-zero at $\mathbf{x}^*$ give smooth contributions to F since $| f_j(\mathbf{x}) |$ is smooth near $\mathbf{x}^*$ when $f_j(\mathbf{x}^*) \neq 0$. Therefore we partition F into a smooth and a non-smooth part,

$$\begin{aligned} F(\mathbf{x}) &= \sum_{j \notin Z} |f_j(\mathbf{x})| + \sum_{j \in Z} |f_j(\mathbf{x})| \\ &= g(\mathbf{x}) + \sum_{j \in Z} |f_j(\mathbf{x})| \end{aligned} \tag{C.8}$$

where $Z = Z(\mathbf{x}^*)$ is defined by (C.4) and $g = g_{\mathbf{x}^*}$ is smooth in a neighbourhood of $\mathbf{x}^*$.

It is easily shown (see for instance Charalambous 1979) that the following set of equations hold at the local minimum $\mathbf{x} = \mathbf{x}^*$

$$\begin{aligned} &\mathbf{g}'(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}_j'(\mathbf{x}) - \sum_{i \in A} \mu_i \mathbf{a}_i = \mathbf{0} , \\ &f_j(\mathbf{x}) = 0 , \qquad\qquad j \in Z , \\ &\mathbf{a}^T \mathbf{x} + b_i = 0 , \qquad\qquad i \in A , \end{aligned} \tag{C.9}$$

where $| \delta_j | \leq 1$, $\mu_i \geq 0$, $Z = Z(\mathbf{x}^*)$ and $A = A(\mathbf{x}^*)$ are defined by (C.4) and (C.5), and

$$g(\mathbf{x}) = \sum_{j \notin Z} |f_j(\mathbf{x})| .$$

This set of equations corresponds to the Kuhn-Tucker conditions for the nonlinear programming problem which is equivalent to (4.1). The unknowns are $\mathbf{x}$, δ_j and μ_i, and it is seen that the number of unknowns equals the number of equations. If we use a vector notation (C.9) can be expressed as follows,

$$\mathbf{R}(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\mu}) = \mathbf{0}. \tag{C.10}$$

Updating the Matrices $\mathbf{J}_k$

The Jacobian of the nonlinear system (C.9) is

$$\mathbf{R}'(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{g}''(\mathbf{x}) + \sum_{j \in Z} \delta_j \mathbf{f}_j''(\mathbf{x}) & \mathbf{E} & \mathbf{F} \\ \mathbf{E}^T & \mathbf{0} & \mathbf{0} \\ -\mathbf{F}^T & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.11}$$

where $\mathbf{E}$ and $\mathbf{F}$ are matrices with columns $\mathbf{f}_j'(\mathbf{x})$, $j \in Z$, and $-\mathbf{a}_i$, $i \in A$, respectively.

In Method 2 we need to find an estimate $\mathbf{J}_{k+1}$ to $\mathbf{R}'(\mathbf{x}^{k+1}, \boldsymbol{\delta}^{(k+1)}, \boldsymbol{\mu}^{(k+1)})$. This is done as follows. At the iterate $\mathbf{x}_k$ estimates Z_k and A_k replace Z and A. The submatrices $\mathbf{E}$ and $\mathbf{F}$ are calculated exactly using $\mathbf{f}_j'(\mathbf{x}_k)$, $j \in Z_k$, and $-\mathbf{a}_i$, $i \in A_k$.

