UNIVERSITY OF LONDON

EXHAUSTIVE SEARCH FOR THE GLOBAL OPTIMUM

OF

LEAST VOLUME STRUCTURES

by

Luis Miguel da Cruz Simões

A Thesis submitted to the University of London

for the Degree of Doctor of Philosophy

Department of Civil and Municipal Engineering

University College London

NOVEMBER 1982

ABSTRACT

The primary objective of this work is to obtain the global solution of a structural synthesis problem namely the minimum volume design of trusses.

This bilinearly constrained problem may present multiple optima and some examples of this nonconvex behaviour are given.In the particular class of structural optimization problems there is an incentive to determine whether a previously obtained local optimum is the global optimum over a given range of the variables in order to terminate the search. The performance of a recently proposed uniqueness test [McCormick,(1980)] is studied within the context of the current problem.

Soland`s (1971) branch and bound solution strategy is directly applied to our problem and several versions based on concepts of both separable and factorable functions are presented.On the basis of computational experience conclusions are drawn on the suitability of the different underestimates taken and of several heuristic modifications tried out. A different algorithm although of branch and bound type initially used by Reeves in 1973 to solve all-quadratic programming problems is redefined in terms of factorable envelopes.

Wolsey's generalization (1981) of Bender's algorithm yields a master problem that will be solved by a process of relaxation. For the bilinearly constrained problem both

support and cutting functions can be linearized by introducing binary variables. Although the relaxed problem that has to be solved at each iteration is still nonconvex it is amenable to standard codes for Ø-1 mixed LP. Using the coupling properties among the variables of the present problem a shortened version of this method is also formulated. Results are compared with those provided by the stategies previously mentioned.

Both solution techniques covered in this study are combinatorial in nature and thus lend themselves to the casting of the discrete variable case (restricted list of available sections). These extensions of the algorithm are demonstrated with examples.

CONTENTS

CHAPTER ONE

INTRODUCTION

1.1. Thesis

The human capability to solve large not necessarly convex mathematical problems arising from structural engineering models is still rather limited . Some of the difficulties are associated with the inability to guarantee the obtaining of a global as opposed to merely local solution to the general nonlinear optimization problem. Minimum volume of trusses are an example of such problems possessing a large field of practical applications and being themselves of intrinsic importance. In fact this one stress-resultant problem can be extended to cover the solution of both frame and plate problems. Bounds on joint displacements,member stresses,buckling restrictions and member sizes are the type of constraints frequentely encountered and to which this work is addressed.

Numerous methods have been proposed to tackle this problem for member sizes chosen from both a continuum and a discrete set of available sections. In none of them statements are made about the convexity of the problem or claims for convergence to the global solution. Nothing prevents an algorithm used for convex programming when applied to nonconvex programming from converging to a local solution or not being able to converge at all. In fact all these methods seem to work well in some problems in the sense that they can generate a good approximation using a reasonable amount of time. Unfortunately this cannot be guaranteed in general.

In the present study the problem is approached as fully nonconvex even if in specific cases the nonconvexity may not manifest itself strongly or at all ; in this way the occurence of local minima can be observed and some guidelines may be established to help in deciding whether the simple optimization strategies should be held as sufficient or the more sophisticated tools of combinatorial optimization should be used. Reliable identification of any local optimum requires search techniques which a) are logically complete in the sense that the process will terminate only after all optima have been found and b) computationaly efficient so that detection of the optima is not prevented by convergence problems. There are classes of nonconvex problems where property b) is not crucial since each local solution can be obtained numerically after a finite number of steps. This is the case of a quadratic nonconvex function subject to a linear inequality system. Unfortunately this property no longer holds when the constraints are bilinear. The difference between these cases will be explained in more detail in the following Chapter.

Two strategies more appropriate for nonconvex minimization were selected. The Branch and Bound approach (B & B) is a globally convergent method that can be applied to the solution of the general separable problem such as functions of one variable and their products. Although historically it is considered a remedy of last resort B & B algorithms are being accepted in structural optimization as competitive since each subproblem can be solved by using linear programming (LP) techniques.

Alternatively a generalization of Bender's algorithm by using dual functions instead of dual variables and LP Duality theory yields a master problem equivalent to the minimum volume design. Even though the master problem is a Ø-1 mixed LP and therefore still nonconvex the importance of this procedure is due to the fact that it can be solved by available globally convergent codes. Several versions of both methods were run trying out possible acceleration schemes. The computational effort of these algorithms is reported as well as the relative efficiency of some heuristic improvements.

1.2. Structural model and computational technique

Trusses are discretized into a finite element model and the Lagrange coordinate technique is used to establish the statics and kinematics of the graph model under the assumptions of linear elastic behaviour. This type of structure is taken for the purpose of presenting the mathematical model although the method is in principle applicable to any finite element discretization of continuous models if the problem-size obstacle can be overcome. The structure is acted upon by either single or multiple loading. In many cases the numerical solution requires a basic reliable Simplex system which can be adopted to the requirements of the problem. Land & Powell's (1973) LP routine has been developed for this purpose. Although not claiming to be very efficient it is reliable and provides sufficient checks and stopping rules to reassure the user about its ouput. The examples presented in this work have been processed with these routines in various adapted forms .

1.3. Organization of the thesis

The object and main purpose of this research are briefly described in Chapter 1 together with the computational techniques used. An historical outline of the related efforts in this field will close the Introduction.

The bilinearly constrained problem that is the subject of this work is described in Chapter 2 and an illustrative example of multiple optima solution is given. Some elements of mathematical programming strategy are also collected together in anticipation of in forthcoming Chapters.

In Chapter 3 mesh and nodal descriptions of the statics and kinematics of a structure are taken as the constraints of an optimization problem. Examples of trusses possessing multiple optima are given. Conditions are stated that ensure that a local minimizer is indeed the global minimizer of a constrained nonconvex program using McCormick (1980) relevant theory.

The Branch and Bound strategy is presented in Chapter 4 in order to solve the minimum volume of the truss with continous design variables. A rotation of the coordinate axes is employed to transform the initially factorable terms into an equivalent separable formulation which can be solved by Soland's algorithm (1971) . Several underestimates each defining a convex hull are used to create the subproblems used in this strategy. The concept of convex envelope of a factorable function defined in a rectangle of bounds is used to provide tighter underestimates. Computational experience in the examples used as testbed will be reported. A different combinatorial scheme based on a modified version of Reeves' work on all quadratic programming (1973) is also tried out.

In Chapter 5 the Resource-Decomposition method is described as a generalization of Bender's algorithm to nonconvex problems. A Master program equivalent to our Bilinearly constrained problem is defined and solved by a process of relaxation. Wolsey's cutting planes (1981) are added at each iteration of the algorithm and their special nature plays an important role in the selection of the methods of solution of each relaxed problem. Several results of the application of this theory are compared with solutions previously obtained.

Both alternative strategies to solve the minimization problem are extended in Chapter 6 to cover the synthesis of structures where member sizes belong to a set of available gauges. Their results are compared and some statements are made about the validity of the methods described in

Chapters 4 and 5 when applied to both continuous and discrete member sizes.

1.4. Historical outline

Historically it is difficult to ascertain the original attempts to determine the least volume of material a structure needs as regard its strength. Probably this coincided with the application of the notion of force and the laws of mechanics as foundation for designing structures. It is generally accepted that this period began in 1638 the year Galileo published the results of his experiments. Since that time theoretical mechanics has constituted a cause for the development of ways of predicting the behaviour of stuctures prior to construction.

On the basis of a theorem prooved by Maxwell,Michell indicated in 1904 the possibility of determining the minimum material design of statically determinate trusses subjected to a single loading condition. He concluded that the member stresses must be at their limiting values for the structure to be optimal. The development of design methods continued in the direction of finding more rational procedures for the needs of practical engineers. The geometry and the materials of the structure were adopted intuitively and for specified design loads the values of the state variables were evaluated by using methods of analysis. The design variables would then be modified in repeated cycles until the calculated behaviour would satisfy certain prescribed requirements.

The design philosophy in structural engineering was changed by the application of matrix methods and the development of electronic computers possessing tremendous storage capacity,speed and accuracy. Optimum design methods replaced design and analysis procedures where the final solution can be highly uneconomical even if the intuitive solution satisfies behavioural constraints. Michell's trusses underlined the importance of fully stressed structures for near optimal designs. More recently research papers have been published which supplement this notion with some newer optimality criteria. These developments have paralled the progress of mathematical programming,also a product of the last quarter of century of research.

These two fundamentally different approaches have been commonly used in systematic structural synthesis. Optimality criteria was shown by Templeman in 1975 to be an indirect method based on obtaining a set of necessary conditions for the optimal design problem. In mathematical programming one starts with an engineering estimate of the optimum design. A direction of search is then computed based on local behaviour of cost and constraint functions;this procedure can be considered a direct method.

1.4.1 Optimality criteria

The simplest type of optimality ctriterion is the fully stressed member condition applicable to statically determinate trusses. For indeterminate structures this approach makes the assumption of a limited redistribution of internal forces due to redundancy. The fully stressed design criterion is relevant to the case when only stress restrictions are taken into account. Its main assumption says that in the optimal structure the maximum allowable stress is attained in each member under at least one of the applied loading cases. The stress ratio formulae is applied recursively after each reanalysis phase as if the structure was isostatic in which case this criterion was rigourous. The main remark about this algorithm presented by Razani (1966) is that it leads to a vertex in the design space which is not necessarily an optimal point and is independent of the objective function (OF) .

Barnet (1968) by introducing a virtual load acting at the point where a deflection is prescribed defines the minimum weight design in terms of the virtual strain energy. Chern and Prager (1971) derived that the ratio of the strain energy density to the volume density is the same for all members in a displacement constrained isostatic truss. In the hyperstatic case this statement no longer applies. Berke (1972) obtained a similar recursive relation to the stress ratio formulae by introducing Lagrange multipliers into the equality constrained problem. Since not all these multipliers are positive the variables are divided into active and passive groups.

By combining the foregoing design relations the problem consists of finding Lagrange multipliers after each reanalysis. Berke and Khot (1974) pointed out that the major drawback associated with these methods consists of the selection of active variables and constraints. To avoid this last difficulty the optimality criteria can be derived from a problem involving inequality instead of equality constraints. The stress ratio formulae can be efficiently substituted by a better approximation of the stresses based on the virtual load procedure as linear combination of displacements. To the resulting problem that has to be recurred iteratively the optimality criteria follows from the application of the K-T necessary conditions. The dual problem contains entirely the selection of active constraints by assigning the Lagrange multipliers positive values if the constraint is active and zero otherwise.

The solution of this approximate problem consists of determining the optimal dual variables subject to sense restrictions by maximizing the Lagrangian of the dual problem. Bartholomew (1978) found a dual bound on the minimum weight based in this procedure. The dual problem partitions the dual space in several subregions corresponding to different divisions in active and passive design variables and accordingly to different definitions of the OF [Fleury (1979)] .

1.4.2 Mathematical programming

The mathematical programming approach attacks directly the nonlinear problem defining the minimum weight/volume design. The classical minimization methods require reanalysis of the structure to evaluate the gradients of the constraints that are nonlinear homogeneous functions of the design variables. The earlier approaches have used primal methods and were considered to be not efficient for large scale systems since the number of iteration steps would increase with the number of design variables. The reader is refered to the works by Schmit (1963) Gellatly (1966) and Brown's (1966) application of Rosen's projected gradient method.

Moses (1964) and Reinschmit (1966) were some of the pioneers in the use of penalty methods that replace the initially constrained form into a sequence of unconstrained problems by adding auxiliary functions reflecting the behaviour of the constraints. This approach is also not efficient since the number of iterations increases just as in the primal methods. Geometric programming has also been used to minimize weight of planar trusses [Templeman (1975)] . Singary and Rao (1975) have made an application of optimal control theory by partitioning the original problem into a number of smaller subproblems and solving them in a sequential manner.

The next step has been the use of a sequence of linear programs obtained by linearizing the objective function and the constraints at the design point. The convergence to a local minimum can only be guaranteed if it exists at a vertex in the design space . Working in the compliance design space in order to reduce linearization errors in the constraints Schmit and Miura (1976) considered approximations obtained by first order Taylor series expansion. They are able to predict dependence relations between state and design variables. Periodically constraint approximations are updated by performing a reanalysis.

Arora and Haug (1976) uses a function space gradient projection technique. The gradient required is computed by using adjoint design variable sensitivity methods and is reported to converge quickly to a local solution of the optimization problem. The future trend of research for large structures is to combine both approaches. The development of hybrid methods was justified by Fleury (1980):"After the mathematical programming has identified a set of constraints close to the optimum the optimality criteria should then be used to find the precise optimum due to its quick convergence once the active set is determined". Rajararam and Schmit (1981) integrated basis reduction concepts to reduce the dimensionality of large scale systems.

This work follows the lines of the latter approach because it affords greater generality in casting the format of the problem.

1.4.3 Discrete optimization

The nonlinear optimization problem when expressed in discrete design variables can be thought of as imbedded in the space of continous solutions. Discrete optimization techniques are relatively new when compared to automatic design of structures and both linear and nonlinear optimization methods. In this area Toackley (1968) and Marcal (1968) have formulated the minimum weight problem directly in discrete variables. Toakley formulated the elastic design problem as a discrete programmming with Ø-1 variables but was forced to terminate the program and resort to bounding procedures due to defficiencies of Gomory's algorithm.

Cella (1971) uses a Branch and Bound type of strategy to optimize trusses in the elastic range. The B and B method is capable of solving mixed LP but efficiency depends on the the type of problem and particular scheme for branching used. Reinschmidt (1971) focused attention on the pure Integer Progamming problem. By using Plastic theory he optimized the member sizes. Imai and Shoji (1981) created a sequence of approximate optimization problems in the compliance design space of the linear elastic range and have solved each subproblem in the dual space. Convergence to a local solution was met afer a small number of reanalysis.

Saglam (1981) used a cutting plane method providing convergence again to a local solution. In order to increase the chances of finding the global solution the procedure should be restarted with different initial trial points and the best solution found would then be selected. Yates,Templeman and Boffey (1982) in their very recent paper have justified the use of methods that approximately solve the problem. They have also shown that they are equivalent in complexity to those used for solving the original problem. This conclusion is elusive because it is possible to have problems such as linear integer programmes that theoretically have exponential complexity and there are algorithms available for solving them that work well in practice.

1.4.4 Literature review

Substantial literature on the optimal design of elastic trusses exists. Papers previously published have generally been directed to test problems involving bars,shafts,beams,plates and trusses of variable section. A variety of performance contraints have been placed on the design problem which serve to segregate the class of problems. The first literature review on stuctural design appeared in 1963 by Wasiutynsky and Brandt. They presented a review of literature though approximately 1962 citing 234 references. Shew and Prager continued with a thorough review until 1967 containing an additional 146 references. They were the last comprehensive reviews that have appeared in the growing field of stuctural optimization.

A discussion of optimal design and the mathematical programming techniques used in this field was presented by Prager (1971). Schmit (1971) pointed out the trend of current and future research in this field by making a distinction between single and multiple purpose structures. In 1973 Niordson and Pedersen gave a critical and analytical review of the literature that appeared between approximately 1968 and 1972. With primary emphasis on finite stuctural optimization a review by Venkayya (1976) and a survey by Krishnamoorthy and Mosi (1978) cover papers appeared until 1977.

1.4.5 Books review

The most substantial text considering a variety of performance constraints is the proceedings of a IUTAM symposium held in Warsaw in 1973 edited by Sawczuck and Mroz. Prager presented in 1974 a series of lectures that provide a discerning introduction to the field of optimality criterion. Haug and Arora published in 1979 a text on numerical methods of structural optimization and have solved the minimum weight design problem under a number of different type of constraints by using a gradient projection technique.

In 1981 a book edited by Haug and Cea as proceedings of a symposium held in Iowa City in distributed parameter optimization has a special emphasis both in shape optimal design and design sensitivity analysis although it includes a chapter in finite dimensional structural optimization. The Russian school seems primarly interested in analytical methods of structural optimization as reported in a monograph published by Banichuk (1980).

1.4.6 Present research

One particular aspect that appears to be significant is the bilinear nature of some of the constraint equations and the mathematical implications of this fact. Our aim was to seek distinguishing features in the optimization problem of hyperstatic trusses of fixed topology which would enable statements of uniqueness to be made about the solution under a combined set of constraints on deflections stress and buckling loads. The ultimate goal of this work is to find the set of either continuous or discrete variables that minimize the volume of the truss. Due to the manifestation of nonconvexity a special emphasis is made on the development of algorithms that permi t ing a better handling of the nonlinearities of the problem will also ensure convergence to the global solution. Some relevant results on their comparative efficiency are also reported.

CHAPTER TWO

BILINEARLY CONSTRAINED PROBLEMS

2.1. Introduction

Bilinear constraints are generally associated with nonconvex behaviour but in special cases the resulting problem may remain "convex" in the sense of having a unique optimum. It is important to distinguish such special cases in order to avoid costly algorithms more suited to convex problems. In the following sections an attempt is made to classify various bilinear problems from this viewpoint.

The bilinerly constrained problem is also related to a class of mixed integer problems and thus complementarity programming problems so that solution methods may be applicable to both. These relationships are described below. For comparative purposes the original least volume truss problem is summarized in the prior instance.

2.2. Truss Optimization Problem

The minimum volume design of indeterminate structures of fixed topology is a particular case of the broader class of bilinear programs bilinearly constrained.

$$\min_{x,y} \quad \underset{\sim}{c}^T \underset{\sim}{x} + \underset{\sim}{d}^T \underset{\sim}{y} + \underset{\sim}{x}^T \underset{\sim}{G} \underset{\sim}{y} \tag{2.1}$$

$$\underline{\text{st}} \quad \underset{\sim}{A}_i \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} + \underset{\sim}{W}_i \underset{\sim}{y} \geq \underset{\sim}{b}_i \tag{2.2}$$

$$\tilde{E}\,\tilde{x} \geq \tilde{e} \qquad \tilde{F}\,\tilde{y} \geq \tilde{f} \qquad (2.3)$$

where $\tilde{G} = \tilde{A} = \tilde{W} = \tilde{\emptyset} \qquad \tilde{c}=\tilde{\emptyset}$

$$\tilde{E}^T = [\ \tilde{E}_1^T \vdots \tilde{E}_2^T \vdots -\tilde{E}_2^T \vdots -\tilde{I}\ ]$$

$$\tilde{H} = [\ \tilde{H}_1^T \vdots -\tilde{H}_1^T\ ] \qquad \tilde{F} = [\ \tilde{I} \vdots -\tilde{I}\ ]$$

$$\tilde{e} = [\ \tilde{e}_1^T \vdots \tilde{e}_2^T \vdots \tilde{e}_2^T \vdots \tilde{e}_3^T\ ] \qquad i = 1,\ldots m$$

$$\tilde{f} = [\ \tilde{f}_1^T \vdots \tilde{f}_2^T\ ] \qquad \tilde{b} = [\ \tilde{b}_1^T \vdots -\tilde{b}_1^T\ ]$$

$\tilde{x}$ unrestricted and $\tilde{y} \geq \emptyset$

$$\tilde{b}\in R^m \quad ;\tilde{x}\in R^{n_x} \quad ;\tilde{y}\in R^{n_y} \quad ;\tilde{A}\in R^{m.n_x}$$

$$\tilde{W}\in R^{m.ny} \; ;\tilde{c}\in R^{n_x} \quad ;\tilde{d}\in R^{n_y} \quad ;\tilde{G}\in R^{n_x.n_y} ;\tilde{H}_i\in R^{n_x.n_y}$$

In special situations such as statical determinacy it was possible to show that the above problem is after all "convex". Such considerations are left to Chapter 3 where the related work of Bayer (1978) is summarized.

2.3. Problem Characteristics. Multiple Optimal Solutions

The optimization problem is therefore one with a linear objective function and possessing both linear and bilinear constraints. Some idea of the morphology of the feasible domain may be gained by examining the domain in projection on selected coordinate space.

The actual domain defined by $\tilde{\emptyset} \leq \tilde{x} \leq \tilde{e}_3$ stands in the n.l dimensional space (where l is the number of alternative loading conditions and n the number of bars) forming a domain with the shape of a hyper rectangle. In addition the linear constraint $\tilde{E}_1\tilde{x} \geq \tilde{e}_1$ define a set which intersects the domain along the n.1 axes giving a set of curtailled parallelopipeds. This set is now intersected by a linear manifold given by $\underset{\sim}{E}_2 \underset{\sim}{x} = \underset{\sim}{e}_2$ and the geometry is yet a convex polytope. It may be embedded in a n+n.1 dimensional coordinate where the new axes are defined by $\underset{\sim}{f}_1 \leq \underset{\sim}{y} \leq \underset{\sim}{f}_2$ giving a prismatic figure. Finally this is intersected by a subspace of bilinear equations this being a curved manifold on the account of the product term appearing in the equations becoming a nonconvex finite sized domain.

In order to illustrate that the above problem is not unlikely to have multiple solutions consider the following problem (2.4)

$$\text{Min} \quad y_1 + y_2 + y_3$$

$$\underline{\text{st}} \quad x_1 y_1 + x_3 y_3 = 0$$

$$3 x_1 y_1 + 1.2 x_2 y_2 - x_3 y_3 = 10$$

$$5 x_1 + x_2 + x_3 \leq 2.5$$

$$.1 \leq y_1 \leq 5. \; ; \; .1 \leq y_2 \leq 5. \; ; \; .1 \leq y_3 \leq 5.$$

$$0. \leq x_1 \leq 2.5 \; ; \; 0. \leq x_2 \leq 2.5 \; ; \; -2.5 \leq x_3 \leq 0.$$

This nonlinear problem presents three local optima namely

$x_1 = 0.0; x_2 = 2.5; x_3 = 0.0; y_1 = 0.1; y_2 = 3.33; y_3 = 0.1$

OF = 3.53

$x_1 = 0.5; x_2 = 2.5; x_3 = -2.5; y_1 = 0.5; y_2 = 3.0 \; ; y_3 = 0.1$

OF = 3.60

$x_1 = 1.0; x_2 = 0.0; x_3 = -2.5; y_1 = 2.5; y_2 = .1; y_3 = 1.0$

$OF = 3.60$

When we want to represent this nonconvex problem graphically we are quickly limited by the dimensions of the problem. In particular we cannot go beyond two equations in three unknowns at least not without placing ourselves in a subspace of the solution space. Fig 2.1 to 2.4 represent axonometric perspective views where all objective function values below a previously given level are printed . They correspond to feasible points of the domain when interpolating the x variables between any two local solutions. The contour of the OF corresponding to the same feasible points is drawn in Fig 2.5 and 2.6 .

Although having a low dimensionality when compared to real life structural synthesis problems this example is an useful testbed for statements about the efficiency of some of the algorithms considered in this work.