Only the upper left hand side part of $\mathbf{R}'$ needs to be approximated. This is done via a modification of the BFGS method, due to Powell (1978). The modification is necessary because the upper left hand side of $\mathbf{R}'$ is not necessarily positive semidefinite at a solution of (4.1). However, it is more stable and not less efficient to keep the approximation positive definite. The updating procedure becomes

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k / [\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k] + \mathbf{z}_k \mathbf{z}_k^T / [\mathbf{s}_k^T \mathbf{z}_k]$$

with

$$
\begin{aligned}
&\mathbf{s}_k = \mathbf{h}_k\,, \\
&\mathbf{z}_k = \theta\,\mathbf{y}_k + (1-\theta)\mathbf{B}_k\,\mathbf{s}_k\,, \qquad 0 < \theta \le 1\,, \\
&\mathbf{y}_k = G(\mathbf{x}_k + \mathbf{h}_k\,, \boldsymbol{\delta}^{(k)}\,, \boldsymbol{\mu}^{(k)}) - G(\mathbf{x}_k\,, \boldsymbol{\delta}^{(k)}\,, \boldsymbol{\mu}^{(k)})\,, \\
&G(\mathbf{x}\,, \boldsymbol{\delta}\,, \boldsymbol{\mu}) = \mathbf{g}'(\mathbf{x}) + \sum_{j \in Z_k} \delta_j\,\mathbf{f}_j'(\mathbf{x})\,,
\end{aligned}
\tag{C.12}
$$

where θ is defined such that $\mathbf{s}_k^T\mathbf{z}_k > 0$ which implies that positive definiteness is maintained. Notice that when θ is close enough to 0 this inequality will hold provided $\mathbf{B}_k$ is positive definite. Normally, however, θ can be chosen to be 1. Powell's formula for calculating θ is

$$
\theta = \begin{cases} 1 & \text{if } \mathbf{s}_k^T\mathbf{y}_k \ge 0.2\,t_k \\ 0.8\,t_k / [\,t_k - \mathbf{s}_k^T\mathbf{y}_k\,] & \text{otherwise} \end{cases}
\tag{C.13}
$$

with $t_k = \mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k$. We have found, however, that Powell's updating procedure becomes unstable when θ is too close to 0 and therefore we have modified (C.13). If θ found by (C.13) becomes less than 0.5 then we use $\theta = 0$. This implies that when (C.13) gives $\theta < 0.5$ then $\mathbf{B}_{k+1} = \mathbf{B}_k$.

This completes the description of the Jacobian approximation procedure.

REFERENCES

H.L. Abdel-Malek and J.W. Bandler (1980a), "Yield optimization for arbitrary statistical distributions - Part I: theory", IEEE Trans. Circuits and Systems, vol. CAS-27, pp.245-253.

H.L. Abdel-Malek and J.W. Bandler (1980b), "Yield optimization for arbitrary statistical distributions - Part II: implementation", IEEE Trans. Circuits and Systems, vol. CAS-27, pp. 253-262.

R.L. Adams and V.K. Manaktala (1975), "An optimization algorithm suitable for computer-assisted network tuning", Proc. IEEE Int. Symp. Circuits and Systems (Newton, MA), pp. 210-212.

D. Agnew (1981), "Improved minimax optimization for circuit design", IEEE Trans. Circuits and Systems, vol. CAS-28, pp. 791-803.

C.J. Alajajian, T.N. Trick and E.I. El-Masry (1980), "On the design of an efficient tuning algorithm", Proc. IEEE Int. Symp. Circuits and Systems (Houston, TX), pp. 807-811.

D.H. Anderson and M.R. Osborne (1977), "Discrete, nonlinear approximations in polyhedral norms: a Levenberg-like algorithm", Num. Math., vol. 28, pp. 157-170.

A.E. Atia (1974), "Computer-aided design of waveguide multiplexers", IEEE Trans. Microwave Theory Tech., vol. MTT-22, pp.332-336.

A.E. Atia and A.E. Williams (1972), "Narrow-bandpass waveguide filters", IEEE Trans. Microwave Theory Tech., vol. MTT-20, pp. 258-265.

J.W. Bandler (1969), "Optimization methods for computer-aided design", IEEE Trans. Microwave Theory Tech., vol. MTT-17, pp. 533-552.

J.W. Bandler (1972), "Optimization of design tolerances using nonlinear programming", Proc. 6th Princeton Conf. Information Sciences and Systems (Princeton, N.J.), pp.655-659. Also in Computer-Aided Filter Design, G. Szentirmai, Ed., New York: IEEE Press, 1973.

J.W. Bandler (1974), "Optimization of design tolerances using nonlinear programming", J. Optimization Theory and Applications, vol. 14, pp. 99-114.

J.W. Bandler and H.L. Abdel-Malek (1978a), "Advances in the mathematical programming approach to design centering, tolerancing and tuning", Proc. Joint Automatic Control Conf. (Philadelphia, PA) , pp. 329-344.

J.W. Bandler and H.L. Abdel-Malek (1978b), "Optimal centering, tolerancing and yield determination via updated approximations and cuts", IEEE Trans. Circuits and Systems, vol. CAS-25, pp. 853-871.

J.W. Bandler and R.M. Biernacki (1980), "Postproduction parameter identification and tuning of analog circuits", Proc. European Conf. Circuit Theory and Design (Warsaw, Poland), vol. 2, pp. 205-220.

J.W. Bandler, R.M. Biernacki, A.E. Salama and J.A. Starzyk (1982), "Fault isolation in linear analog circuits using the ℓ_1 norm", Proc. IEEE Int. Symp. Circuits and Systems (Rome, Italy), pp. 1140-1143 .

J.W. Bandler and C. Charalambous (1972a), "Practical least pth optimization of networks", IEEE Trans. Microwave Theory Tech., vol. MTT-20, pp. 834-840.

J.W. Bandler and C. Charalambous (1972b), "Theory of generalized least pth approximation", IEEE Trans. Circuit Theory, vol. CT-19, pp. 287-289.

J.W. Bandler and C. Charalambous (1973), "New algorithms for network optimization", IEEE Trans. Microwave Theory Tech., vol. MTT-21, pp. 815-818. .

J.W. Bandler and S.H.Chen (1984), "Interactive optimization of multi-coupled cavity microwave filters", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-84-13-R.

J.W. Bandler, S.H. Chen and S. Daijavad (1985), "Exact sensitivity analysis and optimization for multi-cavity microwave communication filters", Proc. IEEE Int. Symp. Circuits and Systems (Kyoto, Japan), pp. 1587-1590

J.W. Bandler, S.H. Chen, S. Daijavad and W. Kellermann (1984), "Optimal design of multi-cavity filters and contiguous-band multiplexers", Proc. 14th European Microwave Conference (Liege, Belgium), pp. 863-868.

J.W. Bandler, S. Daijavad and Q.J. Zhang (1985),"Computer-aided design of branched cascaded networks", Proc. IEEE Int. Symp. Circuits and Systems (Kyoto, Japan), pp. 1579-1582.

J.W. Bandler and M.A. El-Kady (1980a), "A unified approach to power system sensitivity analysis and planning, Part I: family of adjoint systems", Proc. IEEE Int. Symp. Circuits and Systems (Houston, TX), pp. 681-687.

J.W. Bandler and M.A. El-Kady (1980b), "A unified approach to power system sensitivity analysis and planning, Part II: special class of adjoint systems", Proc. IEEE Int. Symp. Circuits and Systems (Houston, TX), pp. 688-692.

J.W. Bandler, M.A. El-Kady, W. Kellermann and W.M. Zuberek (1983a), "BSTALN - a computer implementation of an algorithm for the best mechanical alignment problem", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-83-13.

J.W. Bandler, M.A. El-Kady, W. Kellermann and W.M. Zuberek (1983b), "An optimization approach to the best alignment of manufactured and operating systems", Proc. IEEE Int. Symp. Circuits and Systems (Newport Beach, CA), pp.542-545.

J.W. Bandler, M.A. El-Kady, W. Kellermann and W.M. Zuberek (1984), "A minimax approach to the best mechanical alignment problem", ASME Journal of Mechanisms, Transmissions and Automation in Design, vol. 106, pp. 31-40.

J.W. Bandler, M.A. El-Kady and J. Wojciechowski (1983), "TTM1 - a Fortran implementation of the Tellegen theorem method to power system simulation and design", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-82-12-U2.