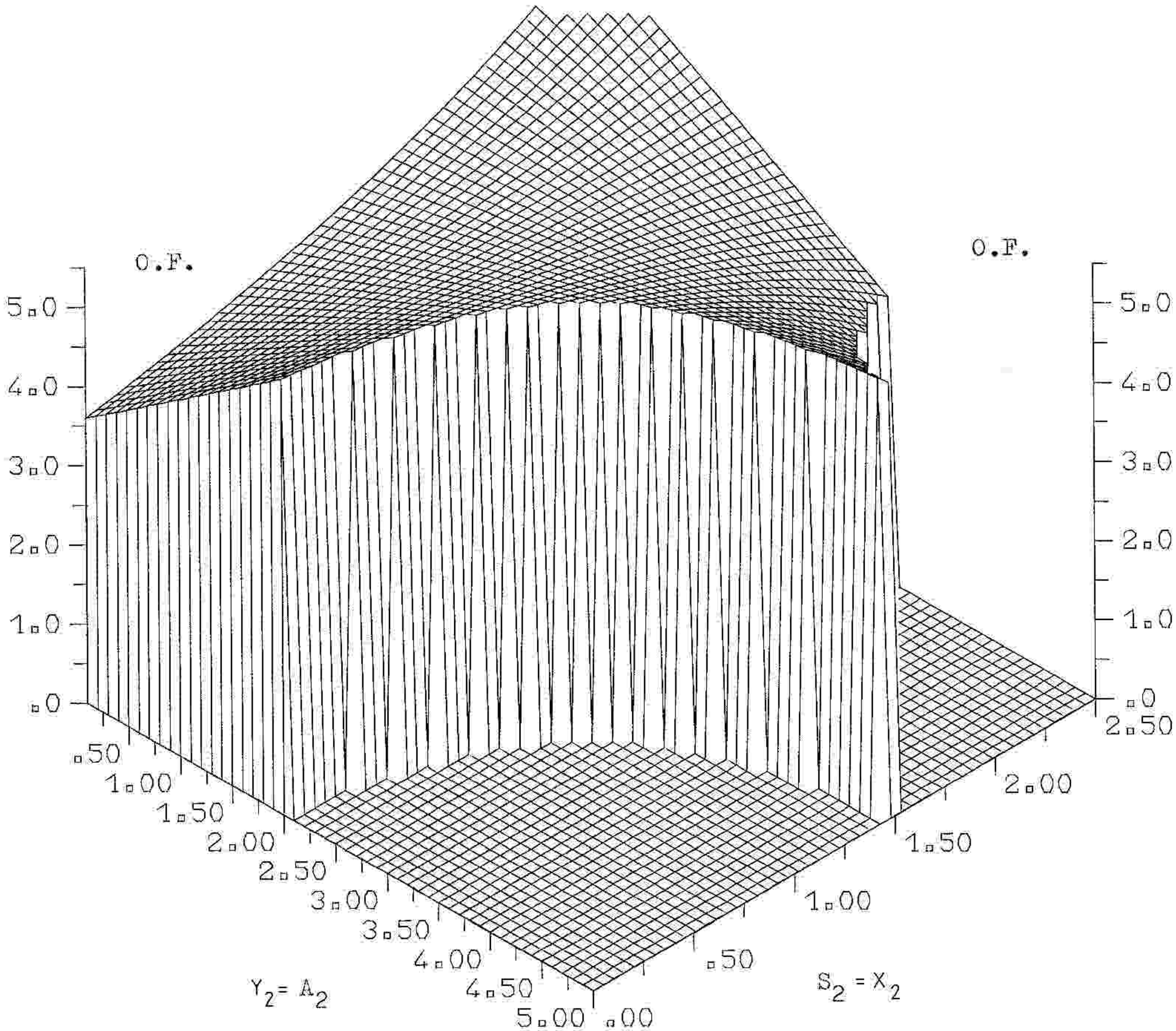


Fig. 2.1 Perspective view of the feasible volumes smaller 5.5 plotted for various A_2 and S_2 and keeping S_3 at its lower bound

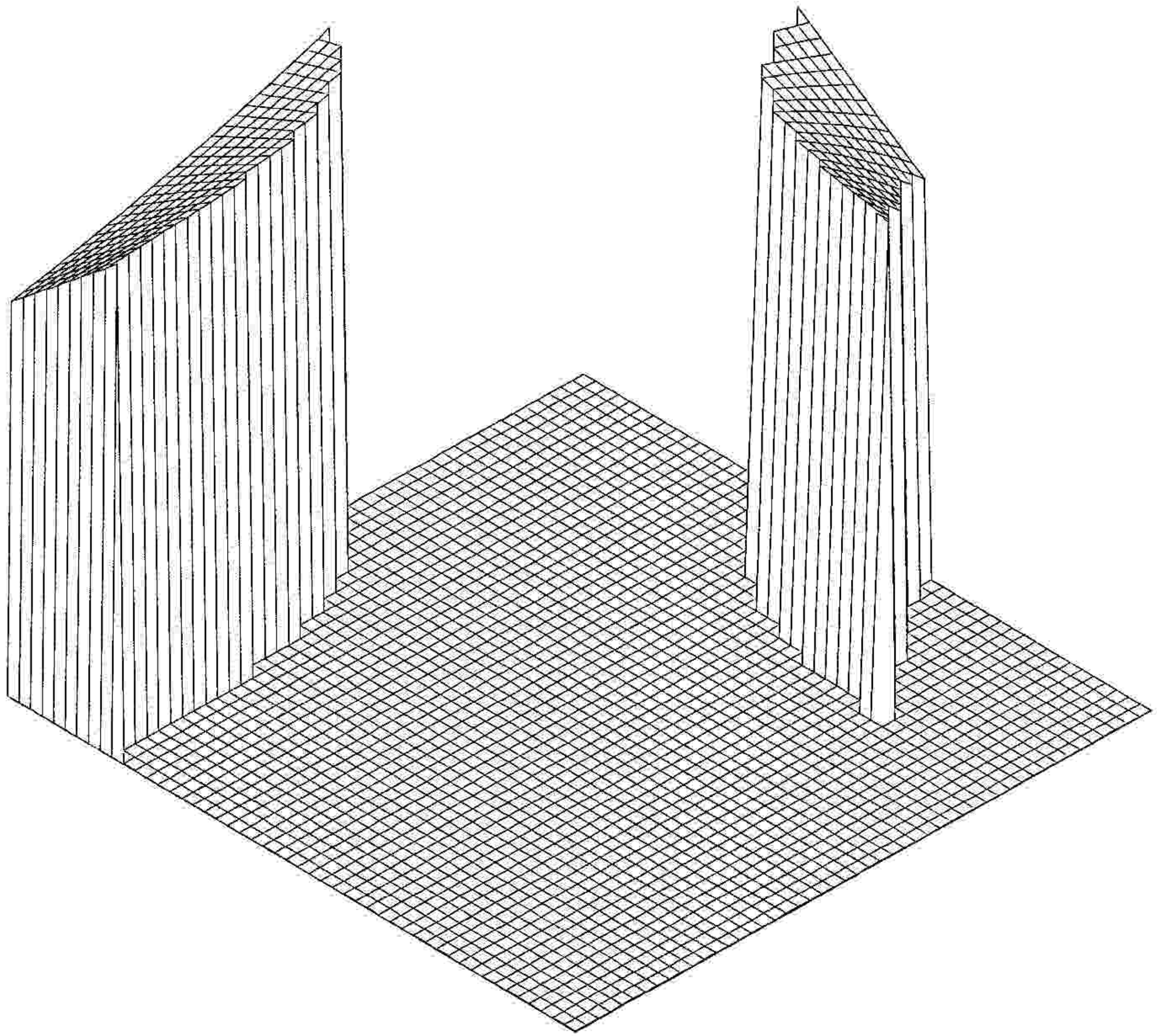


Fig. 2.2 Now the maximum volume is limited to 4.5 . The regions surrounding both local optima are no longer connected.

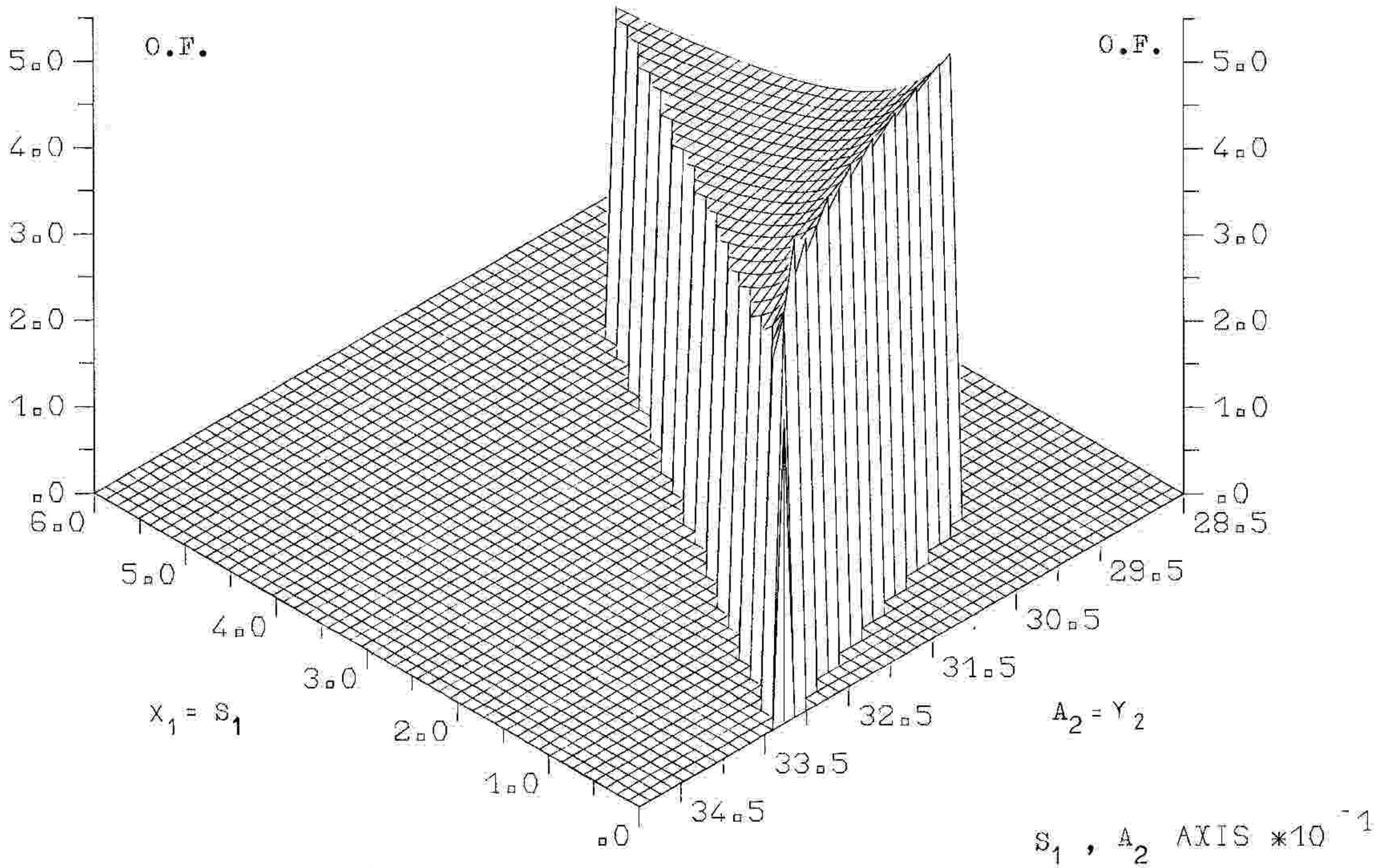


Fig. 2.3 Isometric view of the feasible volume smaller than 5.5 when varying A_2 and S_1 within their bounds and keeping A_3 and S_3 fixed.

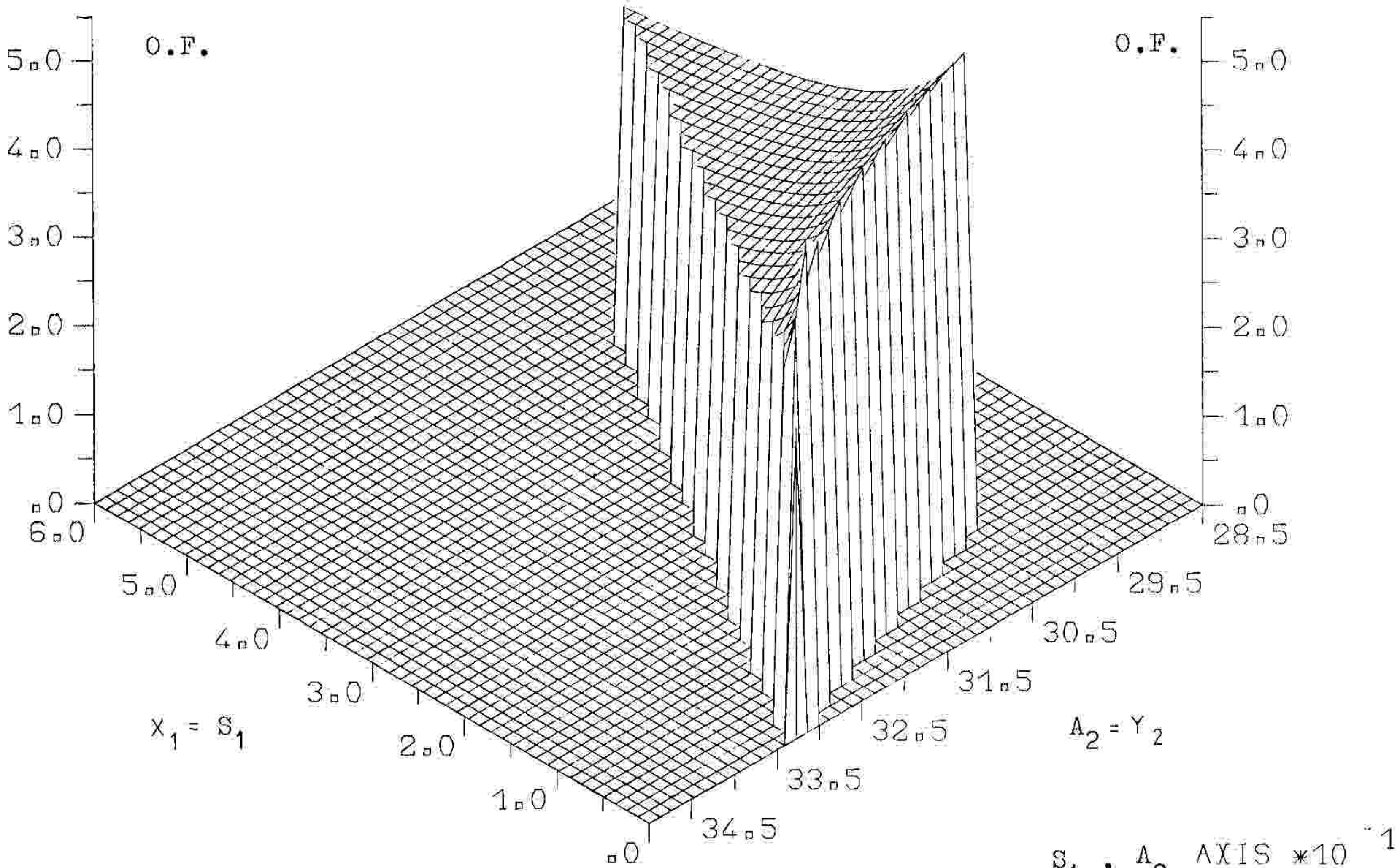


Fig. 2.3 Isometric view of the feasible volume smaller than 5.5 when varying A_2 and S_1 within their bounds and keeping A_3 and S_3 fixed.

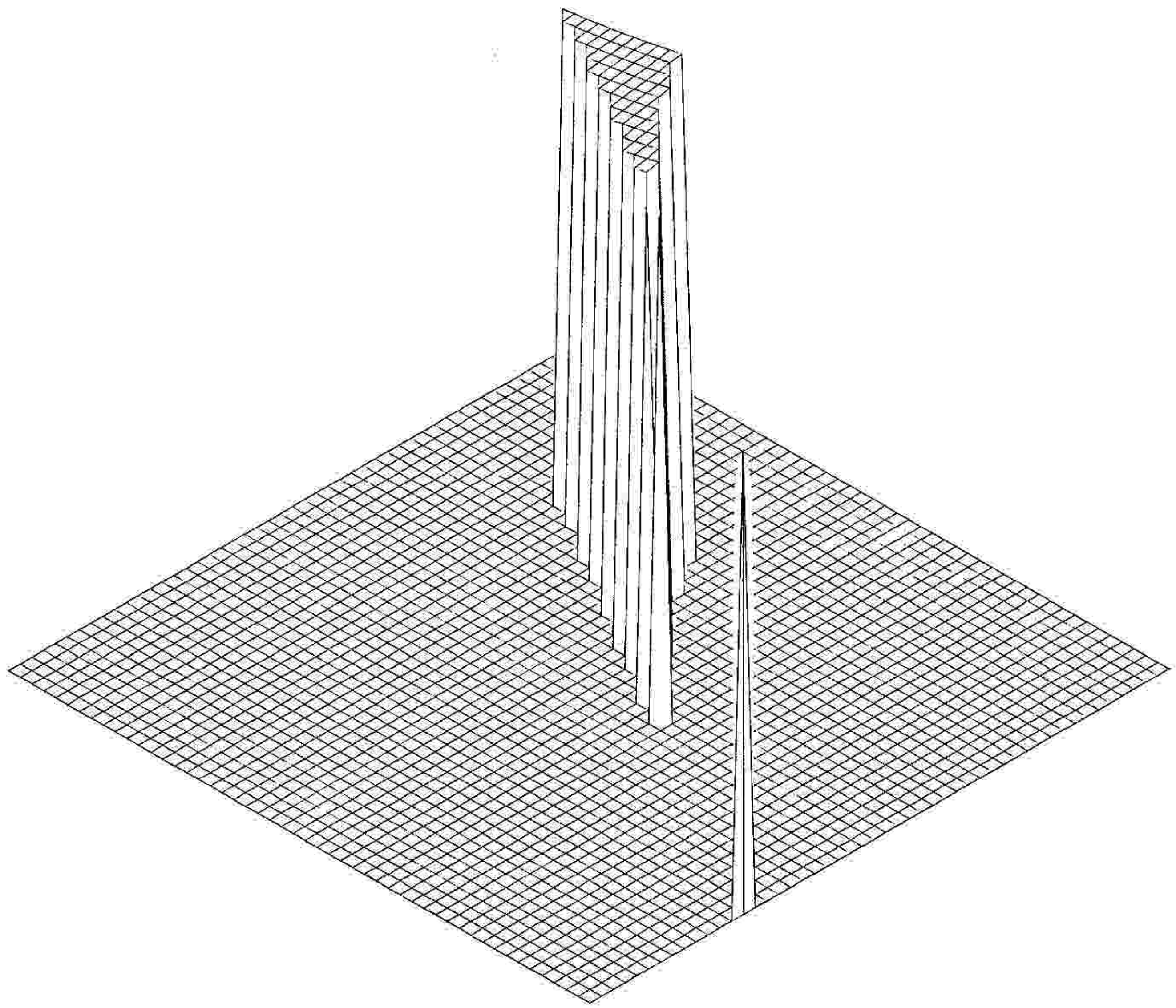


Fig. 2.4 Now the maximum volume is limited to 3.8 . The regions that contain the local optimum and the global optimum are separated.

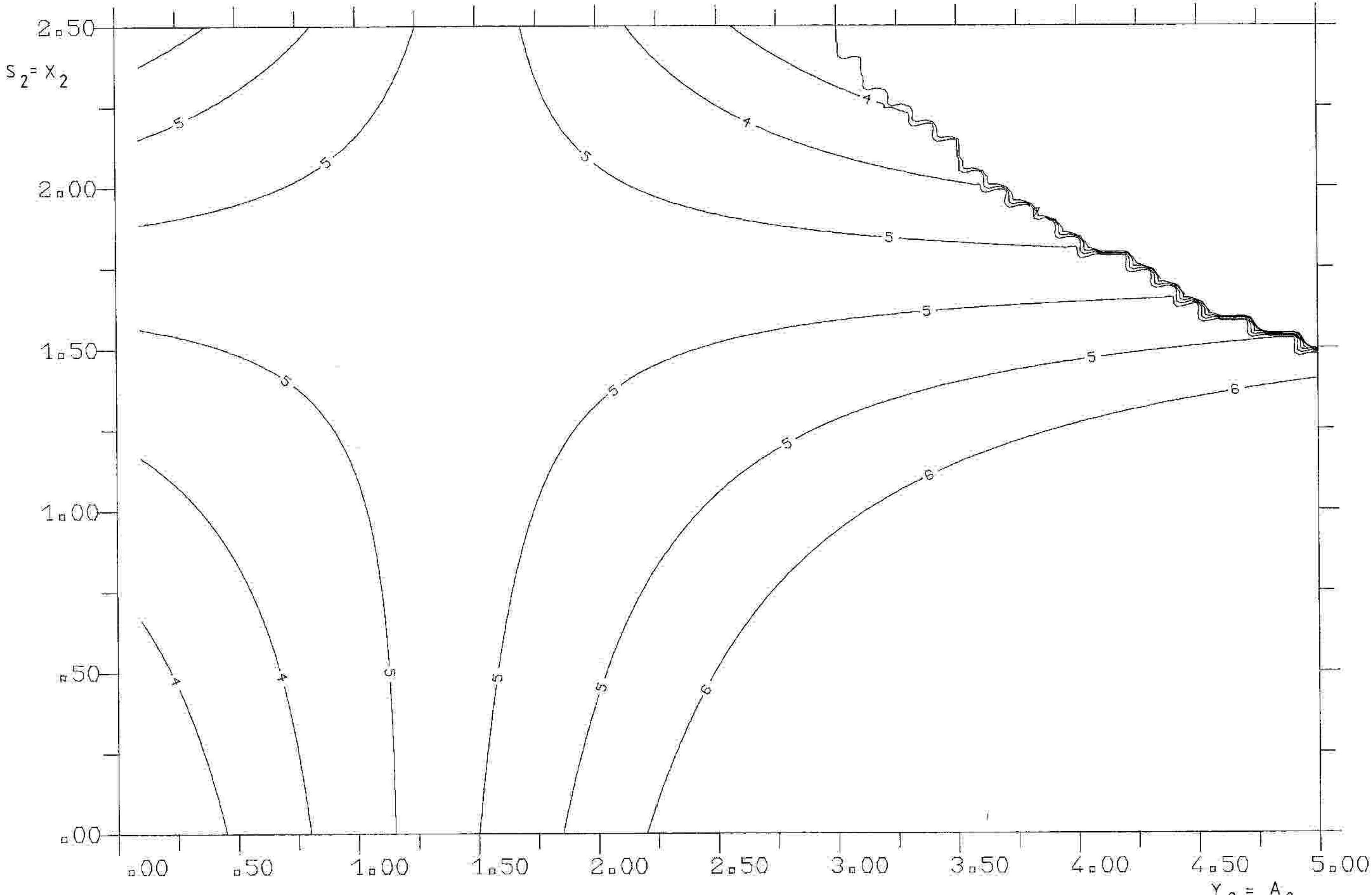


Fig. 2.5 Contours at 7 different levels between the O.F. values of 3.5 and 5.5 are drawn corresponding to the bird's eye view represented in Fig 2.1

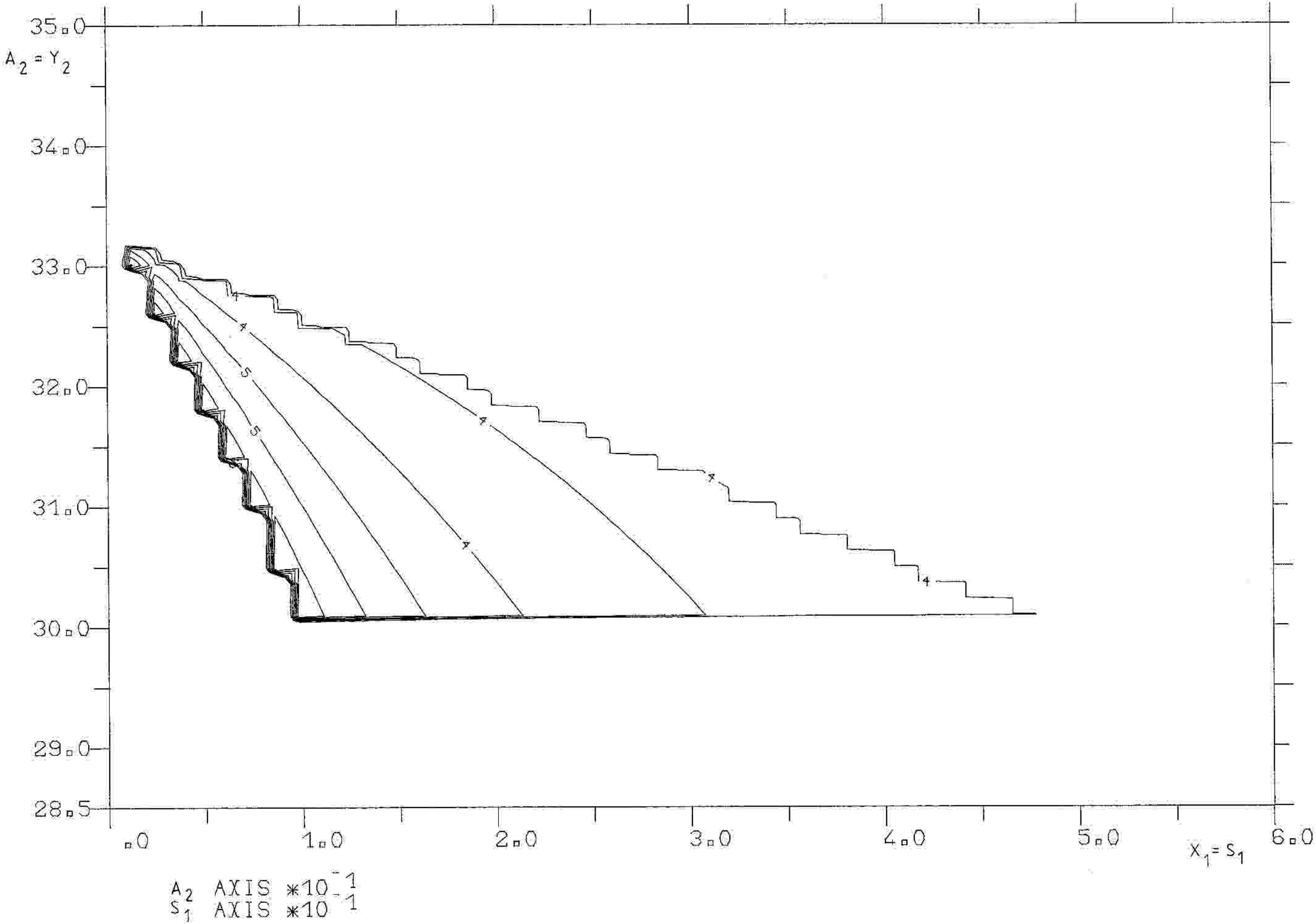


Fig 2.6 Contours at 7 different levels between the O.F. values of 3.8 and 5.4 are drawn corresponding to the perspective represented in Fig 2.3

The optimal design of a three bar truss gives physical meaning to this nonconvex problem. If we let $\tilde{y}$ represent the vector of member areas and $\tilde{x}$ the vector of member stresses the matrix $\tilde{A}$ of the constants appearing in the bilinear equations represents the direction cosine matrix. (see Chapter 3)

$$\tilde{A} = \begin{bmatrix} 1 & . & 1 \\ 3 & 1.2 & -1 \end{bmatrix}$$

A compatibility type of constraint for this problem would be given by

$$x_1 - 3.33\ x_2 - x_3 = 0 \qquad (2.5)$$

The remaining linear inequality in problem (2.4) represents a displacement restraining the structure. Adding the equality (2.5) to the mathematical program (2.4) we obtain a single solution given by

$$x_1 = 0.80; x_2 = 0.99; x_3 = -2.5; y_1 = 3.08; y_2 = 0.1; y_3 = 0.99$$

$$OF = 4.169$$

In Fig 2.7 and 2.8 the effect of the introduction of this hyperplane is shown. In the next Chapter further examples of trusses exhibiting multiple optimality are presented. A local sufficiency test checking whether a previously determined local optima is unique is also applied. But here we proceed by giving another case of multiplicity in the number of solutions that may occur in structural synthesis.