J.W. Bandler and W. Kellermann (1983), "Selected topics in optimal design centering, tolerancing and tuning", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-83-28-R.

J.W. Bandler, W. Kellermann and K. Madsen (1985a), "A superlinearly convergent algorithm for nonlinear ℓ_1 optimization with circuit applications", Proc. IEEE Int. Symp. Circuits and Systems (Kyoto, Japan), pp. 977-980.

J.W. Bandler, W. Kellermann and K. Madsen (1985b), "A superlinearly convergent minimax algorithm for microwave circuit design", IEEE Trans. Microwave Theory Tech., vol. MTT-33.

J.W. Bandler, W. Kellermann and K. Madsen (1985c), "An algorithm for nonlinear ℓ_1 optimization for design and diagnosis of networks", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-85-14-R.

J.W. Bandler and P.A. Macdonald (1969a), "Cascaded noncommensurate transmisssion-line networks as optimization problems", IEEE Trans. Circuit Theory, vol. CT-16, pp. 391-394.

J.W. Bandler and P.A. Macdonald (1969b), "Optimization of microwave networks by razor search", IEEE Trans. Microwave Theory Tech., vol. MTT-17, pp. 552-562.

J.W. Bandler and P.C. Liu (1974), "Automated network design with optimal tolerances", IEEE Trans. Circuits and Systems, vol. CAS-21, pp. 219-222.

J.W. Bandler, P.C. Liu and J.H.K. Chen (1975), "Worst case network tolerance optimization", IEEE Trans. Microwave Theory Tech., vol. MTT-23, pp. 630-641.

J.W. Bandler, P.C. Liu and H. Tromp (1976a), "A nonlinear programming approach to optimal design centering, tolerancing and tuning", IEEE Trans. Circuits and Systems, vol. CAS-23, pp. 155-165.

J.W. Bandler, P.C. Liu and H. Tromp (1976b), "Integrated approach to microwave design", IEEE Trans. Microwave Theory Tech., vol. MTT-24, pp. 584-591.

J.W. Bandler and M.R.M. Rizk (1979), "Optimization of electrical circuits", Mathematical Programming Study on Engineering Optimization, vol. 11, pp. 1-64.

J.W. Bandler, M.R.M. Rizk and H.L. Abdel-Malek (1978),"New results in network simulation, sensitivity and tolerance analysis for cascaded structures", IEEE Trans. Microwave Theory Tech., vol. MTT-26, pp. 963-972.

J.W. Bandler and A.E. Salama (1981), "Postproduction tuning employing network sensitivities", Proc. European Conf. Circuit Theory and Design (The Hague, Netherlands), pp. 704-709.

J.W. Bandler and A.E. Salama (1983), "Integrated approach to microwave postproduction tuning", IEEE Int. Microwave Symp. Digest (Boston, MA), pp. 415-417.

J.W. Bandler and A.E. Salama (1985), "Functional approach to microwave postproduction tuning", IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 302-310.

J.W. Bandler and W.M. Zuberek (1982a), "MMLC - a Fortran package for linearly constrained minimax optimization", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-82-5.

J.W. Bandler and W.M. Zuberek (1982b), "MFNC - a Fortran package for minimization with general constraints", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-82-6.

J.W. Bandler and W.M. Zuberek (1982c), "LFLFD - a Fortran implementation of the fast decoupled load flow technique", Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, Report SOS-82-8.

R.H. Bartels and A.R. Conn (1981), "An approach to nonlinear ℓ_1 data fitting", Computer Science Department, University of Waterloo, Waterloo, Ontario, Report CS-81-17.

R.H. Bartels, A.R. Conn and J.W. Sinclair (1978), "Minimization techniques for piecewise differentiable functions: the ℓ_1 solution to an overdetermined linear system", SIAM J. Numerical Analysis, vol. 15, pp. 224-241.

R.K. Brayton, S.W. Director, G.D. Hachtel and L.M. Vidigal (1979), "A new algorithm for statistical circuit design based on quasi-Newton methods and function splitting", IEEE Trans. Circuits and Systems, vol. CAS-26, pp. 784-794.

H.J. Bremermann (1962), "Optimization through evolution and recombination", in Self-Organizing Systems 1962, M.C. Yovits, G.T. Jacobi and G.D. Goldstein, Eds. Washington D.C.: Spartan Books, pp. 93-106.

E.M. Butler (1971), "Realistic design using large-change sensitivities and performance contours", IEEE Trans. Circuit Theory, vol. CT-18, pp. 58-66.

S.M. Chan and F.C. Schweppe (1979), "A generation reallocation and load shedding algorithm", IEEE Trans. Power Apparatus and Systems, vol. PAS-98, pp.26-34.

S.M. Chan and E. Yip (1979), "A solution of the transmission limited dispatch problem by sparse linear programming", IEEE Trans. Power Apparatus and Systems, vol. PAS-98, pp. 1044-1053.

C. Charalambous (1979), "On conditions for optimality of the nonlinear ℓ_1 problem", Mathematical Programming, vol. 17, pp. 123-135.

C. Charalambous and A.R. Conn (1975), "Optimization of microwave networks", IEEE Trans. Microwave Theory Tech., vol. MTT-23, pp. 834-838.

C. Charalambous and A.R. Conn (1978), "An efficient method to solve the minimax problem directly", SIAM J. Numerical Analysis, vol. 15, pp. 162-187.

M.H. Chen (1983), "A 12-channel contiguous band multiplexer at KU-band", IEEE Int. Microwave Symp. Digest (Boston, MA), pp. 77-79.

M.H. Chen, F. Assal and C. Mahle (1976), "A contiguous band multiplexer", COMSAT Technical Review, vol. 6, pp. 285-306.

A.R. Conn (1979), "An efficient second-order method to solve the (constrained) minimax problem", Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Report CORR 79-5.

S.W. Director and G.D. Hachtel (1977), "The simplicial approximation approach to design centering", IEEE Trans. Circuits and Systems, vol. CAS-24, pp. 363-372.

B.C. Eaves and W.I. Zangwill (1971), "Generalized cutting plane algorithms", SIAM J. Control, vol. 9, pp. 529-542.

R.G. Egri, A.E. Williams and A.E. Atia (1983), "A contiguous-band multiplexer design", IEEE Int. Microwave Symp. Digest (Boston, MA), pp. 86-88.

R.A. El-Attar, M. Vidyasagar and S.R.K. Dutta (1979),"An algorithm for ℓ_1-norm minimization with application to nonlinear ℓ_1-approximation", SIAM J. Numerical Analysis, vol. 16, pp. 70-86.

R. Fletcher and M.J.D. Powell (1963), "A rapidly convergent descent method for minimization", Computer J., vol. 6, pp. 163-168.

M.Z. Ghoneim, M.A.N. Askourah and S.A. Rahman (1977), "Estimation of optimal load shedding in transmission system using mathematical programming techniques", IEEE PES Summer Meeting (Mexico City, Mexico), paper A77 595-2.

T.C. Giras and S.N. Talukdar (1981), "Quasi-Newton method for optimal power flows", Int. Journal of Electrical Power and Energy Systems, vol. 3, pp.59-64.

G.D. Hachtel, T.R. Scott and R.P. Zug (1980), "An interactive linear programming approach to model parameter fitting and worst-case circuit design", IEEE Trans. Circuits and Systems, vol. CAS-27, pp. 871-881.

J. Hald (1981), "MMLA1Q, a Fortran subroutine for linearly constrained minimax optimization", Institute for Numerical Analysis, Technical University of Denmark, Lyngby, Denmark, Report NI-81-01.

J. Hald and K. Madsen (1981), "Combined LP and quasi-Newton methods for minimax optimization", Mathematical Programming, vol. 20, pp. 49-62.