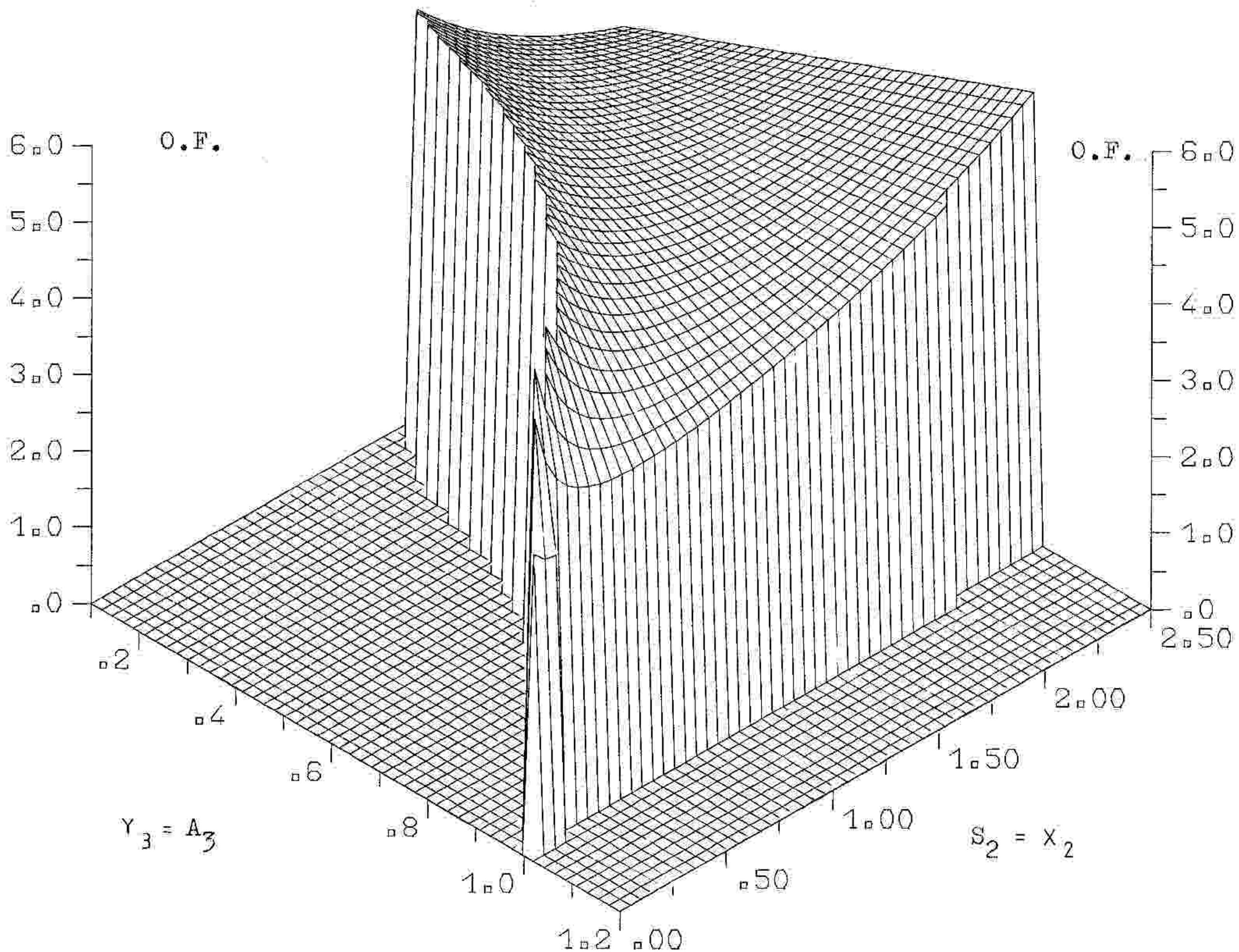


Fig. 2.7 Perspective view of the feasible volumes smaller than 6.0 when A_3 and S_2 are varied within their rectangle of bounds.
S_3 is fixed at the solution value obtained at both optima.

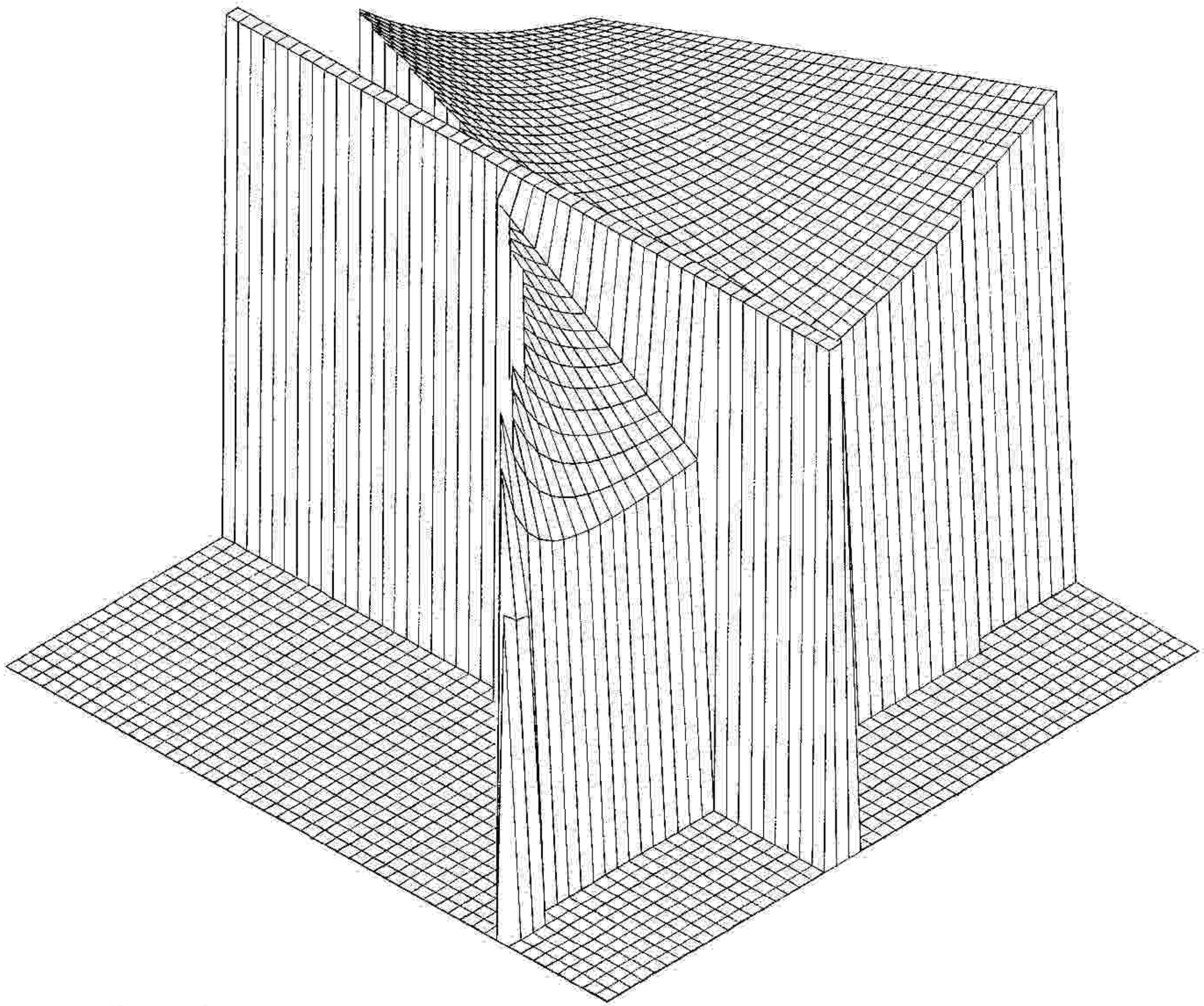


Fig 2.8 Effect of a Compatibility constraint introduced in the bilinearly restricted problem. The intersection of the domain represented in Fig 2.7 and this plane gives the feasible set of the newly defined problem

The following simple examples illustrate some of the problems often encountered in grillage optimal design. A grillage is usually made up of orthogonal beams loaded normal to its plane. The geometry of the structure is assumed to be known including the number of beams span length and support conditions.

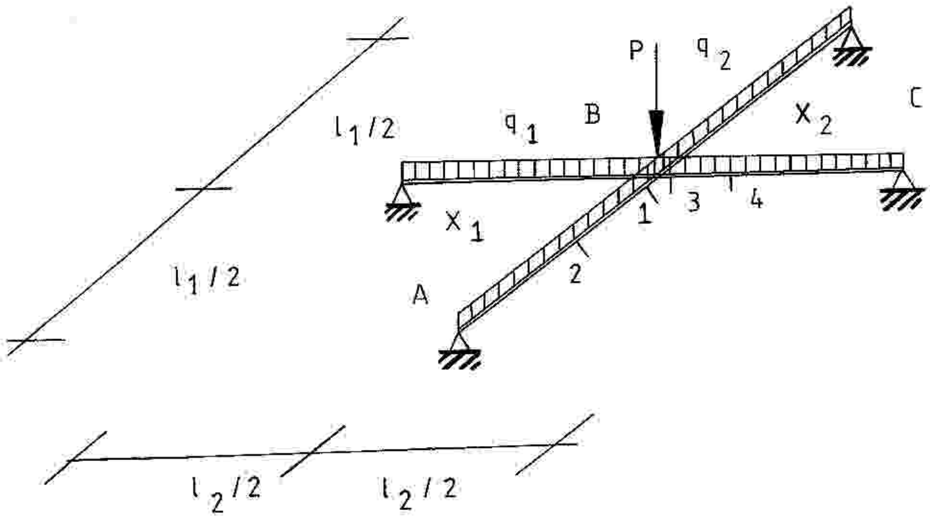


Fig 2.9 Grillage

Constraints are related with bending stresses at node B m_1, m_3 and in critical sections between nodes AB m_2 and BC m_4. The optimal design problem (2.6) is to find the cross sectional areas y_1, y_2 such that

$$\text{Min} \quad l_1 y_1 + l_2 y_2$$

$$\underline{\text{st}} \quad s_{1l} \leq s_1 = m_1/w_1 \leq s_{1u}$$

$$s_{2l} \leq s_2 = m_2/w_1 \leq s_{2u}$$

$$s_{3l} \leq s_3 = m_3/w_2 \leq s_{3u}$$

$$s_{4l} \leq s_4 = m_4/w_2 \leq s_{4u}$$

By using the nodal stiffness method we first obtain the bilinear equation for the deflection at B ,d_B

$$6\ E\ (I_1/(l_1/2)^3 + I_2/(l_2/2)^3)\ d_B = \lambda \quad (2.7)$$

in which E is the modulus of elasticity and $\underset{\sim}{\lambda}$ the vector of nodal forces. The bending moments m_1 and m_3 are computed by

$$m_1 = m_{01} - (3\ E\ I_1/(l_1/2)^2)\ d_B$$

$$m_3 = m_{03} - (3\ E\ I_2/(l_2/2)^2)\ d_B$$

where m_{01} and m_{03} are end moments corresponding to a propped cantilever. For the given bending moments the shear forces

$$F_A = F_{0A} - m_1/(l_1/2)$$

$$F_C = F_{0C} - m_3/(l_2/2)$$

can be determined and we may find the location of critical section 2 and 4 and the corresponding bending moments. The design spaces for the two following cases is shown in Fig 2.10 and 2.11

CASE 1

Sandwich beam possessing the following cross sectional properties

$q_1 = 0.8 \quad q_2 = 0.8 \quad P = 0.$

$l_1 = 30.0 \quad l_2 = 31.0$

$w_j = .5\, y_j \quad I_j = .3\, y_j \quad j = 1,2$

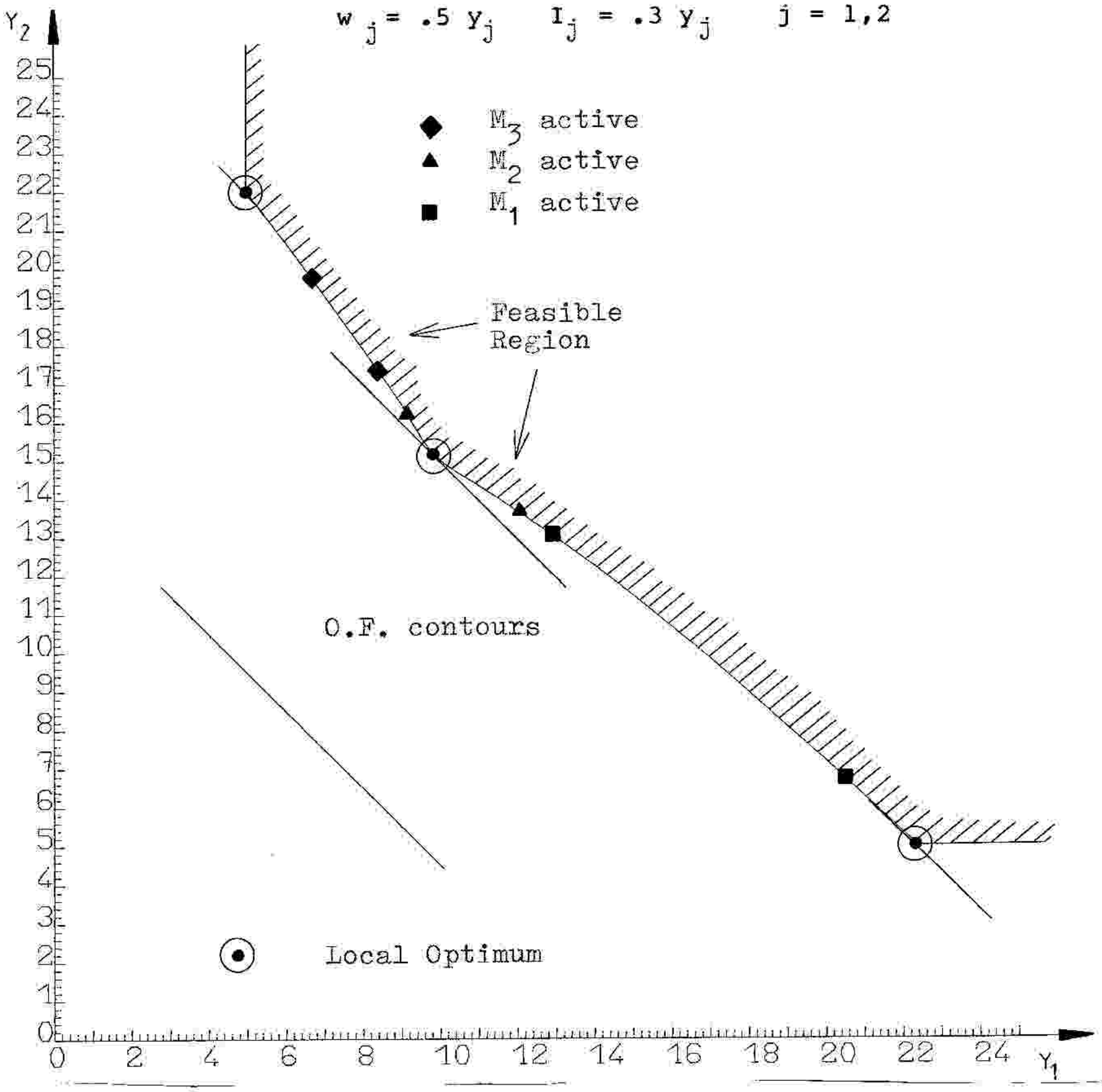


Fig 2.10 Design space - case 1

Here

$$\lambda = 5\,(q_1 l_1 + q_2 l_2)/8 + P$$

and

$$s_l = [-15. \ -15. \ -15. \ -15.] \qquad s_u = -s_l$$

$$y_l = [\ 5. \quad 5.]$$

Critical section 2 and 4 are located at a distance of F_A/q_1 from A and F_C/q_2 from C respectively (but not exceeding the respective member length)

$$m_2 = F_A^2/(2\ q_1)$$

$$m_4 = F_C^2/(2\ q_2)$$

Three local optima are obtained corresponding to

$y_1 = 5.0$	$y_2 = 22.0$	OF = 832.0	
$s_1 = -9.34$	$s_2 = 14.28$	$s_3 = -15.00$	$s_4 = 15.00$
$y_1 = 9.9$	$y_2 = 15.1$	OF = 765.1	
$s_1 = -14.84$	$s_2 = 15.00$	$s_3 = -15.00$	$s_4 = 15.00$
$y_1 = 22.4$	$y_2 = 5.0$	OF = 826.1	
$s_1 = -15.00$	$s_2 = 15.00$	$s_3 = -6.30$	$s_4 = 13.02$

CASE 2

Assume the following relationships

$$q_1 = q_2 = 1.2 \qquad P = 25.0$$

$$l_1 = 21. \qquad l_2 = 24.$$

$$w_j = 5.\ y_j \quad I_j = 25.\ y_j \qquad j = 1,2$$

$s_l = [-5.\ -5.\ -5.]$ $\qquad$ $s_u = -s_l$

$y_l = [\ 1.\ \ 1.]$

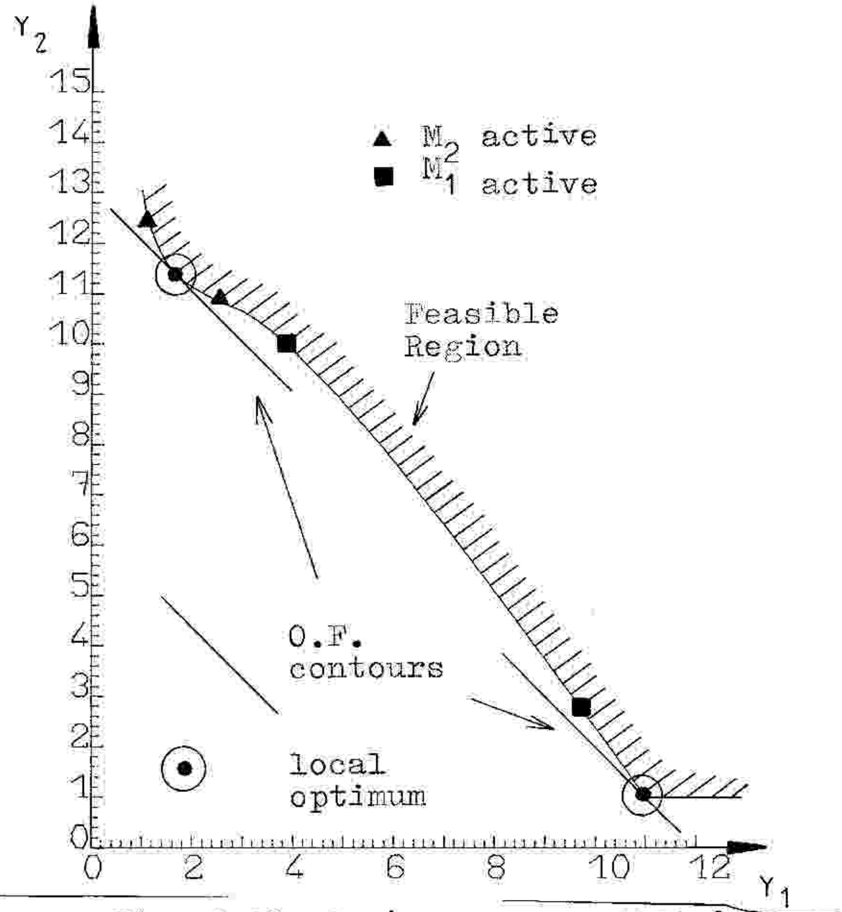


Fig 2.11 Design space - case 2

In this case we would obtain two distinct local minima

$y_1 = 1.70 \qquad y_2 = 11.32 \qquad OF = 307.5$

$s_1 = -4.70 \qquad s_2 = 5.00 \qquad s_3 = -4.70 \qquad s_4 = 4.70$

$y_1 = 11.00 \qquad y_2 = 1.00 \qquad OF = 255.0$

$s_1 = -5.00 \qquad s_2 = 5.00 \qquad s_3 = 0.24 \qquad s_4 = 4.18$

In the remainder of this Chapter we will review some of the extensions of LP that exhibit the closest resemblance to the nonlinearities encountered in the least volume design.

2.4. Linear Programming

The linear programming problem (LP) in its inequality standard form is defined by

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{x} \tag{2.8}$$

$$\underline{\text{st}} \quad \underset{\sim}{A}\,\underset{\sim}{x} \geq \underset{\sim}{b} \tag{2.9}$$

$$\underset{\sim}{x} \geq \underset{\sim}{\emptyset} \tag{2.10}$$

where $\underset{\sim}{x} \in R^n; \underset{\sim}{c} \in R^n; \underset{\sim}{b} \in R^m$ and $\underset{\sim}{A} \in R^{m\times n}$ are given vectors and matrix. The Simplex algorithm (and its revised versions) due to Dantzig (1963) solve this class of problems very efficiently.

Problems capable of being expressed mathematically in the form of a LP are many and varied. The economic interpretation which is given here is not associated with a particular problem but with a terminology that has become quasi-universal language paralleling the more abstract statements used in this work. Consider an entreprise wich has in the general case several "activities" each one using a certain number of "resources". These may be products originated simu ltaneously by the several activities.

Each "commodity" is finally demanded in a quantity b_i . Moreover to each one of the n activities j is attached a cost c_j depending on the intensity of this activity. The implementation of these activities is accompanied by a cost to be minimized while satisfying all the demands without exceeding any availability. The a_j are column vectors the m components of which representing the proportions that the activity j taken in its reference state uses the m resources i . The intensity of the activity j may be represented by the single parameter x_j which will then be called the level of activity j . A set of values of the x_j defines the program of the enterprise under consideration.

2.5. Duality in LP

Associated with every LP called the primal there is another LP called its dual. These problems possess very interesting and closely related properties: If the unique optimal solution to any one is known the optimal solution can readily be obtained. A solution can be found by solving either the primal or the dual whichever easier. Dual LP

$$\text{Max} \quad \underset{\sim}{b}^T \underset{\sim}{y} \tag{2.11}$$

$$\underline{\text{st}} \quad \underset{\sim}{A}^T \underset{\sim}{y} \leq \underset{\sim}{c} \tag{2.12}$$

$$\underset{\sim}{y} \geq \underset{\sim}{\emptyset} \tag{2.13}$$

where $\underset{\sim}{y} \in R^m$ is a vector

It can be seen that

$$\text{Max } \underset{\sim}{b}^T \underset{\sim}{y} = -\text{Min } (-\underset{\sim}{b}^T \underset{\sim}{y})$$

Let us reconsider the economic interpretation of the last section. The implementation of the activities may be accompanied either by a cost to be minimized or a profit to be maximized. Formally the problem may always be reduced to either case to consider a profit as a negative cost. In order to assume homogeneity in the dual relations y must be given the significance of a unit price. The dual problem may then be expressed as :

Given a unit cost c_j for each of the n activities and a demand b_i for each of the m resources i what must be the unit price of each resource b_i such that the total value of the resources produced by j at level 1 should be less than or equal to the cost and the total value of the demanded commodity is maximal. The dual variables y_i will then be called prices. Consider an optimal basis for the primal. The solution of the dual problem may then be interpreted as a system of shadow prices which the resources i must have so that the total value of the "goods" produced by each of the m activities of the basis should be equal to the cost of this activity .

2.6. Quadratic Programming

If the objective function is extended to include quadratic terms ie

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{x} + 1/2\, \underset{\sim}{x}^T \underset{\sim}{D}\, \underset{\sim}{x} \qquad (2.14)$$

where $\underset{\sim}{D} \in R^{n \times n}$ is a positive semi-definite square matrix the mathematical program is termed quadratic programming (QP).

There are several efficient algorithms designed for computing a solution satisfying the Khun Tucker (K-T) conditions for this convex QP. The obtaining of a K-T point is done roughly speaking as easily as by solving a LP by the Simplex method. Since this problem is convex the solution is unique and any local solution is also the global optimum under convexity.

For nonconvex QP that are defined whenever $\tilde{D}$ is indefinite a K-T point is not necessarily a global or even local minimum and it usually requires substantially more effort to determine a global minimum. Most algorithms capable of solving a nonconvex QP employ some combinatorial principles such as Branch and Bound and Cutting Plane strategies.

2.7. Bilinear Programming [Konno (1976)]

Among general QP a natural extension of the LP is to the case where $\tilde{c}$ is not fixed but can be chosen from a certain polyhedral convex set

$$\underset{\tilde{x}}{\text{Min}} \left\{ \underset{\tilde{c}}{\text{Min}} \; \tilde{c}^T \tilde{x} \right\} \qquad \left(\text{or} \quad \underset{\tilde{x},\tilde{c}}{\text{Min}} \; \tilde{c}^T \tilde{x}\right) \qquad (2.15)$$

$$\underline{\text{st}} \quad \tilde{A}\,\tilde{x} \geq \tilde{b} \qquad \tilde{C}\,\tilde{c} \geq \tilde{1} \qquad (2.16)$$

$$\tilde{x} \geq \tilde{\emptyset} \qquad \tilde{c} \geq \tilde{\emptyset} \qquad (2.17)$$

If we vary $\tilde{b}$ in the polyhedral convex set together with $\tilde{c}$ the problem is still reducible to the previous one.

This extended linear program is a special case of the more general problem having linear terms in the objective function together with bilinear terms

$$\min_{\underset{\sim}{x},\underset{\sim}{y}} \quad \underset{\sim}{c}^T \underset{\sim}{x} + \underset{\sim}{d}^T \underset{\sim}{y} + \underset{\sim}{x}^T \underset{\sim}{G}\, \underset{\sim}{y} \tag{2.18}$$

$$\underline{\text{st}} \quad \underset{\sim}{E}\, \underset{\sim}{x} \geq \underset{\sim}{e} \qquad \underset{\sim}{F}\, \underset{\sim}{y} \geq \underset{\sim}{f} \tag{2.19}$$

$$\underset{\sim}{x} \geq \underset{\sim}{\emptyset} \qquad \underset{\sim}{y} \geq \underset{\sim}{\emptyset} \tag{2.20}$$

called bilinear programming (BLP) . It can be reduced to a general QP where $\underset{\sim}{D}$ is a copositive matrix (has symmetric eigenvalues) what makes this problem nonconvex. Since the feasible region defined is a polyhedral convex set methods for finding its optimal solution may use with advantage the knowledge that the optimal point will lie at a vertex of the two separable constraint sets. However it should be noted that whichever strategy used the problem is a nonconvex one and the solution method must be enumerative.

2.8. Generalized Linear Program [Dantzig (1963)]

When b is varied in the initial LP in a convex (usually polyhedral) set then the problem becomes a generalized linear program (GLP) solvable by a Simplex based algorithm (column by column sequential method). In the more general case of Wolfe's problem the objective function is linear and each column of the constraint set is a vector which is to be chosen from a convex domain of its own.

This is an "optimistic" strategy because it seeks the optimal solution feasible for some $a_j \in P_j$ j=1,...,n where P_j is a convex polyhedron.

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{x} \tag{2.21}$$

$$\underline{\text{st}} \quad \underset{\sim}{A}\, \underset{\sim}{x} = a_1 x_1 + a_2 x_2 + \dots + a_n x_n \geq \underset{\sim}{b}$$

$$\text{for } \underline{\text{some}}\ a_j \in P_j \quad \text{and } x \geq \emptyset \qquad (2.22)$$

Here the problem is convex despite the appearance of terms such as $a_j x_j$ (where both a_j and x_j are unknowns). This is most easily realized when reformulating the problem for a Simplex solution.