J. Hald and K. Madsen (1985), "Combined LP and quasi-Newton methods for nonlinear ℓ_1 optimization", SIAM J. Numerical Analysis, vol. 22, pp. 68-80.

S.P. Han (1976), "Superlinearly convergent variable metric algorithms for general nonlinear programming problem", Mathematical Programming, vol. 11, pp. 263-282.

S.P. Han (1981), "Variable metric methods for minimizing a class of nondifferentiable functions", Mathematical Programming, vol. 20, pp. 1-13.

R. Hettich (1976), "A Newton method for nonlinear Chebyshev approximation", in Approximation Theory, R. Schaback and K. Scherer, Eds., Lecture Notes in Mathematics, 556, Springer Berlin, pp. 222-236.

B. Karafin (1974), "The general component tolerance assignment problem in electrical networks", Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA.

B. Krogh and S.H. Javid (1983), "Multi-stage rescheduling of generation, load shedding and short-term transmission capacity for emergency state control", IEEE Trans. Power Apparatus and Systems, vol. PAS-102, pp. 1466-1472.

M.R. Lightner and S.W. Director (1981), "Multiple criterion optimization for the design of electronic circuits", IEEE Trans. Circuits and Systems, vol. CAS-28, pp. 169-179.

P.C. Liu (1975), "A theory for optimal worst-case design embodying centering, tolerancing and tuning with circuit applications", Ph.D. thesis, McMaster University, Hamilton, Canada.

P.V. Lopresti (1977), "Optimum design of linear tuning algorithms", IEEE Trans. Circuits and Systems, vol. CAS-24, pp. 144-151.

E. Lüder and B. Kaiser, (1976), "Precision tuning of minizaturized circuits", Proc. IEEE Int. Symp. Circuits and Systems (Munich, West Germany), pp. 209-211.

K. Madsen (1975), "An algorithm for minimax solution of overdetermined systems of nonlinear equations", J. IMA, vol. 16, pp.321-328.

K. Madsen and H. Schjaer-Jacobsen (1976), "Singularities in minimax optimization of networks", IEEE Trans. Circuits and Systems, vol. CAS-23, pp. 456-460.

K. Madsen and H. Schjaer-Jacobsen (1978), "Linearly constrained minimax optimization", Mathematical Programming, vol. 14, pp. 208-223.

D.Q. Mayne, E. Polak and R. Trahan (1979), "An outer approximations algorithm for computer-aided design problems", J. Optimization Theory and Applications, vol. 28, pp. 331-352.

D.Q. Mayne, E. Polak and A. Voreadis (1982), "A cut-map algorithm for design problems with parameter tolerances", IEEE Trans. Circuits and Systems, vol. CAS-29, pp. 35-45.

R.A. McLean and G.A. Watson (1980), "Numerical methods for nonlinear discrete ℓ_1 approximation problems", in Numerical Methods of Approximation Theory, Birkhauser Verlag, vol. 5, pp. 169-183.

T.K.P. Medicherla, R. Billinton and M.S. Sachdev (1979), "Generation rescheduling and load shedding to alleviate line overloads - analysis", IEEE Trans. Power Apparatus and Systems, vol. PAS-98, pp. 1876-1884.

T.K.P. Medicherla, R. Billinton and M.S. Sachdev (1981), "Generation rescheduling and load shedding to alleviate line overloads - system studies", IEEE Trans. Power Apparatus and Systems, vol. PAS-100, pp. 36-42.

H.M. Merrill (1973), "Failure diagnosis using quadratic programming", IEEE Trans. Reliability, vol. R-22, pp. 207-213.

W. Murray and M.L. Overton (1981), "A projected Lagrangian algorithm for nonlinear ℓ_1 optimization", SIAM J. Scient. Stat. Comp., vol. 2, pp. 207-224.

M.R. Osborne and G.A. Watson (1969), "An algorithm for minimax optimization in the nonlinear case", Computer J., vol. 12, pp. 63 -68.

M.R. Osborne and G.A. Watson (1971), "On an algorithm for discrete nonlinear ℓ_1 approximation", Computer J., vol. 14, pp.184-188.

M.L. Overton (1979), "Projected Lagrangian algorithms for nonlinear minimax and ℓ_1 optimization", Department of Computer Science, Stanford University, Report STAN-CS-79-752.

P. Petersen and R. Johnson (1979, 1980, 1982), Corporate Engineering, Woodward Governor Company, Rockford, Illinois 61101, private communications.

J.F. Pinel (1973), "Tolerance assignment and network alignment in the frequency domain", in IEEE Short Course on Computer-Aided Network Design, 73-SC-06, pp. 17-25.

J.F. Pinel and K.A. Roberts (1972), "Tolerance assignment in linear networks using nonlinear programming", IEEE Trans. Circuit Theory, vol. CT-19, pp. 475-479.

E. Polak and A. Sangiovanni-Vincentelli (1979), "Theoretical and computational aspects of the optimal design centering, tolerancing and tuning problem", IEEE Trans. Circuits and Systems, vol. CAS-26, pp. 795-813.

Lecture Notes in Mathematics 630, G.A. Watson, Ed. Berlin: Springer Verlag, 1978, pp. 144-157.

M.J.D. Powell (1978), "The convergence of variable metric method for nonlinearly constrained optimization calculations", in Nonlinear Programming, O.L. Mangasarian, R.R. Meyer and S.M. Robinson, Eds. New York: Academic Press, pp.27-63.

H. Schjaer-Jacobsen and K. Madsen (1979), "Algorithms for worst-case tolerance optimization", IEEE Trans. Circuits and Systems, vol. CAS-26, pp. 775-783.

T.D. Shockley and C.F. Morris (1973), "Computerized design and tuning of active filters", IEEE Trans. Circuits Theory, vol. CT-20, pp. 438-441.

B. Stott and O. Alsac (1974), "Fast decoupled load flow", IEEE Trans. Power Apparatus and Systems, vol. PAS-93, pp. 859-869.

B. Stott and E. Hobson (1977a), "Power system security control calculations using linear programming, part I", IEEE PES Summer Meeting (Mexico City, Mexico), paper F77 577-0.

B. Stott and E. Hobson (1977b), "Power system security control calculations using linear programming, part II", IEEE PES Summer Meeting, (Mexico City, Mexico), paper F77 573-8.

S.N. Talukdar, T.C. Giras and V.K. Kalyan (1983), "Decompositions for optimal power flows", IEEE Trans. Power Apparatus and Systems, vol. PAS-102, pp. 3877-3884.

S.N. Talukdar and F.F. Wu (1981), "Computer-aided dispatch for electric power systems", Proc. IEEE, vol. 69, pp. 1212-1231.

R. Tong and D. Smith (1984), "A 12-channel contiguous band multiplexer for satellite application", IEEE Microwave Symp. Digest (San Francisco, CA), pp. 297-298.

H. Tromp (1977), "The generalized tolerance problem and worst case search", Conf. Computer-Aided Design of Electronic and Microwave Circuits and Systems (Hull, England), pp. 72-77.

L.M. Vidigal and S.W. Director (1982), "A design centering algorithm for nonconvex regions of acceptability", IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, vol. CAD-1, pp. 13-24.

A. Voreadis and D.Q. Mayne (1982), "A cut-map algorithm for design problems with parameter tolerances and tuning", IEEE Trans. Circuits and Systems, vol. CAS-29, pp. 288-298.

G.A. Watson (1979), "The minimax solution of an overdetermined system of nonlinear equations", J. IMA, vol. 23, pp. 167-180.

G.A. Watson (1984), "Discrete ℓ_1 approximation by rational functions", IMA J. of Numerical Analysis, vol. 4, pp. 275-288.

Woodward Governor Company (1982), Rockford, Illinois 61101, Sample data sent to McMaster University.

AUTHOR INDEX

SUBJECT INDEX