2.9 Inexact Programming [Soyster (1973)]

If a vector $\underset{\sim}{c}$ is constrained to lie in a given convex set C a solution of the following program will be optimal against the worst possible outcome of $c \in C$

$$\underset{\underset{\sim}{x}}{\text{Min}} \left\{ \underset{\underset{\sim}{c} \in C}{\text{Max}} \ \underset{\sim}{c}^T \underset{\sim}{x} \right\} \qquad (2.23)$$

$$\underline{\text{st}} \quad \underset{\sim}{A} \underset{\sim}{x} \geq \underset{\sim}{b} \qquad (2.24)$$

$$\underset{\sim}{x} \geq \underset{\sim}{\emptyset}$$

The solution technique approach is called inexact programming (IEP) . Although superficially equivalent to Bilinear Programming this problem is convex on account of Min Max instead of Min Min. Soyster studied a generalization where the columns of $\underset{\sim}{A}$ and the vector $\underset{\sim}{b}$ were assumed to lie in a given set. Although the feasible region so defined is not of the standard form found in LP this more general problem could be reduced to a problem having the form of the IEP defined. It is interesting to note that the latter is equivalent to the infinitely constrained optimization problem used to idealize large scale systems.

$$\underset{\underset{\sim}{x}}{\text{Min}} \quad t \qquad (2.25)$$

$$\underline{st} \quad \underset{\sim}{A}\,\underset{\sim}{x} \geq \underset{\sim}{b} \qquad t \geq \underset{\sim}{c}^T \underset{\sim}{x} \tag{2.26}$$

$$\text{for } \underline{\text{all}}\ \underset{\sim}{c} \in C \quad \text{and} \quad \underset{\sim}{x} \geq \emptyset \tag{2.27}$$

Thuente (1980) had shown that a dual to the GLP is a IEP in the usual sense of LP having the form

$$\text{Max} \quad \underset{\sim}{b}^T \underset{\sim}{w} \tag{2.28}$$

$$\underline{st} \quad \underset{\sim}{A}^T \underset{\sim}{w} \leq \underset{\sim}{c} \tag{2.29}$$

$$\text{for } \underline{\text{all}}\ a_j \in P_j \quad \text{and} \quad \underset{\sim}{w} \geq \emptyset \tag{2.30}$$

2.10. Variable Factor Programming [Geoffrion (1972)]

A generalization of GLP is Wilson's variable factor programming (VFP)

$$\text{Min} \quad \underset{\sim}{x}^T f(z) \tag{2.31}$$

$$\underline{st} \quad \underset{\sim}{A}\,\underset{\sim}{x} \geq \underset{\sim}{b} \qquad \underset{\sim}{Z}\,\underset{\sim}{x} \geq \underset{\sim}{c} \tag{2.32}$$

$$\underset{\sim}{x} \geq \underset{\sim}{\emptyset} \qquad \underset{\sim}{z} \geq \underset{\sim}{\emptyset} \tag{2.33}$$

where $\underset{\sim}{z}$ is a m-vector and $\underset{\sim}{Z}$ is a n.m vector $[z^1 \ldots z^n]$ and each component of f is assumed strictly concave. This problem gives a unique solution and has been solved by Geoffrion as an extension of Bender's decomposition method.

2.11. Linear Integer Programming

A different type of extension of LP is to the case where $\underset{\sim}{x}$ is allowed to take only integer values

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{x} \tag{2.34}$$

$$\text{st} \quad \underset{\sim}{A}\,\underset{\sim}{x} \geq \underset{\sim}{b} \qquad (2.35)$$

$$\underset{\sim}{x} \geq 0 \text{ and integer}$$

called linear integer programming (LIP) .

2.12. 0-1 Linear Integer Programming

If we let L to be an upper integer bound for $\log_2 x_j$ $j=1,\ldots,n$ then it is possible to write

$$x_j = \sum_{k=0}^{l} y_k^j 2^k \qquad y_k^j \in B = \{0,1\} \qquad (2.36)$$

for any feasible integer x .

Therefore we can reduce a LIP to a linear program in 0-1 variables called binary programming or 0-1 linear integer programming

$$\text{Min} \quad \underset{\sim}{g}^T \underset{\sim}{y} \qquad (2.37)$$

$$\text{st} \quad \underset{\sim}{E}\,\underset{\sim}{y} \geq \underset{\sim}{e} \qquad (2.38)$$

$$\underset{\sim}{y} \in B^{n.L} \qquad (2.39)$$

This program can be shown to be equivalent to the real concave quadratic program (QCM)

$$\text{Min} \quad (\underset{\sim}{g}^T + \mu \underset{\sim}{a}^T)\,\underset{\sim}{y} - \mu\,\underset{\sim}{y}^T \underset{\sim}{y} \qquad (2.40)$$

$$\text{st} \quad \underset{\sim}{E}\,\underset{\sim}{y} \geq \underset{\sim}{e} \qquad (2.41)$$

$$\underset{\sim}{y} \leq \underset{\sim}{a} \qquad \underset{\sim}{y} \geq \underset{\sim}{0} \qquad (2.42)$$

where $\underset{\sim}{a}$ is a n.L vector of 1's and μ is a sufficientely large positive number. This nonconvex minimization problem is equivalent to a Complementarity Programming Problem (CCP) . In fact (QCM) has an optimal solution y^* iff there exists a vector (u^*, z^*, v^*) that is an optimal solution to the following program [Gianessi and Niccolucci (1976)]

$$\text{Min} \quad (g^T + \mu a^T)\, y - \mu\, y^T y \tag{2.43}$$

st
$$-2\mu y + g + \mu a - z^T E + t - u = 0 \tag{2.44}$$

$$E y - e - v = 0 \tag{2.45}$$

$$-y + a - w = 0 \tag{2.46}$$

$$y^T u = 0 \tag{2.47}$$

$$z^T v = 0 \tag{2.48}$$

$$t^T w = 0 \tag{2.49}$$

$$y, z, t, u, v, w \geq 0$$

From (2.44) we have

$$y = 1/2\, (g + \mu a - z^T E + t - u) \tag{2.50}$$

the OF of the latter program becomes

$$(g^T + \mu a^T)\, y - 1/2\, (g + \mu a - z^T E + t - u)^T y =$$

$$= 1/2\, (g^T + \mu a^T)\, y + 1/2\, z^T E y - 1/2\, t^T y - u^T y \tag{2.51}$$

But

$$u^T y = 0 \tag{2.52}$$

$$t^T y = t^T (a - w) = t^T a - t^T w \tag{2.53}$$

$$= t^T a \tag{2.54}$$

$$\underset{\sim}{z}^T \underset{\sim}{E} \underset{\sim}{y} = \underset{\sim}{z}^T (\underset{\sim}{e} + \underset{\sim}{v}) = \underset{\sim}{z}^T \underset{\sim}{e} + \underset{\sim}{z}^T \underset{\sim}{v} \tag{2.55}$$

$$= \underset{\sim}{z}^T \underset{\sim}{e} \tag{2.56}$$

(2.51) gives

$$1/2 \ (\underset{\sim}{g}^T + \mu \underset{\sim}{a}^T) \underset{\sim}{y} + 1/2 \ \underset{\sim}{e}^T \underset{\sim}{z} - 1/2 \ \underset{\sim}{a}^T \underset{\sim}{t} \tag{2.57}$$

That is a linear expression in the variables y,z and t. If we make

$$\bar{\underset{\sim}{y}}^T = (\underset{\sim}{y}^T , \underset{\sim}{z}^T , \underset{\sim}{t}^T) \tag{2.58}$$

$$\bar{\underset{\sim}{u}}^T = (\underset{\sim}{u}^T , \underset{\sim}{v}^T , \underset{\sim}{w}^T) \tag{2.59}$$

$$\bar{\underset{\sim}{E}}^T = \begin{bmatrix} 2\mu \underset{\sim}{I} & \underset{\sim}{E} & -\underset{\sim}{I} \\ -\underset{\sim}{E} & \underset{\sim}{\emptyset} & \underset{\sim}{\emptyset} \\ \underset{\sim}{I} & \underset{\sim}{\emptyset} & \underset{\sim}{\emptyset} \end{bmatrix} \tag{2.60}$$

$$\bar{\underset{\sim}{e}}^T = (\underset{\sim}{g}^T + \mu \underset{\sim}{a}^T , -\underset{\sim}{e}^T , \underset{\sim}{a}^T) \tag{2.61}$$

$$\bar{\underset{\sim}{g}}^T = 1/2 \ (\underset{\sim}{g}^T + \mu \underset{\sim}{a}^T , \underset{\sim}{e}^T , \underset{\sim}{a}^T) \tag{2.62}$$

We obtain the following CCP

$$\text{Min} \quad \bar{\underset{\sim}{g}}^T \bar{\underset{\sim}{y}} \tag{2.63}$$

$$\underline{\text{st}} \quad \bar{\underset{\sim}{E}} \bar{\underset{\sim}{y}} + \bar{\underset{\sim}{u}} = \bar{\underset{\sim}{e}} \tag{2.64}$$

$$\bar{\underset{\sim}{y}}^T \bar{\underset{\sim}{u}} = \underset{\sim}{\emptyset} \tag{2.65}$$

$$\bar{\underset{\sim}{y}}, \bar{\underset{\sim}{u}} \geq \underset{\sim}{\emptyset}$$

CHAPTER THREE

TRUSS OPIMIZATION WITH CONTINUOUS DESIGN VARIABLES

3.1. Statics and Kinematics of the Structural Model

From the reader's point of view it may be advantageous to separate the structural and mathematical aspects of the design problem. We therefore begin by summarizing the standard equations of linear elastic structural theory in order to conveniently refer to them in the sequel. When a structure is discretized into a finite element model its elements and nodes are oriented defining its topology. In analytical mechanics physical events are in general described through discrete coordinates to which all relevant quantities are refered called state variables. There is a second set of variables called design variables that describes the system according to the designer's purpose.

3.2. Nodal (Matrix Displacement) Method

From Hooke's law the member distortions u_j can be expressed in terms of the member forces n_j

$$u_j = [\, l_j / (E_j a_j) \,]\, n_j = f_j\, n_j \qquad (3.1)$$

or conversely as a function of member stiffnesses

$$n_j = [\, E_j a_j / l_j \,]\, u_j = k_j\, u_j \qquad (3.2)$$

Assembling for the whole structure

$$\underset{\sim}{u} = \underset{\sim}{\mathbb{F}} \, \underset{\sim}{n} \quad \text{and} \quad \underset{\sim}{n} = \underset{\sim}{\mathbb{K}} \, \underset{\sim}{u} \tag{3.3}$$

are obtained. The stress in each member must not exceed its permissible limit. The stress constraints take the form

$$s_j = n_j \,/\, a_j = E_j \, u_j \,/\, l_j \leq s_{jw} \tag{3.4}$$

Assembling for the whole structure

$$\underset{\sim}{s} = \underset{\sim}{S} \, \underset{\sim}{u} \leq \underset{\sim}{s}_W \tag{3.5}$$

The member distortions can be represented in any basis of β joint displacements (β is the degree of kinematic freedom) by a displacement transformation matrix for the structure

$$\underset{\sim}{u} = \underset{\sim}{A}^T \, \underset{\sim}{d} \qquad \text{(kinematics)} \tag{3.6}$$

where $\underset{\sim}{d}$ is the vector corresponding to the deflections under the external loads. The stiffness constraints relate external loads with member areas. equating the work done by the former to the work absorbed by the latter

$$\underset{\sim}{\lambda}^T \, \underset{\sim}{d} = \underset{\sim}{n}^T \, \underset{\sim}{u} \tag{3.7}$$

where $\underset{\sim}{\lambda}$ is the vector of external loads.

By expressing distortions as linear combinations of joint displacements

$$\underset{\sim}{u} = \underset{\sim}{A}^T \, \underset{\sim}{d} \tag{3.8}$$

$$\underset{\sim}{\lambda}^T \, \underset{\sim}{d} = \underset{\sim}{n}^T \, \underset{\sim}{A}^T \, \underset{\sim}{d} \quad \Rightarrow \quad \underset{\sim}{\lambda} = \underset{\sim}{A} \, \underset{\sim}{n}$$

(statics) (3.9)

It is therefore possible to express the joint deflections in terms of the external loads in the so called Nodal-Stiffness format

$$\underset{\sim}{\lambda} = \underset{\sim}{A}\,\underset{\sim}{K}\,\underset{\sim}{A}^{T}\,\underset{\sim}{d} = \underset{\sim}{K}\,\underset{\sim}{d} \qquad (3.1\emptyset)$$

3.3. Mesh (Matrix Force) Method

The equation of static equilibrium are not by themselves sufficient for the evaluation of the member forces in a redundant structure. The equilibrium equations can be derived by expressing the member forces separately in terms of the external loads and the unknown hyperstatic forces (p).

$$\underset{\sim}{n} = [\,\underset{\sim}{B}_{O} \,\vdots\, \underset{\sim}{B}\,] \begin{bmatrix} \underset{\sim}{\lambda} \\ \cdots \\ \underset{\sim}{p} \end{bmatrix} \qquad \text{(statics)} \qquad (3.11)$$

A general displacement matrix that is used is the direction cosine matrix $\underset{\sim}{A}$. Its product by $\underset{\sim}{B}$ will be singular ie :

$$\underset{\sim}{A}\,\underset{\sim}{B} = \emptyset \qquad (3.12)$$

The rows of $\underset{\sim}{A}$ (or the columns of $\underset{\sim}{B}_0$) span a subspace of dimension β whereas the columns of $\underset{\sim}{B}$ span a subspace of dimension α (α is the degree of static indeterminacy). Mello (198∅) has shown that $\underset{\sim}{A}$ is related to $\underset{\sim}{B}$ and to $\underset{\sim}{B}_0$ by

$$\underset{\sim}{A} = \underset{\sim}{B}_{O}^{T}\,(\underset{\sim}{B}_{O}\,\underset{\sim}{B}_{O}^{T} + \underset{\sim}{B}\,\underset{\sim}{B}^{T})^{-1} \qquad (3.13)$$

The stress in a member is obtained by dividing the force in it by its cross sectional area

$$s_j = n_j \,/\, a_j \leq s_{jw} \qquad (3.14)$$

The flexibility equations are needed in order to evaluate deflections

$$\begin{bmatrix} \underset{\sim}{d} \\ \cdots \\ \underset{\sim}{v} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{B}_o^T \\ \cdots \\ \underset{\sim}{B}^T \end{bmatrix} \underset{\sim}{F}\, \underset{\sim}{n} \qquad \text{(kinematics)} \qquad (3.15)$$

where $\underset{\sim}{v}$ is a null vector in the linear elastic phase. The equation

$$\underset{\sim}{u} = \underset{\sim}{F}\, \underset{\sim}{n} \qquad (3.16)$$

lists the extension and contraction of each member (distortion) . If the design requirements limit some displacements the constraints will be nonlinear involving the reciprocal of the member areas

$$\underset{\sim}{B}_o^T \underset{\sim}{F}\, \underset{\sim}{B}_o \underset{\sim}{\lambda} + \underset{\sim}{B}_o^T \underset{\sim}{F}\, \underset{\sim}{B}\, \underset{\sim}{p} \leq \underset{\sim}{\Delta} \qquad (3.17)$$

The α compatibility constraints that correspond to annulled discontinuities can be integrated with the equilibrium equations yielding the Mesh-Flexibility format

$$\underset{\sim}{B}^T \underset{\sim}{F}\, \underset{\sim}{B}_o \underset{\sim}{\lambda} + \underset{\sim}{B}^T \underset{\sim}{F}\, \underset{\sim}{B}\, \underset{\sim}{p} = \underset{\sim}{\emptyset} \qquad (3.18)$$

3.4. Miscellaneous Methods

In elastoplastic and large displacement analysis it is computationally more efficient to consider the Mesh description combined with Stiffness relations requiring a smaller up-dating effort (Mesh-Stiffness format)

$$\begin{bmatrix} \underset{\sim}{K} & -\underset{\sim}{B} \\ -\underset{\sim}{B}^T & \cdot \end{bmatrix} \begin{bmatrix} \underset{\sim}{u} \\ \underset{\sim}{p} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{B}_0 \underset{\sim}{\lambda} - \underset{\sim}{n} \\ \cdot \end{bmatrix} \tag{3.19}$$

We remark that $\underset{\sim}{v}$ is no longer $\underset{\sim}{\emptyset}$ when shrinkage or elastoplastic hysteretic behaviour occur.

3.5. Structural Synthesis of Linear Elastic Trusses

Under the assumption that the cost of each section is proportional to its area the problem to be solved consists of the minimum volume design of an elastic truss subject to bounds on nodal displacements, allowable stresses, member areas and buckling constraints. The topology of the truss is assumed to be given. Therefore the nodes are fixed and the member areas are not allowed to vanish. It is also assumed that there will be no geometry changes to be taken into account. The truss is subjected to 1 loading cases.

Since we are minimizing the volume of the structure the objective function will be a linear function of the member areas

$$\text{Min } \underset{\sim}{1}^T \underset{\sim}{a} \tag{3.20}$$

The β.1 equilibrium equations can be represented in a bilinear format involving design variables (cross sections) and state variables (stress/displacements)

$$\underset{\sim}{a}^T \underset{\sim}{H}_i \underset{\sim}{s}^k = \lambda_i^k \quad i=1,\ldots\beta \; ; \quad k=1,\ldots 1 \tag{3.21}$$

where $\underset{\sim}{H}_i$ is a diagonal matrix whose elements are the constants of the ith row in the direction cosine matrix.

The compatibility constraints are α linear relations in the state variables

$$\underset{\sim}{B}^T \underset{\sim}{L} \underset{\sim}{s}^k = \emptyset \tag{3.22}$$

where $\underset{\sim}{L}$ is a diagonal matrix whose elements are the quotient of the length of the members by the Young modulus. The bounds on member areas and member stresses can be written respectively

$$\underset{\sim}{a}_l \leq \underset{\sim}{a} \leq \underset{\sim}{a}_u \qquad \underset{\sim}{s}_l^k \leq \underset{\sim}{s}^k \leq \underset{\sim}{s}_u^k \tag{3.23}$$

A linear relationship exists linking stresses and displacements

$$\underset{\sim}{L}^{-1} \underset{\sim}{A}^T \underset{\sim}{d}^k = \underset{\sim}{s}^k \tag{3.24}$$

From this last overdeterminate system it is possible to express displacements in terms of the stresses by

$$\underset{\sim}{d}^k = \underset{\sim}{D} \underset{\sim}{s}^k \tag{3.25}$$

where $\underset{\sim}{D}$ is the inverse of a square nonsingular submatrix of $\underset{\sim}{L}^{-1} \underset{\sim}{A}^T$ so that we can write the bounds on nodal displacements by

$$\underset{\sim}{d}_l^k \leq \underset{\sim}{D} \underset{\sim}{s}^k \leq \underset{\sim}{d}_u^k \tag{3.26}$$

Buckling constraints may be defined by using Euler-Johnson (Shechler,E.E. and Dunn,L.G. (1963)) stability analysis : For long columns compressive failure stress is given by

$$\bar{s}_{jw} = s_{cr} = (c_j \pi^2 E_j) / (l_j/r_j)^2 \tag{3.27}$$

where c_j is the end fixity coefficient (c_j=1 for pin joined frameworks) and r_j is the radius of giration of the cross section.

Buckling loads are therefore

$$P_{cr} = (\pi^2 E_j I_j) / l_j^2 \tag{3.28}$$

since $r_j = I_j/a_j$.

Assuming all cross sections changing by the same amount

$$P_{cr} = (\gamma \pi^2 E_j a_j^2) / l_j^2 \tag{3.29}$$

In the compressed members the stress constraints can be substituted by

$$-a_j s_j^k - (\gamma \pi^2 E_j a_j^2) / l_j^2 \leq 0$$

or

$$-s_j^k - (\gamma \pi^2 E_j a_j) / l_j^2 \leq 0$$

or

$$-\underset{\sim}{s}^k - \underset{\sim}{E}\, \underset{\sim}{a} \leq 0 \tag{3.30}$$

We obtain a linear relation in both state and design variables. The failure stress for short columns can also be checked. The Johnson parabolic approximation gives

$$s_{\bar{w}} = s_I^* - s_{II}^*(s_I^* - s_{II}^*) / s_{cr} \tag{3.31}$$

where s_I^* and s_{II}^* are specified stresses

Graphically it can be represented by a parabola that intersects the Euler curve at the point $(\sqrt{c\pi^2 E\,(s_{II}^*)^{-1}},\ s_{II}^*)$ and has its vertex at $(0, s_I^*)$. s_{II}^* represents the allowable compressive stresses determined either by yielding or local instability. s_{II}^* is taken 1/2 s_I^*. Failure by skin wrinkling is avoided following our assumption of proportional change of all member sections and choosing stable proportions at the outset. If necessary stresses of this type could be included without changing the nature of the problem but we are not going to do it here.

The truss optimization problem has a linear objective function in terms of the member areas. The equilibrium equations are bilinear in member areas and stresses the compatibility equations are linear in the member stresses. The bounds on areas stresses and displacements are linear range constraints. The surface which spans in the simplest way a twisted rectangle has the equation

$$z = x.y \,/\, c \qquad \text{(hyperbolic paraboloid)} \qquad (3.32)$$

It reflects the behaviour of each term of the equilibrium equations. In order to split this factorable term into a separable form in the sense that each term would become a function of one variable a rotation of $\pi/4$ is induced to the coordinate axes. The equation of each term in the equilibrium equations assumes the form

$$z = u^2/(\ c\ ) - v^2/(\ c\ ) \qquad (3.33)$$

where $u = (x+y)/2$ and $v = (x-y)/2$.

It is clear from this that the equilibrium equations are a sum of strictly convex and strictly concave terms and therefore nonconvex.

In statically determinate trusses it can be shown that after a transformation of variables the problem can be reformulated as a convex programming problem and hence there will be a unique solution to it. When the truss has a determinate layout the equilibrium equations directly yield the force vector and this is an important piece of information because it means that the basic bilinear variables of the problem are known . Consider the reciprocal of the member areas. These are nonnegative variables and it is possible to express both stresses and displacements as a linear function of them. We have therefore a linearly constrained problem where the objective function is a sum of convex functions and is therefore convex. By contrast in a indeterminate structure it is not possible to know the member forces before the structure is designed. Taking 1/a type of variables together with member stresses does not fully succeed in making the problem linear although this is often done to improve convergence. Alternatively to this variable combination member force/area variable choice seems also to work well [Johson (1982)] if a convex problem is assumed "a priori".

A particularly vexing problem in the solution of nonconvex optimization is that algorithms for solving the problem will converge to local as opposed to global solutions. In the following section examples of multiple optima in minimum volume truss design will be discussed. Forthcoming chapters are dedicated to examine two alternative proposals that overcome this difficulty.

3.6. Multiple Optimal Solutions in Trusses

3.6.1 Variable topology solutions

A simple structural optimization problem that usually serves as testbed for synthesis algorithms is the three bar truss represented in Fig 3.1.

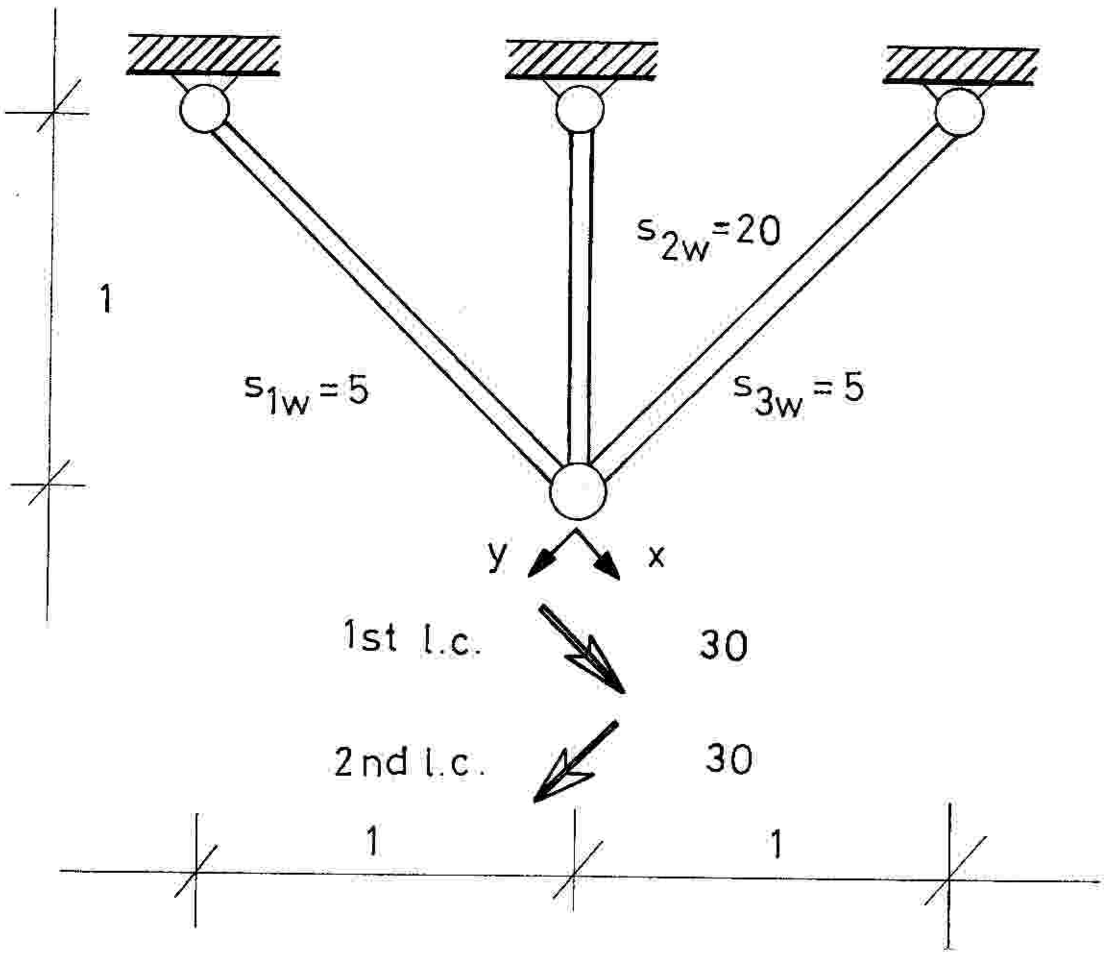


Fig 3.1 Three Bar Truss

The design objective is to choose a set of cross sectional areas so that the truss has a minimal volume while satisfying the constraints on stress. The cost function is therefore

$$\sqrt{2}\, a_1 + a_2 + \sqrt{2}\, a_3 \qquad (3.34)$$

By liberating the structure the matrix B is given by (Fig 3.2)

$$\underset{\sim}{B}^T = [\,-\sqrt{2}/2 \quad 1 \quad -\sqrt{2}/2\,] \tag{3.35}$$

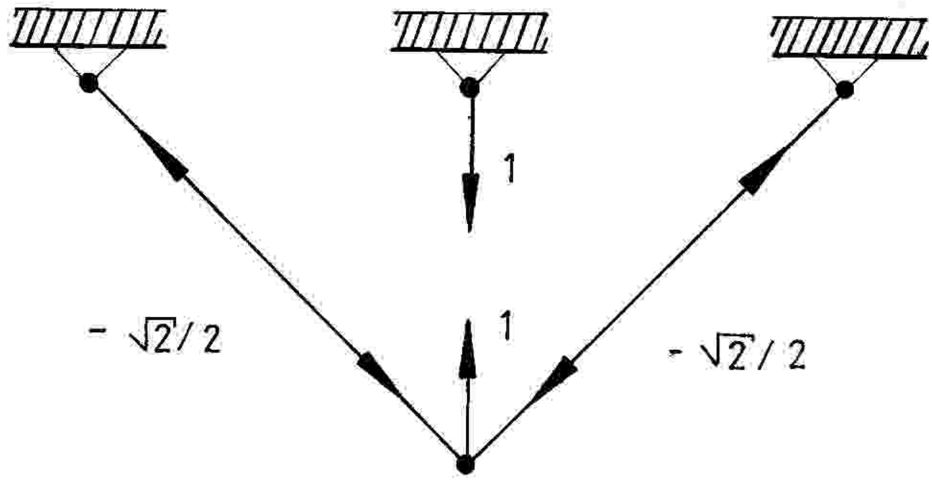


Fig 3.2 Unit Load Applied

to the Liberated Structure

The direction cosine matrix can be obtained after specifying a basis for displacements. Supposing they coincide with the two remaining bars

$$\underset{\sim}{B}_O = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.36}$$

$\underset{\sim}{A}$ can be obtained by simple matrix inversion.

$$\underset{\sim}{A} = \underset{\sim}{B}_O^T (\underset{\sim}{B}_O \underset{\sim}{B}_O^T + \underset{\sim}{B}\,\underset{\sim}{B}^T)^{-1}$$

$$A = \begin{bmatrix} 1 & \sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 & 1 \end{bmatrix} \tag{3.37}$$

Assuming the Young modulus is unity ie E = 1 the linear elastic displacements can be written in terms of the member stresses

$$\underset{\sim}{d}^{k} = \underset{\sim}{B}_{o}^{T} \underset{\sim}{L} \underset{\sim}{s}^{k} \tag{3.38}$$

$$\begin{bmatrix} d_x^k \\ d_y^k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & & \\ & 1 & \\ & & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_1^k \\ s_2^k \\ s_3^k \end{bmatrix} = \begin{bmatrix} \sqrt{2}\, s_1^k \\ \sqrt{2}\, s_3^k \end{bmatrix}$$

Alternatively the member stresses can be written in terms of the displacements

$$\underset{\sim}{s}^{k} = \underset{\sim}{L}^{-1} \underset{\sim}{A}^{T} \underset{\sim}{d}^{k} \tag{3.39}$$

$$\begin{bmatrix} s_1^k \\ s_2^k \\ s_3^k \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 & & \\ & 1 & \\ & & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_x^k \\ d_y^k \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2\ d_x^k \\ \sqrt{2}/2\,(d_x^k + d_y^k) \\ \sqrt{2}/2\ d_y^k \end{bmatrix}$$

The compatibility constraints have the form

$$\underset{\sim}{B}^{T} \underset{\sim}{L} \underset{\sim}{s}^{k} = \underset{\sim}{0} \tag{3.40}$$

$$[-\sqrt{2}/2 \quad 1 \quad -\sqrt{2}/2] \begin{bmatrix} \sqrt{2} & & \\ & 1 & \\ & & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_1^k \\ s_2^k \\ s_3^k \end{bmatrix} = 0$$

$$-s_1^k + s_2^k - s_3^k = 0$$

The β.1 equilibrium equations are bilinear expressions in terms of stresses and areas

$$\underset{\sim}{A}\ (a.\underset{\sim}{s}^k) = \underset{\sim}{\lambda}^k \qquad (3.41)$$

$$a_1\ s_1^k + \sqrt{2}/2\ a_2\ s_2^k = \lambda_x^k$$

$$\sqrt{2}/2\ a_2\ s_2^k + a_3\ s_3^k = \lambda_y^k$$

where $\beta = 2$ and λ_x^k, λ_y^k are the load components in the directions x and y respectively can be expressed identically in terms of displacements and areas

$$\underset{\sim}{K}\ \underset{\sim}{d}^k = \underset{\sim}{\lambda}^k \qquad (3.42)$$

$$(\sqrt{2}/2\ a_1 + 1/2\ a_2)\ d_x^k + 1/2\ a_2\ d_y^k = \lambda_x^k$$

$$1/2\ a_2\ d_x^k + (\sqrt{2}/2\ a_3 + 1/2\ a_2)\ d_y^k = \lambda_y^k$$

where $\underset{\sim}{K}$ is the assembled stiffness matrix. The synthesis problem subject to stress constraints only can be stated as

$$\text{Min}\ \sqrt{2}\ a_1 + a_2 + \sqrt{2}\ a_3 \qquad (3.43)$$

$$\underline{\text{st}} \quad (\sqrt{2}/2\ a_1 + 1/2\ a_2)\ d_x^k + 1/2\ a_2\ d_y^k = \lambda_x^k \qquad (3.44)$$

$$1/2\ a_2\ d_x^k + (\sqrt{2}/2\ a_3 + 1/2\ a_2)\ d_y^k = \lambda_y^k \qquad (3.45)$$

$$s_{1l}^k \leq \sqrt{2}/2\ d_x^k \leq s_{1u}^k \qquad (3.46)$$

$$s_{2l}^k \leq \sqrt{2}/2\ (d_x^k + d_y^k) \leq s_{2u}^k \qquad (3.47)$$

$$s_{3l}^k \leq \sqrt{2}/2\ d_y^k \leq s_{3u}^k \qquad (3.48)$$

$$a_1, a_2, a_3 \geq 0 \qquad k=1,\dots l$$

l represents the number of loading conditions.

This small scale example can be solved analytically. It is based on Sved and Ginos (1968) and a fully nonconvex behaviour is manifested. Haug,E. and Arora,E.J. (1979) also solved a similar truss analyticaly but for different loading conditions. In their case the problem turns out to be convex after all.

The displacements are uniquely given by the inverse of the assembled stiffness matrix times the loading vector. They can be viewed as being determined by the structural equations once the design variables are specified.

$$\begin{bmatrix} d_x^k \\ d_y^k \end{bmatrix} = (\det)^{-1} \begin{bmatrix} (\sqrt{2}/2\ a_3 + 1/2\ a_2) & -1/2\ a_2 \\ -1/2\ a_2 & (\sqrt{2}/2\ a_1 + 1/2\ a_2) \end{bmatrix} \qquad \begin{matrix} (3.49) \\ (3.50) \end{matrix}$$

We remark that the determinant of K

$$\det = 1/2\ a_1 a_3 + \sqrt{2}/4\ a_1 a_2 + \sqrt{2}/4\ a_2 a_3 > 0 \qquad (3.51)$$

for any two out of a_1, a_2, a_3 greater than 0.

The constraints on the stresses can be written explicitly in terms of the design variables

$$s_{1l}^k \leq [(\sqrt{2}/2\ a_3 + 1/2\ a_2)\ \lambda_x^k - 1/2\ a_2 \lambda_y^k]$$

$$[\sqrt{2}/(2.\det)] \leq s_{1u}^k \qquad (3.52)$$

$$s_{2l}^k \leq [\sqrt{2}/2\ a_3 \lambda_x^k + \sqrt{2}/2\ a_1 \lambda_y^k]$$

$$[\sqrt{2}/(2.\det)] \leq s_{2u}^k \qquad (3.53)$$

$$s_{3l}^k \leq [-1/2\ a_2 \lambda_x^k + (\sqrt{2}/2\ a_1 + 1/2\ a_2)\ \lambda_y^k]$$

$$[\sqrt{2}/(2.\det)] \leq s_{3u}^k \qquad (3.54)$$

However even in this simple problem such manipulation is impractical. In larger scale it is impossible.

A simplification may occur if we subject the truss to a symmetric loading corresponding to two alternative loading conditions.

The constraints on the stresses corresponding to loading condition 1

$$s_{11}^{1} \leq [\sqrt{2}/2\ a_3 + 1/2\ a_2]\ [15.\sqrt{2}/det] \leq s_{1u}^{1} \qquad (3.55)$$

$$s_{21}^{1} \leq [\sqrt{2}/2\ a_3]\ [15.\sqrt{2}/det] \leq s_{2u}^{1} \qquad (3.56)$$

$$s_{31}^{1} \leq [-1/2\ a_2]\ [15.\sqrt{2}/det] \leq s_{3u}^{1} \qquad (3.57)$$

$$s_{1}^{1} = [0\ 0\ -5] \qquad s_{u}^{1} = [5\ 20\ 0]$$

$$s_{1}^{2} = [-5\ 0\ 0] \qquad s_{u}^{2} = [0\ 20\ 5]$$

Due to symmetry of layout loading and maximum allowable stresses

$$a_1 = a_3 \qquad (3.58)$$

We note that for any a_1, a_2 greater than 0 the lower bounds on the first two stresses and the upper bound on the third stress are irrelevant. We can rewrite the equivalent MP

$$\text{Min} \quad 2\sqrt{2}\ a_1 + a_2 \qquad (3.59)$$

$$\underline{st} \quad s_1 = (30\ a_1 + 15\sqrt{2}\ a_2)/(a_1^2 + \sqrt{2}\ a_1\ a_2) \leq 5 \qquad (3.60)$$

$$s_2 = (30\ a_1)/(a_1^2 + \sqrt{2}\ a_1\ a_2) \leq 20 \qquad (3.61)$$

$$s_3 = (15\sqrt{2}\ a_2)/(a_1^2 + \sqrt{2}\ a_1\ a_2) \leq 5 \qquad (3.62)$$

$$a_1, a_2 \geq 0$$

The OF and the nonnegativity conditions on the design variables are represented by linear functions and are therefore convex. The second derivative matrices for the constraint functions s_2 and s_3 are positive definite and positive semidefinite respectively. Thus s_2 and s_3 are both convex. s_1 is the sum of s_2 and s_3 that are convex functions and it is therefore convex. The mathematical minimization involves convex functions and any local minimizer on the domain defined is also the global minimizer on that domain. Local minimizers are obtained by solving the system obtained by considering the K-T conditions. The minimum volume is therefore 15.824 corresponding to the following design variables

$$a_1 = 4.725 \qquad a_2 = 2.460$$

state variables

$$s_1 = 5. \qquad s_2 = 3.66 \qquad s_3 = -1.34$$

Placing these results into the original variables

member areas

$$a_1 = 4.725 \qquad a_2 = 2.46 \qquad a_3 = 4.725$$

stress/displacement resultant

$$d_x^1 = 7.071 \quad d_y^1 = -1.845 \quad s_1^1 = 5. \quad s_2^1 = 3.66 \quad s_3^1 = -1.34$$

$$d_x^2 = -1.845 \quad d_y^2 = 7.071 \quad s_1^2 = -1.34 \quad s_2^2 = 3.66 \quad s_3^2 = 5.$$

We remark that although the expressions that give the member stresses as a function of the design variables are in general nonconvex the convex problem resulting from the assumption $a_1 = a_3$ posseses a unique solution.

The three bar truss would seem to be completely solved in this particular case. But if we set either a_1 or a_3 equal to Ø the resulting statically determinate two bar truss (with a_3 = Ø) is considerably smaller than the three bar truss (OF = 1Ø.6Ø7) being able to withstand the same loading.

$$a_1 = 6. \quad a_2 = 2.12$$

$$s_1^1 = 5. \quad s_2^1 = 0. \quad s_1^2 = -5. \quad s_2^2 = 20.$$

The multiplicity of solutions in this case is not due to the nonconvexity inherent to the constraint functions. The hypothetical stresses on bar 3 corresponding to the latter design (and an infinitelly small bar 3) are $s_3^1 = -5.$ $s_3^2 = 25.$ This stress would violate the constraint limiting the maximum allowable stress in member 3. For the two bar truss this is absurd since the member does not exist. If one changes the upper limit on the stress in member 2 to into 5 the statically indeterminate design would give the global solution.

This situation is equivalent to the minimization of a linear OF subject to a constraint discontinuity. This is to say that the domain is composed of disjoint admissible design sets. From this example it is possible to conclude that all optimal designs will have this type of disjoint solution.

One may try to find if there exists a path joining the indeterminate solution and a statically determinate design.

Consider the three bar truss represented in Fig 3.1 subjected to two symmetric alternative loading conditions where the load angle has been modified

$$\underset{\sim}{\lambda}^{1} = [\ 40.\ \ 10.]^{T} \qquad \underset{\sim}{\lambda}^{2} = [\ 10.\ \ 40.]^{T}$$

Assuming $a_1 = a_3$ and if the stress/displacement in member 1 under loading condition 1 is at its upper value it is possible to write the member areas in terms of the stress in member 3 by using the nodal stiffness equations. For any $s_1 = x$ and $s_3 = y$, s_2 is determined by the compatibility relations

$$s_2 = s_1 + s_3$$

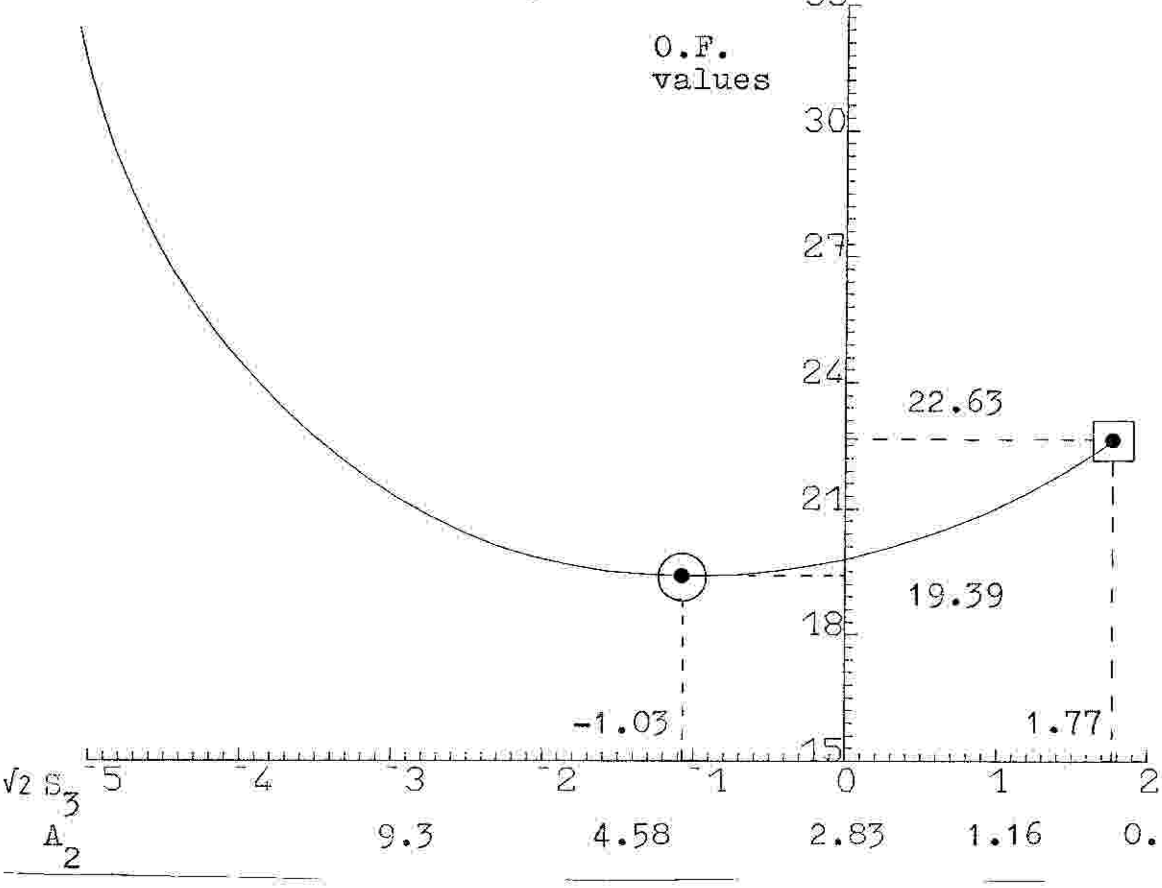
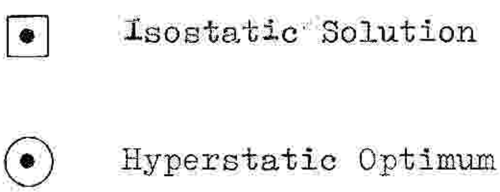


Fig 3.3 Variation of the OF with the stress in member 3



The optimal indeterminate truss is

$$a_1 = 5.237 \quad a_2 = 4.577 \qquad OF = 19.389$$

$$s_1 = 5. \qquad s_2 = 4.270 \qquad s_3 = -\ .73$$

If the stress in member 3 is increased keeping the same active set for the stress limits we obtain the results represented in Fig 3.3 .

In this case the statically determinate design corresponding to

$$a_1 = 8. \qquad a_2 = 0. \qquad\qquad OF = 22.627$$

$$s_1 = 5. \qquad s_2 = 6.25 < 20. \qquad s_3 = 1.25$$

is a feasible point of the domain.

We now change the upper limit on the stress in member 2 in order to show that this path may not be feasible. Whenever this happens, algorithms that find the global solution of the structural design problem have to consider several separate domains each of which possesssing its own solution.

Assume $s_{2u} = 5.$ The same plotting would be obtained until

$$a_1 = 6. \qquad a_2 = 2.828 \qquad OF = 19.799$$

$$s_1 = 5. \qquad s_2 = 5. \qquad\qquad s_3 = 0.$$

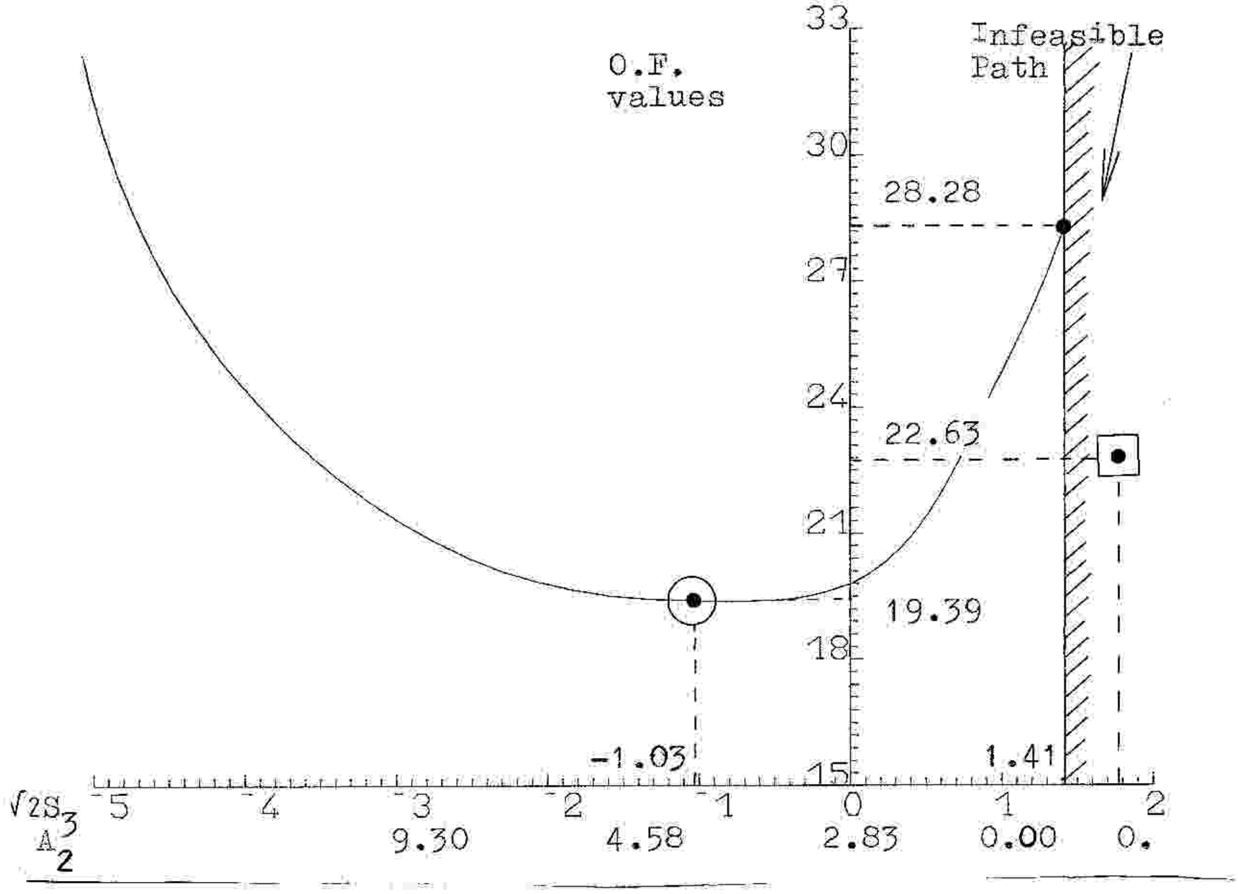


Fig 3.4 Variation of the OF with the stress in member 3

After this point the stress in member 3 under loading condition 1 can only be increased if s_1 is also reduced (with s_2 at its upper limit defining the active set) [Fig 3.4] . It is not possible to increase s_3 after the point

$$a_1 = 10. \quad a_2 = 0. \qquad OF = 28.284$$

$$s_1 = 4. \quad s_2 = 5. \qquad s_3 = 1.$$

For bigger s_3 the system of equations would give a negative value for a_2 that has no physical meaning.

This situation can be viewed graphically in Fig 3.5 where there is a linear path in the stress domain to which corresponds a nonlinear feasible path in the area hypercube. In the latter instance [Fig 3.6] the upper stress limit at bar 2 becomes active and then the proposed nonlinear path in the area domain is no longer feasible. Therefore the statically determinate solution corresponding to the elimination of a_2 cannot be reached by continuous variation of the state variables throughout the feasible set.

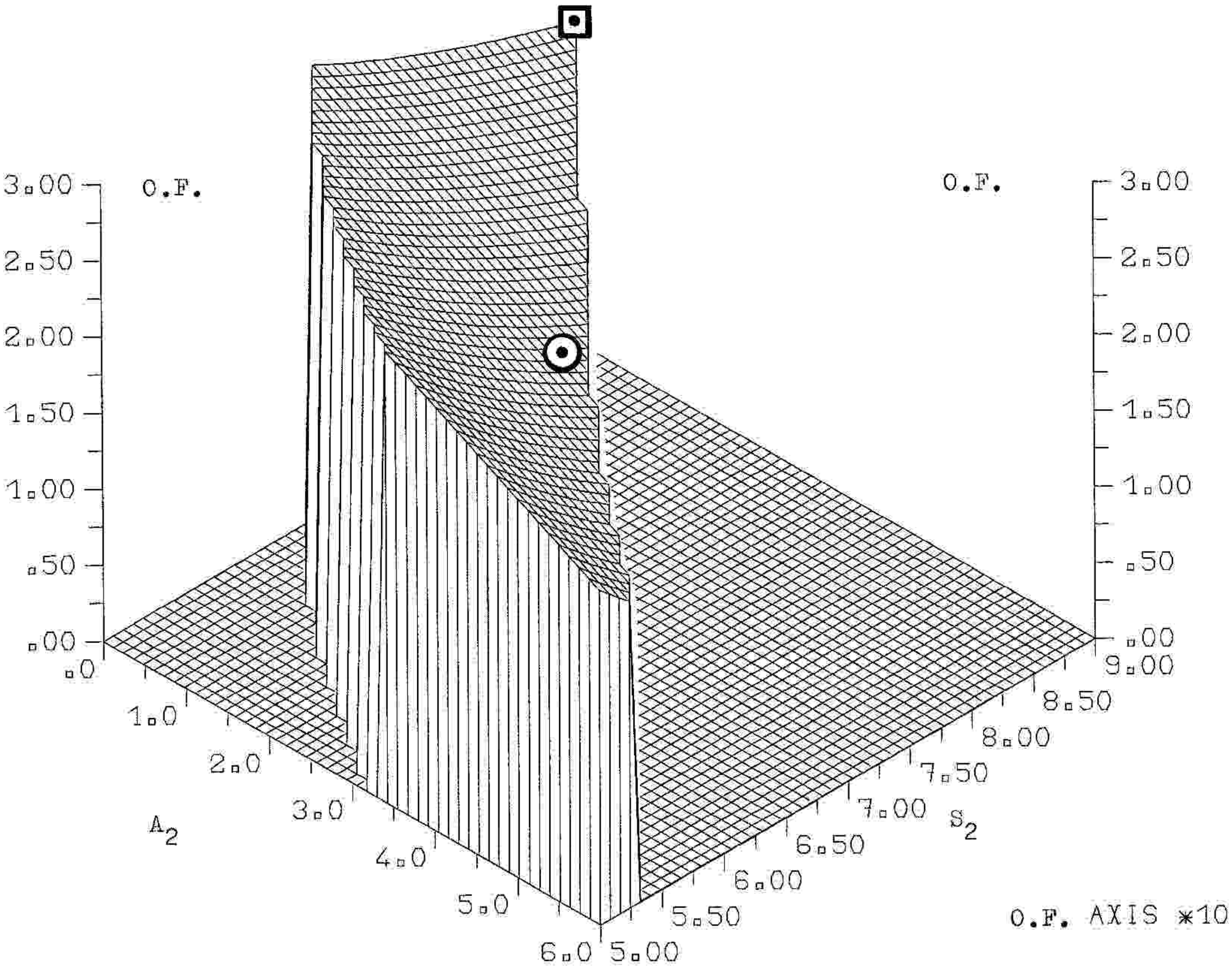


Fig 3.5 Perspective view corresponding to the feasible volumes obtained by varying S_2 and A_2 within their rectangle of bounds.
It is possible to yield $A_2 = 0$ by continuous variation of S_2

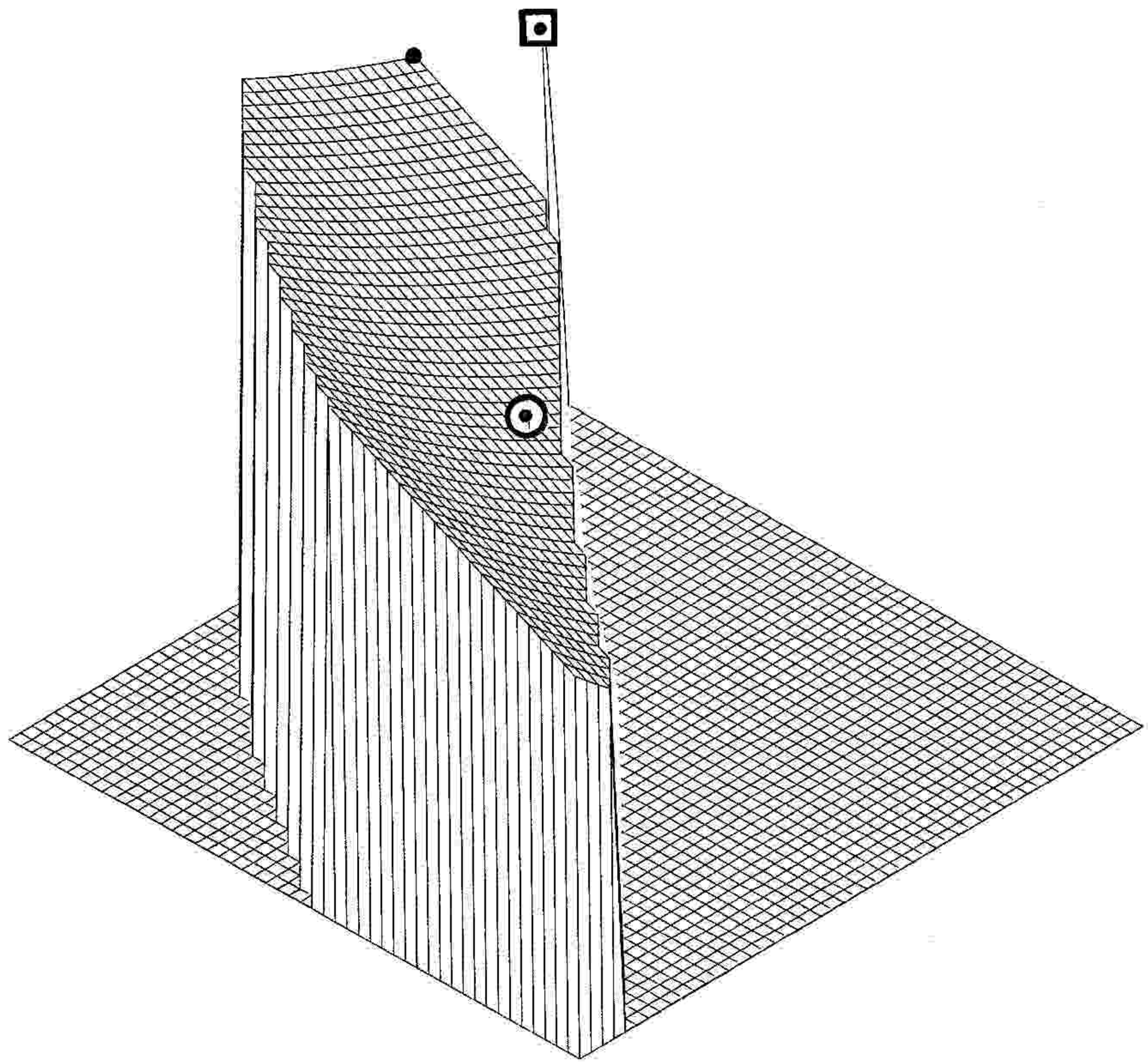


Fig. 3.6 By changing the upper limit in S_2 the path that would yield the isostatic solution is no longer feasible.

The multiplicity of solutions described in this paragraph is due to the change in topology caused by the addition of a member. One sure way of avoiding this behaviour is to suboptimize within each configuration that is admissible. However such total enumeration is prohibitive in terms of computational time and the further requires more economical methods of solution . One such approach is introducing Ø-1 variables and using a branch and bound strategy in the way described in Chapter 4 . Alternatively these variables can be integrated in a master problem that will be defined in Chapter 5.

The n Ø-1 variables δ_j will denote wether a member is present in wich case $\delta_j = 1$ or absent $\delta_j = 0$. The constraint on the upper bounds of the member areas become

$$\underset{\sim}{a} \leq \underset{\sim}{\delta}\, \underset{\sim}{a}_u \tag{3.63}$$

The upper and lower bounds on the stresses are respectively

$$\underset{\sim}{s}^k \leq \underset{\sim}{\delta}\, \underset{\sim}{s}_u^k + (1-\underset{\sim}{\delta})\, \underset{\sim}{s}_{BIG} \tag{3.64}$$

$$\underset{\sim}{\delta}\, \underset{\sim}{s}_1^k - (1-\underset{\sim}{\delta})\, \underset{\sim}{s}_{BIG} \leq \underset{\sim}{s}^k \tag{3.65}$$

where $\underset{\sim}{s}_{BIG}$ represents a large enough value.

In order to make the design structurally admissible the number of members present in the structure must be not less than number of degrees of kinematic freedom

$$\sum_{j=1}^{n} \delta_j \geq \beta \tag{3.66}$$

This relation makes sure that there must be at least β members in a truss and this means that any combination which has less than ß members will not be considered. When there are ß members or more in a truss however this does not mean that the structure will always carry the loads because it may still contain local mechanisms. Although these combinations may be considered they are not valid answers to the problem.

Changing topologies would lead to the the existence of disjoint domains defined by the presence and absence of each member together constituting a nonconvex set. Such problems are better treated combinatorially by introducing 0-1 variables and transforming the general problem into a 0-1 mixed bilinear constrained program.

3.6.2 Curvature of the bilinear constraints

By introducing lower bounds on the design variables the topology of the structure will remain fixed throughout the optimization process. A question arises related with the occurence of multiple optimal solutions due to the curvature of the bilinear constraints. Consider the following ten bar truss drawn in Fig 3.3 subject to a single loading condition (the author is grateful to Dr. Bartholomew for passing details of another variation of this case by private communication). In fact multiple loading cases do not alter the nature of the problem but complicate notation.

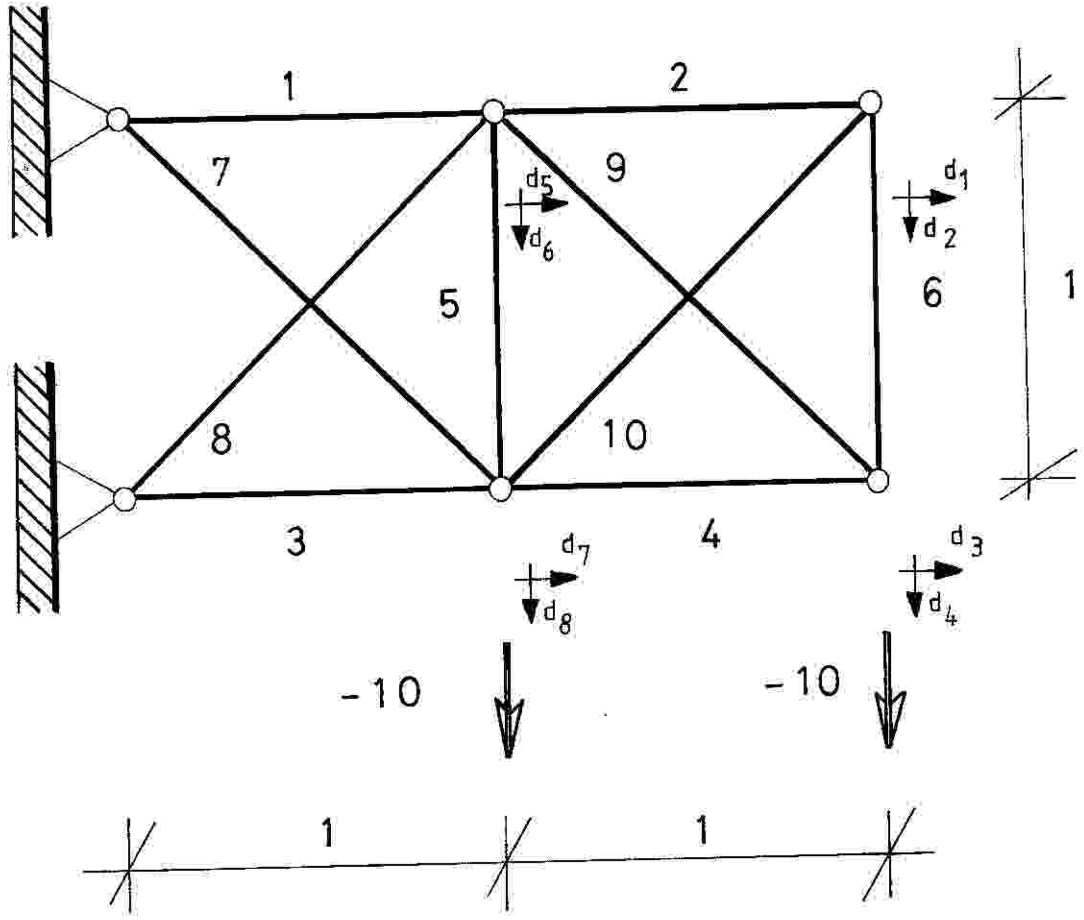


Fig 3.7 Ten Bar Truss

The degree of kinematic freedom in this system is $\beta = 8$

Loading vector

$$\underset{\sim}{\lambda}^T = [0 \quad 0 \quad 0 \quad -10 \quad 0 \quad 0 \quad 0 \quad -10] \tag{3.68}$$

Lower bounds on the areas

$$\underset{\sim}{a}_1^T = [.1 \quad .1 \quad .1 \quad .1 \quad .1 \quad .1 \quad .1 \quad .1 \quad .1 \quad .1] \tag{3.69}$$

The stresses can be written in terms of nodal displacements

$$\underset{\sim}{s} = \underset{\sim}{L}^{-1} \underset{\sim}{A}^{T} \underset{\sim}{d} \tag{3.70}$$

$$s = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \end{bmatrix} = \begin{bmatrix} d_5 \\ d_1 - d_5 \\ d_7 \\ d_3 - d_7 \\ d_6 - d_8 \\ d_2 - d_4 \\ (d_4 - d_8)/2 \\ (d_5 + d_6)/2 \\ (d_3 - d_4 - d_5 + d_6)/2 \\ (d_1 + d_2 - d_7 - d_8)/2 \end{bmatrix}$$

The direction cosine β.n matrix $\underset{\sim}{A}$ is given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2}/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \sqrt{2}/2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \sqrt{2}/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -\sqrt{2}/2 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \sqrt{2}/2 & \sqrt{2}\;2 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & \sqrt{2}/2 & 0 & 0 & -\sqrt{2}/2 \\ 0 & 0 & 0 & 0 & -1 & 0 & -\sqrt{2}/2 & 0 & 0 & -\sqrt{2}/2 \end{bmatrix} \quad (3.67)$$

The equilibrium equations can be presented in the more condensed Assembled Stiffness format

$$\underset{\sim}{K}\ \underset{\sim}{d} = \underset{\sim}{\lambda} \tag{3.71}$$

The problem with no displacement constraints considered and subjected to the maximum allowable stresses of ±2.5 gives a single optimal solution of 44.25 corresponding to the following design variables

$a_1 = 7.94$ $a_2 = .1$ $a_3 = 8.06$ $a_4 = 3.94$ $a_5 = .1$

$a_6 = .1$ $a_7 = 5.74$ $a_8 = 5.57$ $a_9 = 5.57$ $a_{10} = .1$

The optimal stresses and displacements are respectively

$s_1 = 2.5$ $s_2 = 1.6$ $s_3 = -2.5$ $s_4 = -2.5$ $s_5 = 0.$

$s_6 = 1.55$ $s_7 = 2.5$ $s_8 = -2.5$ $s_9 = 2.5$ $s_{10} = -2.2$

$d_1 = 4.05$ $d_2 = -18.45$ $d_3 = -5.$ $d_4 = -20.$

$d_5 = 2.5$ $d_6 = -7.5$ $d_7 = -2.5$ $d_8 = -7.5$

If the absolute value of the displacement is limited to 3.5 a fully nonconvex behaviour will occur. Indeed two optimal solutions were obtained

Global Solution 219.93

$a_1 = 48.68$ $a_2 = .1$ $a_3 = 35.55$ $a_4 = 24.1$ $a_5 = .1$

$a_6 = 1.21$ $a_7 = 9.37$ $a_8 = 34.35$ $a_9 = 34.09$ $a_{10} = .1$

$s_1 = .42$ $s_2 = -.25$ $s_3 = -.56$ $s_4 = -.42$ $s_5 = 2.24$

$s_6 = -.02$ $s_7 = 2.94$ $s_8 = -.84$ $s_9 = .83$ $s_{10} = .72$

$d_1 = .16$ $d_2 = -3.5$ $d_3 = -.97$ $d_4 = -3.48$

$d_5 = .42$ $d_6 = -1.26$ $d_7 = -.56$ $d_8 = -3.50$

Local Solution 223.34

$a_1=48.67$ $a_2=.1$ $a_3=38.09$ $a_4=23.33$ $a_5=.1$

$a_6=.1$ $a_7=13.66$ $a_8=33.11$ $a_9=32.99$ $a_{10}=.1$

$s_1=.41$ $s_2=.00$ $s_3=-.52$ $s_4=-.42$ $s_5=1.29$

$s_6=.00$ $s_7=2.04$ $s_8=-.87$ $s_9=.86$ $s_{10}=.00$

$d_1=.41$ $d_2=-3.5$ $d_3=-.95$ $d_4=-3.5$

$d_5=.41$ $d_6=-1.28$ $d_7=-.52$ $d_8=-2.56$

These points are solutions to the K-T equations corresponding to two different sets of Lagrange multipliers. Both locally minimize the volume of the structure . Any feasible direction linking both solutions would determine (at least) one point having a greater volume than the local optima. For a fixed set of design variables the stress/displacement resultant is uniquely determined. A infeasible displacement in d_2 arises when the member areas are linearly varied between the solution points.

Stresses may also be continuously varied between the global and local solution coordinates. It is known that for a fixed set of stresses the minimum volume can be obtained by solving an LP. The locus of minimal volumes will represent a convex function starting at the coordinates corresponding to the global solution until the local solution coordinates are attained. At this point the matrix that is a function of the stresses and expresses the equilibrium relations will become singular ie: the same set of displacements (and stresses) is obtained by multiplying the inverse of the assembled stiffness matrix by the load vector for several realizations of a_2 and keeping the remaining member areas at their local solution coordinates. a_2 is free to assume any value so that it will be considered at its lower bound in order to minimize the total volume of the structure. This type of nonconvexity is truly due to the nonconvexity of the domain induced by the bilinear relations. The manifestation of this behaviour justifies the use of methods more appropriate for nonconvex programming that will be discussed latter. But first conditions will be given that verify the existence of an isolated global minimizer of the structural synthesis problem with continuous design variables.

3.7. Characterization of Local Solutions

Considerable theory is available that characterizes local minima. Perhaps the most suitable theoretical development for our purposes is that of McCormick (1980). Unfortunately it happens to be rather abstruse. It has been quoted at length in the following pages in order to reflect its full scope,but sucessful applications of the theory depend on problem specific simplifications. The approach in local theory is to suppose that the problem

$$\text{Min} \quad f(\underset{\sim}{x}) \tag{3.72}$$

$$\underline{\text{st}} \quad g_j(\underset{\sim}{x}) = 0 \qquad j=1,\ldots q \tag{3.73}$$

$$g_j(\underset{\sim}{x}) \geq 0 \qquad j=q+1,\ldots Q \tag{3.74}$$

$$\underset{\sim}{x} \in X$$

has a local extremum at point x^* and then to find conditions among $f(x)$ and $g(x)$ that must hold at this point. In this way many points in the constraint region can be eliminated as candidates for a relative minimum. Such conditions are therefore called necessary. In some problems it will be possible to obtain a set of conditions that guarantee that a point yields a local minima. Conditions of this kind will be termed sufficient. Before meaningful results applicable to our problem are given some conditions and theorems will be stated next followed by their application to the least volume design.

A feasible point x_o is called a regular point of the domain if $f(x)$ is differentiable at x_o and if the gradient $\nabla_x g_j(x_o)$ for only those j>q with $g_j(x_o) = \emptyset$. (active inequality constraints) and all $\nabla_x g_j(x_o)$ j=1,...q (equality constraints) are linearly independent. This definition is equivalent to the requirement that the matrix of the active constraint derivatives $g_j'(x_o)$ has full rank where j=1,...q and will include those j>q with $g(x_o)=\emptyset$.

THEOREM (Kuhn-Tucker Necessary Conditions)

Let the functions $f(x)$ and $g_j(x)$ j=1,...q be differentiable and let $\bar{x}$ be a feasible regular point. For $\bar{x}$ to be a local minimum of (P) it is necessary that there exist a multiplier vector $\gamma \in R^Q$ such that

$$\gamma_j \geq 0 \qquad j=q+1,\ldots Q \qquad (3.75)$$

$$\gamma_j g_j(\underset{\sim}{x}) = 0 \qquad j=1,\ldots q \qquad (3.76)$$

$$\nabla_{x_i} L(\underset{\sim}{x},\underset{\sim}{\gamma}) = 0 \qquad i=1,\ldots n \qquad (3.77)$$

where $L(\underset{\sim}{x},\underset{\sim}{\gamma}) = f(\underset{\sim}{x}) - \underset{\sim}{\gamma}^T g(\underset{\sim}{x})$ is termed the Lagrangian.

Define the function

$$\phi(\underset{\sim}{z}) = \begin{bmatrix} \nabla_x L(\underset{\sim}{x},\underset{\sim}{\gamma}) \\ -C(\underset{\sim}{x}) \end{bmatrix} \qquad (3.78)$$

where $\underset{\sim}{z}^T = [\underset{\sim}{x}^T; \underset{\sim}{\gamma}^T]$ and $C'(\underset{\sim}{x})$ is the k.n matrix derivative of the k-vector $C(\underset{\sim}{x})$ whose elements are the active constraints at the feasible point x. In order to satisfy the K-T conditions at a local minimum it is necessary to have $\phi(\underset{\sim}{z}) = \underset{\sim}{0}$ and γ_j, j>q positive.

THEOREM (2nd order Sufficiency Conditions for Local Optimality)

Let $f(\underset{\sim}{x})$ and $g_j(\underset{\sim}{x})$; j=1,...q have two continuous derivatives. Let $\bar{\underset{\sim}{x}}$ be a feasible regular point satisfying the K-T necessary conditions for (P); and for every $\underset{\sim}{v} \neq \underset{\sim}{0}$ in R^k such that $\nabla_x g_j(\bar{\underset{\sim}{x}})\, v_j = 0$, j=1,...q and $\nabla_x g_j(\bar{\underset{\sim}{x}})\, v_j = 0$ j=q+1,...Q for each j>q with $\bar{\gamma}_j > 0$ let

$$\underset{\sim}{v}^T \nabla_x^2 L(\bar{\underset{\sim}{x}},\bar{\underset{\sim}{\gamma}})\, \underset{\sim}{v} > 0 \qquad (3.79)$$

where $\nabla_x^2 L(\bar{\underset{\sim}{x}},\bar{\underset{\sim}{\gamma}})$ is the matrix representing the 2nd derivative of the Lagrangian. Then $\bar{\underset{\sim}{x}}$ is an isolated local minimum

Suppose that $S(\tilde{x})$ denotes the continuously differentiable n.(n-k) matrix function wich gives the null space of $C'(\tilde{x})$. A necessary and sufficient condition for $C'(\tilde{x})\ \tilde{v} = 0$ is that $\tilde{v} = S(\tilde{x})\ \tilde{u}$ for some $\tilde{u}$. It is therefore possible to express the condition

$$\tilde{v}^T\ \nabla_x^2 L(\bar{\tilde{x}},\bar{\tilde{\gamma}})\ \tilde{v} > 0 \qquad (3.80)$$

as to have the matrix $H(\bar{\tilde{x}},\bar{\tilde{\gamma}}) = S(\bar{\tilde{x}})^T\ \nabla_x^2 L(\bar{\tilde{x}},\bar{\tilde{\gamma}})\ S(\bar{\tilde{x}})$ positive definite (PD) .

It is also important to investigate the effect of variations of $\tilde{x}$ on the solution value. This is done on the basis of

THEOREM (Implicit Function)

If at $\tilde{z} = \bar{\tilde{z}}$ there is a solution $\emptyset(\bar{\tilde{z}}) = \bar{\tilde{\eta}} = \tilde{0}$ and if the (n+k).(n+k) derivative matrix of $\emptyset(\tilde{z})$ has an inverse at every point $\tilde{z}$ then there is a continuously differentiable solution to the equations $\tilde{\eta} = \emptyset(\tilde{z})$ in a neighbourhood of $\bar{\tilde{z}}$.

The importance of this theorem is reflected in the next theorem that was originally pro ved in McCormick`s paper (1980) . Let $\tilde{x}$ be any point in $X \subset R^n$. Define for $0 \le t \le 1$ and $0 \le s \le 1$

$$\tilde{y}(t) = \tilde{x}^*\ (1-t) + \tilde{x}\ t \qquad (3.81)$$

$$\hat{\tilde{y}}(s,t) = \tilde{x}^*\ (1-s) + [\tilde{x}^*(1-t) + \tilde{x}\ t]\ s \qquad (3.82)$$

THEOREM (Isolated Global Optimality)

Let $f(\underset{\sim}{x})$ and $g_j(\underset{\sim}{x})$ j=1,...Q have two continuous derivatives and $\underset{\sim}{x}^*$ be a feasible point satisfying both the K-T necessary conditions and the 2nd order sufficiency conditions. Assume further

(i) The matrix $C'(\underset{\sim}{y}(t))$ of the active constraint derivatives has full rank.

(ii) the matrix $H[\hat{\underset{\sim}{y}}(s,t),\underset{\sim}{\gamma}^*]$ is a positive definite matrix

where

$$H[\hat{\underset{\sim}{y}}(s,t),\underset{\sim}{\gamma}^*] = S[y(t)]^T \int_0^1\int_0^1 \nabla_x^2 L[\underset{\sim}{y}(s,t),\underset{\sim}{\gamma}^*]\, ds\, dt\; S[\underset{\sim}{y}(t)] \qquad (3.83)$$

Then the derivative matrix of $\varnothing(\underset{\sim}{z})$

$$\left[\varnothing'(\underset{\sim}{z})\right] = \begin{bmatrix} \int_0^1\int_0^1 \nabla_x^2 L[\underset{\sim}{y}(s,t),\underset{\sim}{\gamma}^*]\, ds\, dt & -\int_0^1 C'[\underset{\sim}{y}(t)]^T dt \\ -\int_0^1 C'[\underset{\sim}{y}(t)]\, dt & \emptyset \end{bmatrix} \qquad (3.84)$$

has an inverse

$$\left[\varnothing'\right]^{-1} = \begin{bmatrix} \underset{\sim}{\mathbb{A}} & \underset{\sim}{\mathbb{B}}^T \\ \underset{\sim}{\mathbb{B}} & \underset{\sim}{\mathbb{C}} \end{bmatrix} \qquad (3.85)$$

where

$$\underset{\sim}{\mathbb{A}} = \underset{\sim}{S}\,\underset{\sim}{H}^{-1}\underset{\sim}{S}^T$$

$$\underset{\sim}{\mathbb{B}} = -\,[\,\underset{\sim}{I} - \underset{\sim}{S}\,\underset{\sim}{H}^{-1}\underset{\sim}{S}^T \int_0^1\int_0^1 \nabla_x^2 \underset{\sim}{L}\, ds\, dt]\; C'^{\#}$$

$$\underset{\sim}{\mathbb{C}} = -\,\underset{\sim}{C}'^{\#}\,[\int_0^1\int_0^1 \nabla_x^2 \underset{\sim}{L}\, ds\, dt - \int_0^1\int_0^1 \nabla_x^2 \underset{\sim}{L}\, ds\, dt\; \underset{\sim}{S}\,\underset{\sim}{H}^{-1}\underset{\sim}{S}^T \int_0^1\int_0^1 \nabla_x^2 \underset{\sim}{L}\, ds\, dt]\; \underset{\sim}{C}'^{\#}$$

and $\underset{\sim}{C}'^{\#}$ is the n.k matrix generalized inverse of $\int_0^1 C'[\underset{\sim}{y}(t)]\, dt$ ie $\underset{\sim}{C}'_x\, \underset{\sim}{C}'^{\#} = \underset{\sim}{I}_k$

(iii) The quantity

$$\int_0^1 f'[y(t)]\, dt\ \Big[I - S[y(t)]\ H[\hat{y}(s,t),\gamma^*]^{-1}\ S[y(t)]^T \int_0^1\!\!\int_0^1 \nabla_x^2 L[\hat{y}(s,t),\gamma^*]\, ds\, dt\Big] C'[y(t)]^{\#}$$

$$= -\int_0^1 f'[y(t)]\, dt\ B \qquad (3.86)$$

is positive component by component. Then $f(x) > f(x^*)$ for any feasible x. If (i) (ii) (iii) hold for all $x \in X \subset R$ and $f'_x C'^{\#} = (\gamma, \mu) \in R^m$ then x^* is the unique global minimizer of f over g

Proof

The existence of the inverse $[\varnothing'(z)]^{-1}$ is guaranteed by theorem hypothesis. The matrix inverse is invariant with the choice of particular null space matrix and generalized inverse $C'^{\#}$.

We have

$$\begin{bmatrix} \int_0^1 \nabla_x L[y(t),\gamma^*]\, dt \\ -C(x) \end{bmatrix} = \begin{bmatrix} \nabla_x L(x^*,\gamma^*) \\ -C(x^*) \end{bmatrix} + \left[\varnothing'(z)\right] \begin{bmatrix} \Delta x \\ \Delta \gamma \end{bmatrix} \qquad (3.87)$$

Taking x from (3.87) with the help of the inverse (3.85) and placing in the above expression

$$\Delta x = x - x^* \quad \text{and} \quad \Delta\gamma = \gamma^* - \gamma^* = \emptyset$$

$$C(x^*) = \emptyset$$

$$\nabla_x L(x^*,\gamma^*) = \emptyset \qquad (\varnothing(z^*) = \emptyset)$$

The difference between the OF values at two different points may be found by integrating the gradients of the OF along the incremental path Δx.

$$f(\underset{\sim}{x}) - f(\underset{\sim}{x}^*) = \int_0^1 f'[\underset{\sim}{y}(t)]\,dt\,\Delta x \tag{3.88}$$

$$= \int_0^1 f'[\underset{\sim}{y}(t)]\,dt\,S[\underset{\sim}{y}(t)]\,H[\hat{\underset{\sim}{y}}(s,t),\underset{\sim}{\gamma}^*]^{-1}\,S[\underset{\sim}{y}(t)]^T$$

$$\int_0^1 \nabla_x L[\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*]\,dt - \int_0^1 f'[\underset{\sim}{y}(t)]\,dt\,\mathbb{B}^T\,C(\underset{\sim}{x}) \tag{3.89}$$

Now

$$\int_0^1 C'[\underset{\sim}{y}(t)]\,dt\,S[\underset{\sim}{y}(t)] = \emptyset \tag{3.90}$$

so that

$$\int_0^1 f'[\underset{\sim}{y}(t)]\,dt\,S[\underset{\sim}{y}(t)] = \int_0^1 [f'[\underset{\sim}{y}(t)] - \underset{\sim}{\gamma}^{*T} C'[\underset{\sim}{y}(t)]]\,dt\,S[\underset{\sim}{y}(t)]$$

$$= \int_0^1 \nabla_x L[\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*]\,dt^T\,S[\underset{\sim}{y}(t)] \tag{3.91}$$

Finally

$$f(\underset{\sim}{x}) - f(\underset{\sim}{x}^*) = \int_0^1 \nabla_x L[\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*]\,dt^T\,S[\underset{\sim}{y}(t)]\,H[\hat{\underset{\sim}{y}}(s,t),\underset{\sim}{\gamma}^*]^{-1}\,S[\underset{\sim}{y}(t)]^T$$

$$\int_0^1 \nabla_x L[\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*]\,dt - \int_0^1 f'[\underset{\sim}{y}(t)]\,dt\,\mathbb{B}^T\,C(\underset{\sim}{x}) \tag{3.92}$$

Since $\underset{\sim}{x}$ is feasible $g(\underset{\sim}{x}) \geq \emptyset$. This implies that $C(\underset{\sim}{x}) \geq \emptyset$. thus both terms on the RHS are nonnegative. If $\int_0^1 \nabla_x L(\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*)\,dt\,\underset{\sim}{S} \neq \emptyset$ the first term above is positive. The second term is positive if $C(\underset{\sim}{x}) > \emptyset$ for one of j = 1,...k. The only way both terms can be equal to zero is if $\int_0^1 \nabla_x L(\underset{\sim}{y}(t),\underset{\sim}{\gamma}^*)\,dt\,\underset{\sim}{S} = \emptyset$ and $C(\underset{\sim}{x}) = \emptyset$. If both these condition hold it is easy to show that $\Delta\underset{\sim}{x} = \underset{\sim}{x} - \underset{\sim}{x}^* = \underset{\sim}{\emptyset}$. Thus $f(\underset{\sim}{x}) > f(\underset{\sim}{x}^*)$ when $\underset{\sim}{x} \neq \underset{\sim}{x}^*$.

QED

Refering to expression (3.92) the difference $f(\underset{\sim}{x})-f(\underset{\sim}{x}^*)$ is seen to consist of two contributions. The first is due to a step from $\underset{\sim}{x}$ to $\bar{\underset{\sim}{x}}$ which drives to $\underset{\sim}{\emptyset}$ the Lagrangian derivative ;the second from $\bar{\underset{\sim}{x}}$ to $\underset{\sim}{x}^*$ seeks to satisfy the active constraints violated in the first step.

$$f(\underset{\sim}{x}^{*}) - f(\underset{\sim}{x}) = [f(\underset{\sim}{\bar{x}}) - f(\underset{\sim}{x})] + [f(\underset{\sim}{x}^{*}) - f(\underset{\sim}{\bar{x}})] \qquad (3.93)$$

In terms of the inverse function approach $\underset{\sim}{\bar{x}}$ consists of the first n components of $\emptyset^{-1}[\underset{\sim}{0}^{T};g(\underset{\sim}{x})^{T}]^{T}$. The exact difference can be computed by integrating from $[\underset{\sim}{x}^{T},\underset{\sim}{\gamma}^{T}]$ to $[\underset{\sim}{\bar{x}}^{T},\underset{\sim}{\bar{\gamma}}^{T}]$ along $\emptyset^{-1}[t\nabla_{x}L(\underset{\sim}{x},\underset{\sim}{\gamma})^{T};g(\underset{\sim}{x})^{T}]^{T}$. The exact second term $f(\underset{\sim}{x}) - f(\underset{\sim}{\bar{x}})$ can then be obtained by integrating from $[\underset{\sim}{\bar{x}}^{T};\underset{\sim}{\bar{\gamma}}^{T}]$ to $[\underset{\sim}{x}^{*T};\underset{\sim}{\gamma}^{*T}]$ along the curve $\emptyset^{-1}[\underset{\sim}{0}^{T};g(\underset{\sim}{x})^{T}]^{T}$.
In order to illustrate the meaning of the difference $f(\underset{\sim}{x}) - f(\underset{\sim}{x}^{*})$ consider the optimal solution (if there is one) of the following problem

$$\text{Min} \quad f(\underset{\sim}{w}) \qquad \underline{\text{st}} \qquad g(\underset{\sim}{w}) = g(\underset{\sim}{x});\underset{\sim}{w}\in X \qquad (3.94)$$

Indeed the first term under the integral is an approximation to $f(\underset{\sim}{x}) - f(\underset{\sim}{\bar{x}})$. The difference in the optimal objective function value resulting from a perturbation in the RHS of the constraints is approximated by the sum of the Lagrange multipliers times the perturbations. The vector premultiplying $g(\underset{\sim}{x})$ is the integral of the vector of 2nd order multiplier estimates for the minimization of the Lagrangian function. This can be viewed as the appropriate vector of Lagrange multipliers needed to approximate $f(\underset{\sim}{x}^{*}) - f(\underset{\sim}{\bar{x}})$.

3.7.1. Minimum volume design

Consider the synthesis problem subject to stress/displacement constraints and area bounds in the nodal stiffness format. The only nonlinear equations of the system correspond to the equilibrium between member forces and loads. In the hypothesis of a single loading condition they may be written as

$$\underset{\sim}{K}\,\underset{\sim}{d} = \underset{\sim}{\lambda} \tag{3.95}$$

we recall that $\underset{\sim}{\lambda}$ is a β-vector $\underset{\sim}{d}$ is a β-vector $\underset{\sim}{K}$ is a $\beta.\beta$ matrix derived from

$$\underset{\sim}{A}\,\underset{\sim}{\mathbf{K}}\,\underset{\sim}{A}^{T}$$

where $\underset{\sim}{A}$ is the direction cosine $\beta.m$ matrix and $\underset{\sim}{\mathbf{K}}$ is a diagonal m.m matrix whose elements are member stiffnesses. By operating in the subspace defined by these equality constraints some properties of the problem may be derived. The active constraint vector $C(\underset{\sim}{a},\underset{\sim}{d})$ has β elements. Its derivative is the $\beta.(m+\beta)$ matrix $C'(\underset{\sim}{a},\underset{\sim}{d})$

$$\nabla_{a} C(\underset{\sim}{a},\underset{\sim}{d}) = \underset{\sim}{A}\,\underset{\sim}{V} \tag{3.96}$$

$$\nabla_{d} C(\underset{\sim}{a},\underset{\sim}{d}) = \underset{\sim}{K} \tag{3.97}$$

where $\underset{\sim}{V}$ is a diagonal matrix whose elements v_j are member stresses written as linear combinations of the nodal displacements.

$$\underset{\sim}{v} = \underset{\sim}{s} = \underset{\sim}{L}^{-1}\,\underset{\sim}{A}^{T}\,\underset{\sim}{d} \tag{3.98}$$

$$\underset{\sim}{C}'(a,d) = [\,\underset{\sim}{A}\,\underset{\sim}{V} \;\vdots\; \underset{\sim}{K}\,] \tag{3.99}$$

The matrix function giving the null space of $C'(a,d)$ is a $(m+\beta).m$ matrix $S(a,d)$ such that

$$C'(\underset{\sim}{a},\underset{\sim}{d})\; S(\underset{\sim}{a},\underset{\sim}{d}) = 0 \tag{3.100}$$

or

$$S(a,d) = \begin{bmatrix} I \\ \hline -K^{-1} A V \end{bmatrix} \qquad (3.101)$$

S(a,d) is a choice of all solutions satisfying these homogeneous equations and has rank m. The $(m+\beta).(m+\beta)$ matrix giving the second derivative of the Lagrangian is a copositive matrix function of the values assumed by the Lagrange multipliers.

$$\nabla^2_{(a,d)} L[(a,d),\gamma] = \begin{bmatrix} 0 & -P A^T \\ -A P & 0 \end{bmatrix} \qquad (3.102)\ (3.103)$$

where P is a diagonal matrix whose elements p_j are a linear combination of the Lagrange multipliers corresponding to the bilinear equalities.

$$p = L^{-1} A^T \gamma \qquad (3.104)$$

It is well known that a copositive matrix has a set of symmetric eigenvalues so that convexity properties of the points lying on the equality constrained subspace will depend on positive definiteness of the matrix

$$H[(a,d),\gamma] = S(a,d)^T \nabla^2_{(a,d)} L[(a,d),\gamma]\ S(a,d) \qquad (3.105)$$

Now we will check the regularity assumption for any feasible point (a,d). The assembled stiffness matrix is a $\beta.\beta$ real symmetric positive definite matrix of rank β. We can infer that

$$C'(a,d) = [\ A V \ \vdots\ K\ ] \qquad (3.106)$$

has rank β for any feasible $(\underset{\sim}{a},\underset{\sim}{d})$.

We also have $\underset{\sim}{K}^{-1}$ PD because the same holds true for $\underset{\sim}{K}$. In order to define the matrix $H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]$ it is necessary to have not only a feasible point but also a set of multipliers $(\underset{\sim}{\gamma},\underset{\sim}{\mu})$ satisfying the K-T necessary conditions.

$$H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}] = S(\underset{\sim}{a},\underset{\sim}{d})^T \nabla^2_{(a,d)} L[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]\ S(\underset{\sim}{a},\underset{\sim}{d})$$

$$= [\underset{\sim}{I}_m \ \vdots\ -\underset{\sim}{V}\,\underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}] \begin{bmatrix} \underset{\sim}{0} & -\underset{\sim}{P}\,\underset{\sim}{A}^T \\ -\underset{\sim}{A}\,\underset{\sim}{P} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{I}_m \\ \hline -\underset{\sim}{K}^{-1}\,\underset{\sim}{A}\,\underset{\sim}{V} \end{bmatrix}$$

$$= \underset{\sim}{V}\,\underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}\underset{\sim}{A}\,\underset{\sim}{P} + \underset{\sim}{P}\,\underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}\underset{\sim}{A}\,\underset{\sim}{V} \qquad (3.107)$$

$\underset{\sim}{Q} = \underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}\underset{\sim}{A}$ is a m.m symmetric matrix positive semidefinite and singular of rank β since $\alpha = m - \beta$ of its rows (or columns) can be expressed as a linear combination of the remaining rows (columns). The general element h_{ij} of $H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]$ can be related to the corresponding element q_{ij} of $\underset{\sim}{Q}$ by

$$h_{ij} = q_{ij}(p_i v_j + p_j v_i) \qquad i=1,\dots m \quad j=1,\dots m \qquad (3.108)$$

$H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]$ is a symmetric matrix given by the sum of two matrices where the second is the transpose of the first. A relationship linking the elements of $\underset{\sim}{V}$ and $\underset{\sim}{P}$ is found if the Lagrangian is differentiated with respect to the design variables

$$\nabla_a L[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}] = 0 \qquad (3.109)$$

we have

$$\underset{\sim}{1} - \underset{\sim}{\mu}_a^l + \underset{\sim}{\mu}_a^u - \underset{\sim}{V}\,\underset{\sim}{A}^T \underset{\sim}{\gamma} = \emptyset \qquad (3.110)$$

where $\underset{\sim}{\mu}_a^l$, $\underset{\sim}{\mu}_a^u$ are the Lagrange multipliers corresponding to the design variables touching their lower and upper boundaries respectively.

If we suppose that no active bounds on the design variables would occur at the local optima all $\underset{\sim}{\mu}_a^l = \underset{\sim}{\mu}_a^u = \underset{\sim}{\emptyset}$ and

$$\underset{\sim}{V}\,\underset{\sim}{A}^T \underset{\sim}{\gamma} = \underset{\sim}{1} \qquad (3.111)$$

ie

$$\underset{\sim}{V}\,\underset{\sim}{P}\,\underset{\sim}{L} = \underset{\sim}{L} \qquad (3.112)$$

or

$$\underset{\sim}{P} = -\underset{\sim}{V}^{-1} \qquad (3.113)$$

In this case

$$H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}] = \underset{\sim}{V}\,\underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}\,\underset{\sim}{A}\,\underset{\sim}{V}^{-1} + \underset{\sim}{V}^{-1}\underset{\sim}{A}^T\,\underset{\sim}{K}^{-1}\underset{\sim}{A}\,\underset{\sim}{V} \qquad (3.114)$$

Although both matrices result from similarity transformation on $\underset{\sim}{Q}$ sharing their eigenvalues with it the symmetric matrix representing their sum need not be positive definite or even PSD. Besides,the case considered above is unrealistic in the sense that it would lead to an inconsistent set of relations (3.112) : These would be m equations for the unknowns in the $\underset{\sim}{V}$ array with $m > \beta$ and it may be verified that the system is insolvable. We now consider the inclusion of a number of active constraints on the design variables which make the system compatible and we attempt to analyse their effect on the character of the matrix $\underset{\sim}{H}$.With the hypothesis of a single loading condition we would need to have at least α active bounds on the member areas in order to make soluble the system given by the KT necessary conditions related to the design variables.

We have

$$p_i v_i = 1 - \mu^l_{ia}/l_i + \mu^u_{ia}/l_i \qquad i=1,...m \qquad (3.115)$$

where $\mu^l_{ia}, \mu^u_{ia} \geq 0$

Suppose any $1 - \mu^l_{ia}/l < 0$ ie at least one of the Lagrange multipliers corresponding to active lower bounds on the area exceeds the member length. A necessary condition for the positive definiteness of a real symmetric matrix states that all diagonal elements must be positive. Following the assumption that h_{ii} is negative, this is not met. $\underset{\sim}{H}$ would become an indefinite matrix when restricting the domain to the subspace defined by the equality equilibrium relations. It is therefore necessary to investigate the effect caused by the insertion of any active area bound inequality in the derivative matrix $C'(\underset{\sim}{a},\underset{\sim}{d})$ of the active constraint vector $C(\underset{\sim}{a},\underset{\sim}{d})$. Assume the columns of $C'(\underset{\sim}{a},\underset{\sim}{d})$ numbered so that the first m columns would correspond to derivatives with respect to the design variables and the remaining β are associated with the derivatives wrt the state variables.

By including an active inequality corresponding to the lower area bound l, $C'(\underset{\sim}{a},\underset{\sim}{d})$ would be increased by one row from the matrix previously defined. In this row the element of the column I will be a unity and all other elements will be zero. The rank of $C'(\underset{\sim}{a},\underset{\sim}{d})$ is increased by 1 since this new row is linearly independent of the rows of the basis defined by the $\underset{\sim}{K}$ matrix

$$\underset{\sim}{C}'(\underset{\sim}{a},\underset{\sim}{d}) = \left[\begin{array}{ccc} \underset{\sim}{A} \ \underset{\sim}{V} & & \underset{\sim}{K} \\ \hline \emptyset..1\ \emptyset.... & & \emptyset \end{array}\right] \quad (3.116)$$

The matrix that gives its null space can be derived from S $(\underset{\sim}{a},\underset{\sim}{d})$ by suppressing its column 1

$$\underset{\sim}{C}'(\underset{\sim}{a},\underset{\sim}{d})\ \underset{\sim}{S}(\underset{\sim}{a},\underset{\sim}{d}) = \emptyset \quad (3.117)$$

ie

$$\left[\begin{array}{ccc} \underset{\sim}{A}^1 & \underset{\sim}{V}^1 & \underset{\sim}{K} \\ \hline \emptyset\ 1 & \emptyset & \emptyset \end{array}\right] \left[\begin{array}{c} \underset{\sim}{I} \\ \hline \emptyset\emptyset\emptyset\emptyset \\ \hline \underset{\sim}{I} \\ \hline \underset{\sim}{K}^{-1}\underset{\sim}{A}^1\underset{\sim}{V}^1 \end{array}\right] = \emptyset \quad (3.118)$$

where $\underset{\sim}{A}^1$ and $\underset{\sim}{V}^1$ are obtained from $\underset{\sim}{A}$ and $\underset{\sim}{V}$ respectively by also deleting column 1 (and row 1 in V).

The matrix H $[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]$ is now a (m-1).(m-1) matrix given by

$$H[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}] = \underset{\sim}{V}^1\ \underset{\sim}{A}^{1^T}\ \underset{\sim}{K}^{-1}\underset{\sim}{A}^1\ \underset{\sim}{P}^1 + \underset{\sim}{P}^1\ \underset{\sim}{A}^{1^T}\ \underset{\sim}{K}^{-1}\ \underset{\sim}{A}^1\ \underset{\sim}{V}^1 \quad (3.119)$$

where P^1 is a diagonal matrix derived from $\underset{\sim}{P}$ by suppressing its row and column 1. In $\underset{\sim}{H}$ the row and column corresponding to the index of the active area constraint would vanish. Therefore the negative eigenvalue due to the negative diagonal element should no longer be considered. If more than one lower bounds on the areas are active they can be equally inserted in the active constraint derivative matrix. The final matrix $\underset{\sim}{H}$ is obtained from the one derived in the equality constrained subspace by deleting all the rows and columns corresponding to indexes of active areas. Let their number be r. For a single loading case $\underset{\sim}{H}$ is a square matrix whose dimension (m-r) is not greater than β . $\underset{\sim}{A}^{r^T} \underset{\sim}{K}^{-1} \underset{\sim}{A}^r$ is a symmetric positive definite (m-r).(m-r) matrix of rank $m-r \leq \beta$ whose rows and columns are linear combinations of the rows and columns of $\underset{\sim}{K}^{-1}$.

We may recall Gershgorin's circle theorem which states the following :

Each eigenvalue of a n.n matrix $\underset{\sim}{B}$ satisfies at least one of the equalities

$$\left|\lambda - b_{ii}\right| < r_i \; ; \; r_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} b_{ij} \tag{3.120}$$

for i=1,..n

In other words every eigenvalue of $\underset{\sim}{B}$ lies in at lest one of the circles with centre b_{ii} and radius r_i in the complex λ-plane.

Since we are dealing with real matrices every eigenvalue must lie within an interval of the real axes where the mid point is defined by one of the diagonal elements of the matrix and its length does not exceed the double of the sum of the absolute values of the off-diagonal elements along the corresponding row. In practice Gershgorin's theorem is often used to estimate the eigenvalues of a matrix $\underset{\sim}{B}$ where the eigenvalues are much smaller than the diagonal elements. Instead of applying the theorem directly to B much more accurate bounds can often be found by first applying a similarity transformation $\underset{\sim}{C}^{-1}\underset{\sim}{B}\,\underset{\sim}{C}$ where $\underset{\sim}{C}$ is diagonal.

Let

$$\underset{\sim}{C} = \begin{bmatrix} r & & \\ & 1 & \\ & & 1 \end{bmatrix} \tag{3.121}$$

Suppose for r=1 the intervals are disjoint. By increasing r the length of the first interval will decrease whereas the length of the other intervals will increase. Clearly there will be a value of r for which the length of the first interval will be as small as possible while still defining a lower bound on the eigenvalue of the matrix B corresponding to that diagonal element. The same procedure can be extended to determine lower bounds on the remaining eigenvalues of the matrix.

This procedure could be applied to the matrix in (3.114) in order to determine lower bounds on the values of the eigenvalues . Its diagonal elements will coincide with those of $2(\underset{\sim}{A}^{r^T}\underset{\sim}{K}^{-1}\underset{\sim}{A}^{r})$ that is a symmetric positive definite matrix while the remaining elements will suffer slight changes. We may expect all the eigenvalues to be still bracketed in the nonnegative real axes. While the above argument for the positive definiteness of $\underset{\sim}{H}$ in general is not as rigorous as might be desired the property can at least be checked numerically in particular instances and this idea is followed up in subsquent sections.

The remaining assumptions of the theorem that ensures uniqueness require that (for all set of variables and Lagrange multipliers, related to them by a first order estimate) the matrix $H\,[(\underset{\sim}{a},\underset{\sim}{d}),\underset{\sim}{\gamma}]$ to be positive definite. By defining a set of design variables the set of state variables is uniquely determined. The matrix $\underset{\sim}{P}$ of Lagrange multipliers is also uniquely determined if the system is soluble. In this case due to Gershgorin's theorem we may expect positive definiteness in $\underset{\sim}{H}$ when $C\ (\underset{\sim}{a},\underset{\sim}{d})$ incorporates the constraints corresponding to lower bounds on the member areas.

If the Lagrange multipliers corresponding to lower bounds on the areas are smaller than the corresponding member length this can be regarded as a special case of the $\underset{\sim}{H}$ matrix where negative eigenvalues are not apparent. Here the active lower bounds on the areas should be also considered in the matrix $C\ (\underset{\sim}{a},\underset{\sim}{d})$. We also remark that the nature of matrix $\underset{\sim}{H}$ is not altered if the structure is subjected to a number of loading cases.

$$H[(\underset{\sim}{a},\underset{\sim}{d}),\gamma] = \sum_{k=1}^{l} (\underset{\sim}{V}^{k}\,\underset{\sim}{A}^{T}\,\underset{\sim}{K}^{-1}\underset{\sim}{A}\,\underset{\sim}{P}^{k} + \underset{\sim}{P}^{k}\underset{\sim}{A}^{T}\underset{\sim}{K}^{-1}\underset{\sim}{A}\,\underset{\sim}{V}^{k}) \qquad (3.122)$$

We remark that the contents of this section are not applicable to the general bilinearly constrained problem unless it is cast in the special least volume design format. Its special features are provided by the inclusion of the compatibility constraints themselves depending on the direction cosine matrix and on the member lengths.

3.7.2. Convergence to nonoptimal solutions

When primal feasible methods are applied to our nonconvex problem a new sort of trouble may occur. It is the case when the algorithm converges to a solution possessing negative Lagrange multipliers corresponding to linear inequalities. For the ten bar truss possessing multiple solutions we will give an example of such behaviour:

member areas

$a_1=65.94$ $a_2=.1$ $a_3=27.11$ $a_4=25.58$ $a_5=8.45$

$a_6=.1$ $a_7=.1$ $a_8=53.78$ $a_9=36.17$ $a_{10}=.1$

member stresses

$s_1=.45$ $s_2=.00$ $s_3=-.37$ $s_4=-.39$ $s_5=1.17$

$s_6=.00$ $s_7=1.15$ $s_8=-.52$ $s_9=.39$ $s_{10}=.00$

nodal displacements

$d_1=.45$ $d_2=.00$ $d_3=-.76$ $d_4=-3.5$

$d_5=.45$ $d_6=-1.5$ $d_7=-.37$ $d_8=-2.68$

objective function ...254.755

The solution would therefore correspond to a dual infeasible problem where the multipliers coresponding to active area inequality bounds are:

$$\mu^1_{2a}=1.\ ;\ \mu^1_{6a}=1.\ ;\ \mu^1_{7a}=-1.99 < 0\ ;\ \mu^1_{10a}=1.414$$

The algorithm should be corrected by the introduction of an alternative iteration up-date formula.

3.7.3. Multiple solution due to nonconvex behaviour

Much more serious is the manifestation of the fully nonconvex behaviour exemplified in 3.6.. In the case of the local solution 223.339 the Lagrange multipliers corresponding to lower bounds on the areas are

$$\mu^{1}_{2a}=1. \; ; \mu^{1}_{5a}=5.07 \; ; \mu^{1}_{6a}=1. \; ; \mu^{1}_{10a}=1.414$$

whereas the global solution 219.929 is associated with

$$\mu^{1}_{2a}=4.91 \; ; \mu^{1}_{5a}=9.93 \; ; \mu^{1}_{10a}=12.47$$

By inspection the $\underset{\sim}{H}$ matrix is in both cases indefinite when restricting the domain to the equality constrained subspace. By integrating in the active constraint vector the inequalities corresponding to lower bounds on the areas $\underset{\sim}{H}$ would become positive definite.

3.7.4. Ten bar truss with one member stronger

It is possible to have an indefinite matrix $\underset{\sim}{H}$ in the subspace defined by the equilibrium equations and the local optimum to be unique in the domain of the problem. It is the case when the displacement constraints are absent and the limit stress on bar 9 is increased to 3.75 (instead of 2.5). We have

member areas

$a_1=7.9$ $a_2=.1$ $a_3=8.1$ $a_4=3.9$ $a_5=.1$

$a_6=.1$ $a_7=5.8$ $a_8=5.52$ $a_9=3.68$ $a_{10}=.14$

member stresses

$s_1=2.5$ $s_2=2.5$ $s_3=-2.5$ $s_4=-2.5$ $s_5=0.00$

$s_6=2.5$ $s_7=2.5$ $s_8=-2.5$ $s_9=3.75$ $s_{10}=-2.5$

nodal displacements

$d_1=5.$ $d_2=-20.$ $d_3=-5.$ $d_4=-22.5$

$d_5=2.5$ $d_6=-7.5$ $d_7=-2.5$ $d_8=-7.5$

corresponding to the minimum volume ... 41.6

The Lagrange multipliers corresponding to the lower bounds on the variables are

$$\mu^1_{2a}=1.66\ ;\ \mu^1_{5a}=1.\ ;\ \mu^1_{6a}=0.$$

The matrix H in the subspace defined by the equilibrium relations would have one negative eigenvalue corresponding to a negative diagonal element in the 2nd row of this real symmetric matrix. In order to check whether this minimum is unique we would increase the rank of the active constraint derivative matrix by 3 by inserting the active inequalities related to the lower bounds on a_2, a_5, a_6. The result

$$\int_0^1 f'[\underset{\sim}{y}(t)]\ \underset{\sim}{B}^T\, dt$$

of the product of the constant

$$\int_0^1 f'[\underset{\sim}{y}(t)]\, dt = [\ \underset{\sim}{1}^T \mid \underset{\sim}{0}^T]^T \qquad (3.123)$$

by the matrix given by

$$[\underset{\sim}{I} - S[\underset{\sim}{y}(t)]\, H[\hat{\underset{\sim}{y}}(s,t),\gamma]^{*-1}\, S[\underset{\sim}{y}(t)]^{T}$$

$$\nabla_x^2 L\, [\hat{\underset{\sim}{y}}(s,t),\underset{\sim}{\gamma}]^{*}\, C[\underset{\sim}{y}(t)]' \qquad (3.124)$$

is a (β+3) vector.

For $t = 0$ we have $\underset{\sim}{y}(t) = \underset{\sim}{x}^*$ and all components are positive. We remark that the first β components can always be taken as positive since they are related to equality constraints so that only the inequalities giving a positive value for γ_i should be selected. This vector would correspond to the 2nd order estimates used by Gill and Murray (1977) to determine whether or not a constraint should be considered binding at the local solution.

The same vector when multiple optima solution is exhibited is no longer positive component by component. In fact the components corresponding to a_2 and a_{10} are negative whereas the vector obtained from the local solution has the component corresponding to a_{10} negative. This appear to be a necessary condition for multiple optimality.

3.7.5. Iterative procedure

A Lagrangian based iterative procedure for convex programming problems with a fast rate of convergence (Gill and Murray (1977)) will be briefly described in this section. Suppose a current value of $\underset{\sim}{x}^* = (\underset{\sim}{a},\underset{\sim}{d})$ is $\hat{\underset{\sim}{x}}$ and $\hat{\underset{\sim}{x}}$ is not $\underset{\sim}{x}^*$. An estimate of the multipliers corresponding to the bilinear constraints could be taken as

$$= f'(\underset{\sim}{x})^T \ C'(\underset{\sim}{x}) \# \qquad (3.125)$$

This could be used in the computation of the Lagrangian Hessian matrix $\nabla_x^2 L [\underset{\sim}{x}, \underset{\sim}{\gamma}]$

The signs of the Lagrange multipliers will identify the inequality constraints whose removal from the active set will make it possible to compute a search direction both feasible and descendent. A second estimate of the multipliers is therefore given by

$$f'(\underset{\sim}{x}) \ [\underset{\sim}{I} - \underset{\sim}{S}(x) \ \underset{\sim}{H}(\underset{\sim}{x}, \underset{\sim}{\gamma}) \ \underset{\sim}{S}(x)^T \nabla_x^2 L [(\underset{\sim}{x}, \underset{\sim}{\gamma})] \ C'(\underset{\sim}{x}) \# \qquad (3.126)$$

This is the one which determines whether or not a constraint should be considered binding at the local solution since the only possibility of obtaining a better estimate is to move off one or more of the active constraints. When a local solution is finally computed local optimality tests are carried out and statements about uniqueness can be made.

The range of applicability of this method is limited to structures with continuous area variables and the problem of finding the global solution may become awkward when multiple solution are present. We will develop methods that although more appropriate when the optimization is made with respect to discrete design variables are also globally convergent with continuous variables.

CHAPTER FOUR

BRANCH AND BOUND METHODS

4.1. Introduction

The existence of multiple solutions for the bilinearly constrained problem lead us to consider methods of solution more appropriate for this nonconvex behaviour. The Branch and Bound strategy is a globally convergent procedure consisting of the transformation of the general nonconvex domain into a sequence of intersecting convex domains by the use of underestimating convex functions. The two main ingredients are a combinatorial tree with appropriately defined nodes and some upper and lower bounds to the final solution associated with each node of the tree. It is then possible to eliminate a large number of possible solutions without evaluating them. It is well known that a local solution to a problem possessing a convex objective function and being restricted by convex inequalities is also its global solution.

Separable nonconvex programming by Branch and Bound was initially described by Lawler and Wood (1966) but latter Soland (1971) has generalized the method to include nonconvex separable constraint functions. His algorithm is applicable to the present study because the bilinear constraint expressions can be converted into separable form ie: a sum of functions of one variable possibly nonconvex. If in the more general case of a continuous function over a compact set one replaces the objective function by the largest convex function that fits below it [Fig 4.1] and the domain by the convex hull corresponding to the compact set it is known that the solution set of the latter convex problem contains the solution set of the original problem

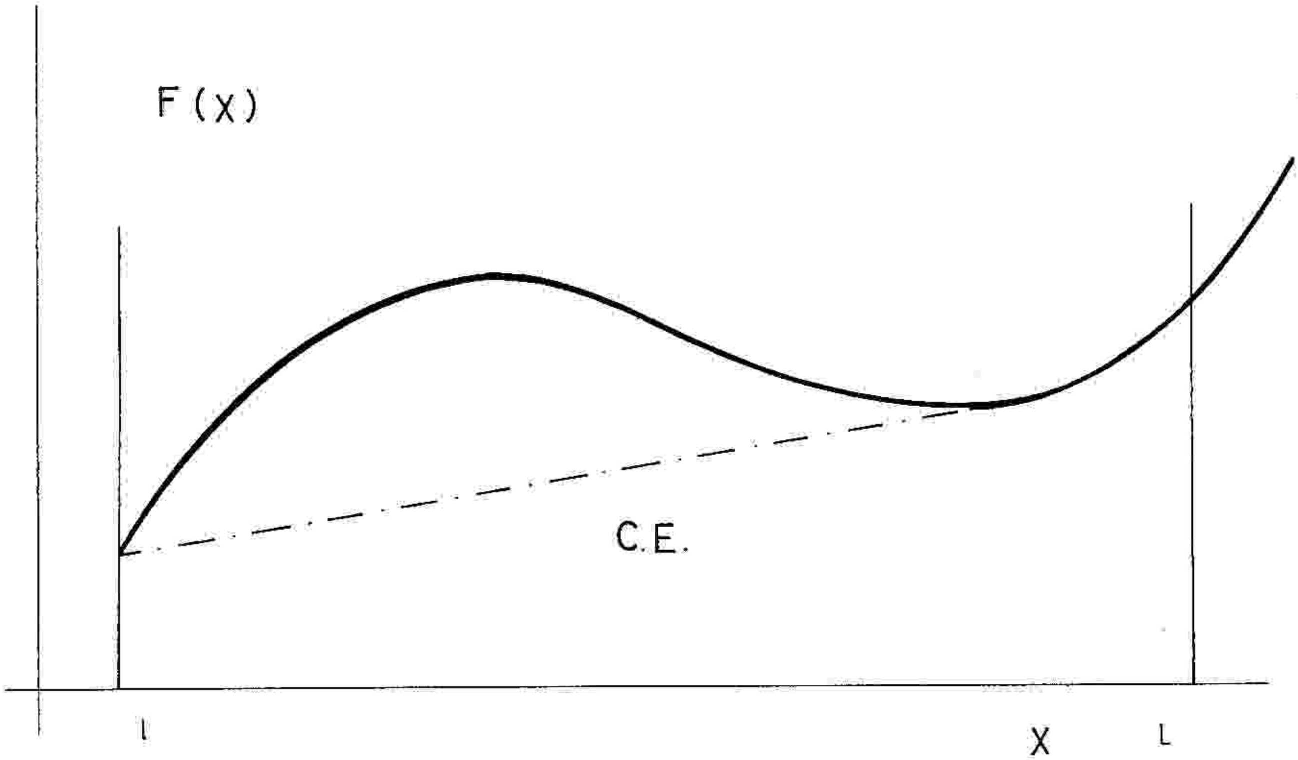


Fig 4.1 Convex Underestimate of F(X)

Reeves published in 1973 a different branch and bound based algorithm for the minimization of nonconvex allquadratic programming. The algorithm uses local minima and elimination intervals surrounding them prior to branching and bounding. His approach can be classified as "inside-out" ie: for how large a region is a local minimizer global. In contrast , Soland's "outside-in" strategy determines by how much can the region of possible optimality be reduced by partitioning the domain defining the hyper-rectangle of bounds on the variables.

We will start this Chapter by generally describing the algorithm that historically was first and we will apply it to the least volume design. Several underestimating functions will be tried out and on the basis of computational experience some conclusions will be drawn. Reeves' algorithm is presented in the last section of this Chapter and its advantages and drawbacks will be discussed when compared with Soland's method.

4.2. Separable Programming

As it has been previously mentioned the factorable terms in the equilibrium equations could be transformed into a sum of functions of one variable by inducing a rotation of $\pi/4$ to the axis of the system. Let the new variables be x_i^k and y_i^k

$$a_i = x_i^k + y_i^k \qquad s_i^k = x_i^k - y_i^k \qquad (4.1)$$

$$x_i^k = (a_i + s_i^k)/2 \qquad y_i^k = (a_i - s_i^k)/2 \qquad (4.2)$$

$$i=1,\dots n \qquad k=1,\dots l$$

We remark that for multiple loading we are therefore penalized by an increase in the dimensionality of the problem. Range constraints on the new variables will correspond to the simple bounds on the variables $\underset{\sim}{a}$ and $\underset{\sim}{s}^k$.

$$\underset{\sim}{a}_l \leq \underset{\sim}{x}^k + \underset{\sim}{y}^k \leq \underset{\sim}{a}_u \qquad (4.3)$$

$$\underset{\sim}{s}_l^k \leq \underset{\sim}{x}^k - \underset{\sim}{y}^k \leq \underset{\sim}{s}_u^k \qquad (4.4)$$

The bounds on the new variables (that will be necessary to the definition of convex envelope) are equally obtained from the bounds on the variables $\underset{\sim}{a}$ and $\underset{\sim}{s}^k$ ie:

$$\underset{\sim}{x}_l^k = (\underset{\sim}{a}_l + \underset{\sim}{s}_l^k)/2 \leq \underset{\sim}{x}^k \leq \underset{\sim}{x}_u^k = (\underset{\sim}{a}_u + \underset{\sim}{s}_u^k)/2 \qquad (4.5)$$

$$\underset{\sim}{y}_l^k = (\underset{\sim}{a}_l - \underset{\sim}{s}_u^k)/2 \leq \underset{\sim}{y}^k \leq \underset{\sim}{y}_u^k = (\underset{\sim}{a}_u - \underset{\sim}{s}_l^k)/2 \qquad (4.6)$$

All the remaining linear constraints will be transformed in expressions of the same type

$$\underset{\sim}{B}^T \underset{\sim}{L} \; (\underset{\sim}{x}^k - \underset{\sim}{y}^k) = \emptyset \qquad \text{Compatibility} \qquad (4.7)$$

$$\underset{\sim}{d}_l^k \leq \underset{\sim}{D} \; (\underset{\sim}{x}^k - \underset{\sim}{y}^k) \leq \underset{\sim}{d}_u^k \qquad \text{Displacements} \qquad (4.8)$$

$$-\underset{\sim}{x}^k + \underset{\sim}{y}^k - \underset{\sim}{E} \; x^k - \underset{\sim}{E} \; y^k \leq \underset{\sim}{\emptyset} \qquad \text{Buckling} \qquad (4.9)$$

Each separable equality equilibrium constraint will be represented by a sum of strictly convex and strictly concave functions. Only the convex envelopes of the latter will be required. First we replace each equality bilinear constraint by a pair of inequality constraints ie

$$\underset{\sim}{A} \; (\underset{\sim}{x}^{k^2} - \underset{\sim}{y}^{k^2}) = \underset{\sim}{\lambda}^k \qquad (4.10)$$

can be written

$$\underset{\sim}{A} \; (\underset{\sim}{x}^{k^2} - \underset{\sim}{y}^{k^2}) \leq \underset{\sim}{\lambda}^k \qquad (4.11)$$

$$\underset{\sim}{A} \; (\underset{\sim}{y}^{k^2} - \underset{\sim}{x}^{k^2}) \leq - \underset{\sim}{\lambda}^k \qquad (4.12)$$

The convex envelope of a concave function of a single variable is a linear function passing through the endpoints of the graph of the given function [Fig 4.2]

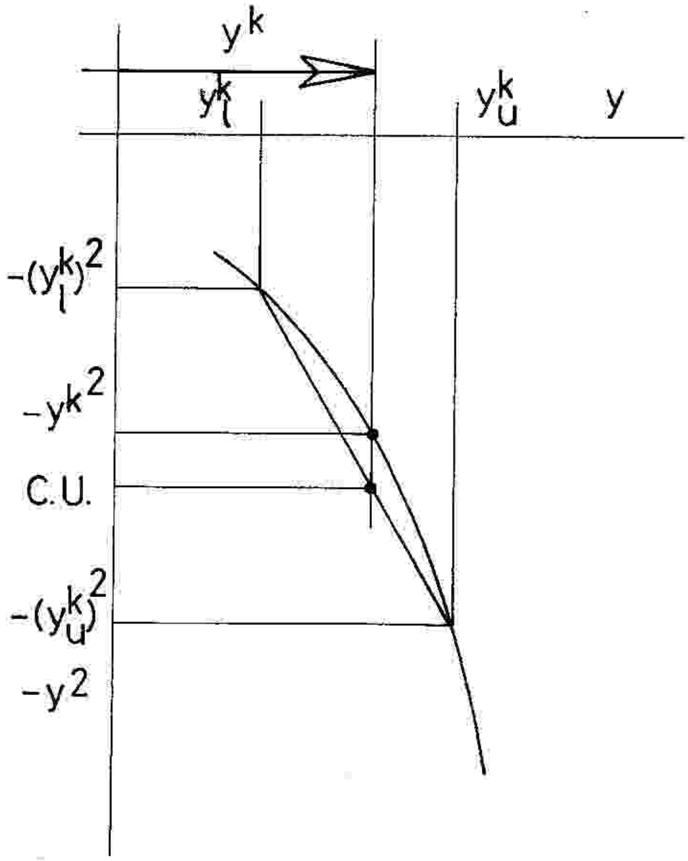


Fig 4.2 Convex Envelope of a Concave Function

$$CE = [(-y_u^{k^2} + y_l^{k^2})/(y_u^k - y_l^k)]\ (y^k - y_l^k) - y_l^{k^2}$$

$$= -(y_u^k + y_l^k)\ (y^k - y_l^k) - y_l^{k^2}$$

$$= -(y_u^k + y_l^k)\ y^k + y_l^k\, y_u^k \leq -y^{k^2} \qquad (4.13)$$

similarly

$$-(x_u^k + x_l^k)\ x^k + x_l^k\, x_u^k \leq -x^{k^2} \qquad (4.14)$$

The convex envelopes corresponding to the equilibrium constraints would still be bilinear although strictly convex

$$\mathbf{A}\ \mathbf{x}^{k^2} - \mathbf{A}\ (\mathbf{y}_u^k + \mathbf{y}_l^k)\ \mathbf{y}^k \leq \boldsymbol{\lambda}^k - \mathbf{y}_u^k \mathbf{y}_l^k \qquad (4.15)$$

$$\underset{\sim}{A}\,\underset{\sim}{y}^{k^2} - \underset{\sim}{A}\,(\underset{\sim}{x}_u^k + \underset{\sim}{x}_l^k)\,\underset{\sim}{x}^k \leq - \underset{\sim}{\lambda}^k - \underset{\sim}{x}_u^k\,\underset{\sim}{x}_l^k \qquad (4.16)$$

Tight convex underestimate subproblems although having linear objective function would still present curvature in the constraint set. A Newton iteration method using second derivatives of the Lagrangian could solve each subproblem sucessfully. Due to the fact that the number of subproblems to be solved cannot be neglected and a Newton iteration procedure is necessary for each such problem the whole process would become very inefficient. A more crude approximation to the constraints is therefore justifiable if the underestimating subproblem could be solvable by a code for LP.

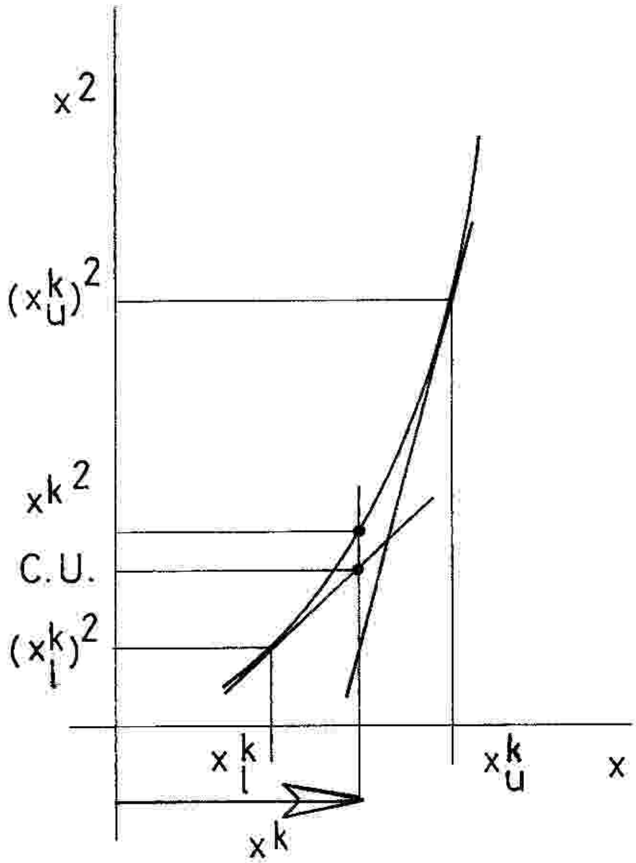


Fig 4.3 Convex Envelope given by Endpoint tangents

The envelope that gives the maximum over the two endpoints tangents may represent a linear approximation of the quadratic coinciding with the function at its boundaries. Such convex underestimate of the nonlinear convex terms presents a big disadvantage: The problem dimensionality will increase since a new constraint will arise from each such term [Fig 4.3]

$$2\, x_u^k x^k - x_u^{k^2} \leq x^{k^2} \qquad (4.17)$$

$$2\, x_l^k x^k - x_l^{k^2} \leq x^{k^2} \qquad (4.18)$$

$$2\sum_{i=1}^{\xi} A\, x_{iu}^k x_i^k + 2\sum_{j=1}^{\eta} A\, x_{jl}^k x_j^k - \underset{\sim}{A}\,(\underset{\sim}{y}_u^k + \underset{\sim}{y}_l^k)\, \underset{\sim}{y}^k \leq \sum_{i=1}^{\xi} x_{iu}^{k^2} + \sum_{j=1}^{\eta} x_{jl}^{k^2} - \underset{\sim}{y}_u^k \underset{\sim}{y}_l^k + \underset{\sim}{\lambda}^k \qquad (4.19)$$

$$2\sum_{i=1}^{\xi} A\, y_{iu}^k y^k + 2\sum_{j=1}^{\eta} A\, y_{jl}^k y^k - \underset{\sim}{A}\,(\underset{\sim}{x}_u^k + \underset{\sim}{x}_l^k)\, x^k \leq \sum_{i=1}^{\xi} y_{iu}^{k^2} + \sum_{j=1}^{\eta} y_{jl}^{k^2} - \underset{\sim}{x}_u^k \underset{\sim}{x}_l^k - \underset{\sim}{\lambda}^k \qquad (4.20)$$

where ξ and η are mutually exclusive subsets of $\{1,2,...n\}$

$$\xi \cup \eta = 1,2...n \ \underline{\text{and}}\ x_{il} \cap x_{ju} = \emptyset$$

Alternatively in order to avoid this increase in the number of constraints and variables in the linear subproblems the quadratic terms may be approximated by mid-point tangents [Fig 4.4]

$$(x_l^k + x_u^k)\, x^k - (x_u^k + x_l^k)^2/4 \leq x^{k^2} \qquad (4.21)$$

In this latter case the linear underestimating constraints become

$$\underset{\sim}{A}\,(\underset{\sim}{x}_l^k + \underset{\sim}{x}_u^k)\, \underset{\sim}{x}^k - \underset{\sim}{A}\,(\underset{\sim}{y}_l^k + \underset{\sim}{y}_u^k)\, \underset{\sim}{y}^k$$

$$\leq (\underset{\sim}{x}_l^k + \underset{\sim}{x}_u^k)^2/4 - \underset{\sim}{y}_l^k\, \underset{\sim}{y}_u^k + \underset{\sim}{\lambda}^k \tag{4.22}$$

$$\underset{\sim}{A}\,(\underset{\sim}{y}_l^k + \underset{\sim}{y}_u^k)\, \underset{\sim}{y}^k - \underset{\sim}{A}\,(\underset{\sim}{x}_l^k + \underset{\sim}{x}_u^k)\, \underset{\sim}{x}^k$$

$$\leq (\underset{\sim}{y}_l^k + \underset{\sim}{y}_u^k)^2/4 - \underset{\sim}{x}_l^k\, \underset{\sim}{x}_u^k - \underset{\sim}{\lambda}^k \tag{4.23}$$

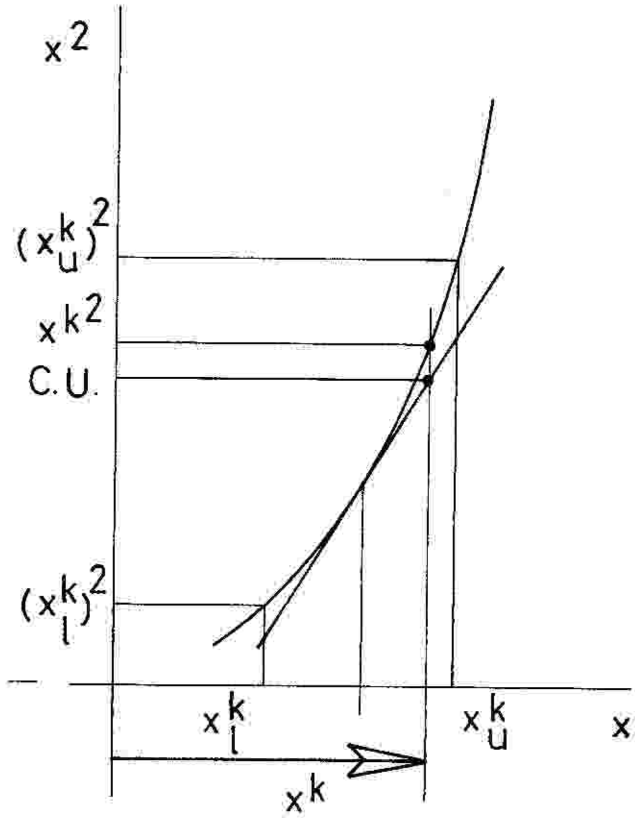


Fig 4.4 Mid-Point Tangent Approximation

4.3. Outline of the Algorithm

Let $\underset{\sim}{x}^p$ be the solution of the linear underestimating subproblem (P_p) with $k=1$

$$\text{Min } \underset{\sim}{c}^T \underset{\sim}{x} \tag{4.24}$$

$$\underline{\text{st}} \quad \underset{\sim}{A}\, \underset{\sim}{x} \geq \underset{\sim}{b} \tag{4.25}$$

$$\underset{\sim}{l}_p \leq \underset{\sim}{x} \leq \underset{\sim}{L}_p \tag{4.26}$$

If $\underset{\sim}{x}^p$ is not a feasible solution of the original problem we may try to strengthen the constraint or to restrict the domain of optimization of the subproblem (P_p) in order to make the solution x^p infeasible. Problem (P_p) is replaced by a set W^p of problems that bound the original problem in the sense that there exist one optimal solution x^* for at least one problem $j \in W^p$. Suppose an optimal solution to each problem $j \in W^p$ is obtained and let

$$x^s = \min_{j \in W^p} c^T x^j \qquad (4.27)$$

If x^s is not a feasible solution of the original problem we replace one of the problems of the bounding set W^p by a set of new problems.

Make p=p+1. We replace problem s by a set W^p such that $W^p = (W^{p-1} - \{s\}) \cup W^s$ contains an optimal solution of the original problem x^* feasible for at least one problem W^p. For each problem $j \in W^p$ either x^j is infeasible for j or $c^T x^j > c^T x^*$. This is a condition ensuring that some progress towards the final solution is made. The sequence of subproblems generated yield x^* at the limit so that a convergence condition limiting the maximum infeasibility to ε may have to be employed.

The combinatorial tree has each node identified with a subproblem j. The problems that replace j in the bounding set W^p are pointed to by the branches directed outward from that node. At any intermediary point in the calculations the set W^p of current bounding problems is identified with the set of nodes that are the leaves of the tree. We associate with each node of the tree an incumbent bound v. We say that any leaf node of the tree whose bound is

strictly less than v is active. Otherwise it is designated as terminated and need not be considered in any further computation. The Branch and Bound tree will be developped until every leaf can be terminated. Since all functions that define the domain of the original problem are continuous Soland's weak refining rule for splitting the bounds on the variables will be used:
Choose the index i of the variable that maximizes the difference between the quadratic form and the linear approximation made out of the constraints that are violated by $\underset{\sim}{x}^s$ in the original problem.

Now divide the corresponding interval $[l_i, L_i]$ into two new intervals $[l_i, x_i]$ and $[x_i, L_i]$. Therefore as soon as a node is selected to be branched the partition of its intervals is only dependent on its solution value and is not related to other partitions at the same level of the tree.

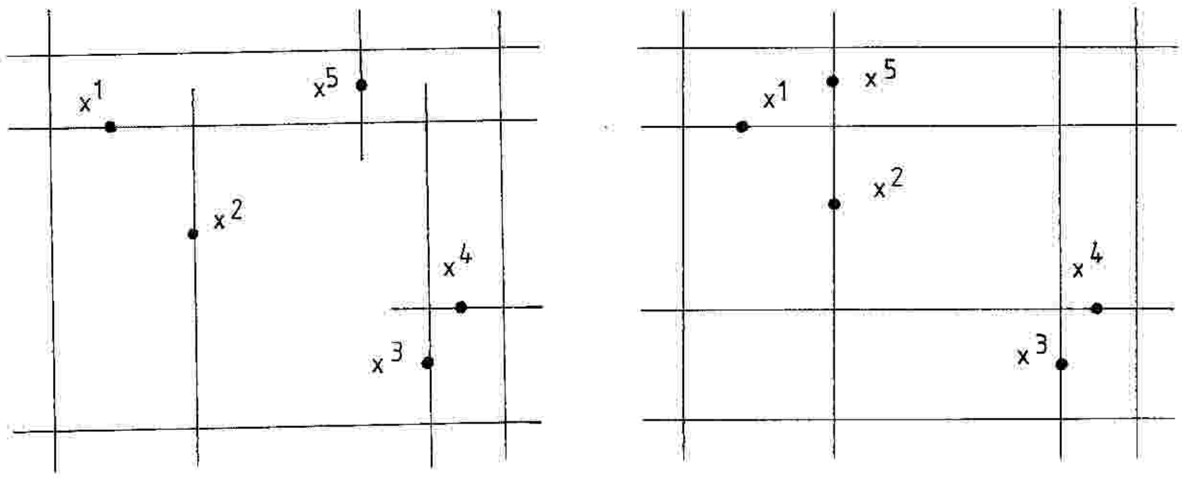


Fig 4.5 Partition and its completion

Thus the relaxed algorithm not requiring completion of the partition of the intervals at the same level of the tree will be used. It is based on a weaker form of Soland's convergence theorem [Fig.4.5] .

We remark that whenever the quadratic terms are approximated by mid-point tangents the underestimate fails to agree with the original function at its endpoints. Fig 4.6 shows how the underestimating function adopted may lead to to a solution at the boundary of the original bounds. If in this case the difference between a function and its approximation is a maximum along the axis corresponding to variable i a scheme to tighten up the underestimating functions by subdividing the whole box close to the solution point will fail to be interior so that another splitting rule (eg:a mid-point subdivision) has to be used.

These partitioning rules are not fixed and there is room for some heuristic alterations : the index of the variable that maximizes the difference from the original function of its approximation may be choosen not out of all violated constraints but taking the more violated constraint. In fact the least volume design is restricted by bilinear equalities and by doing so any solution of the subproblems should be brought up closer to feasibility.

4.4. Upper Bounds (Incumbent Solutions)

This is the value of the best feasible solution so far which is recorded and used as a thereshold value for accepting or rejecting any subsequent solutions or nodes of the tree. One of the generally available upper bounds at

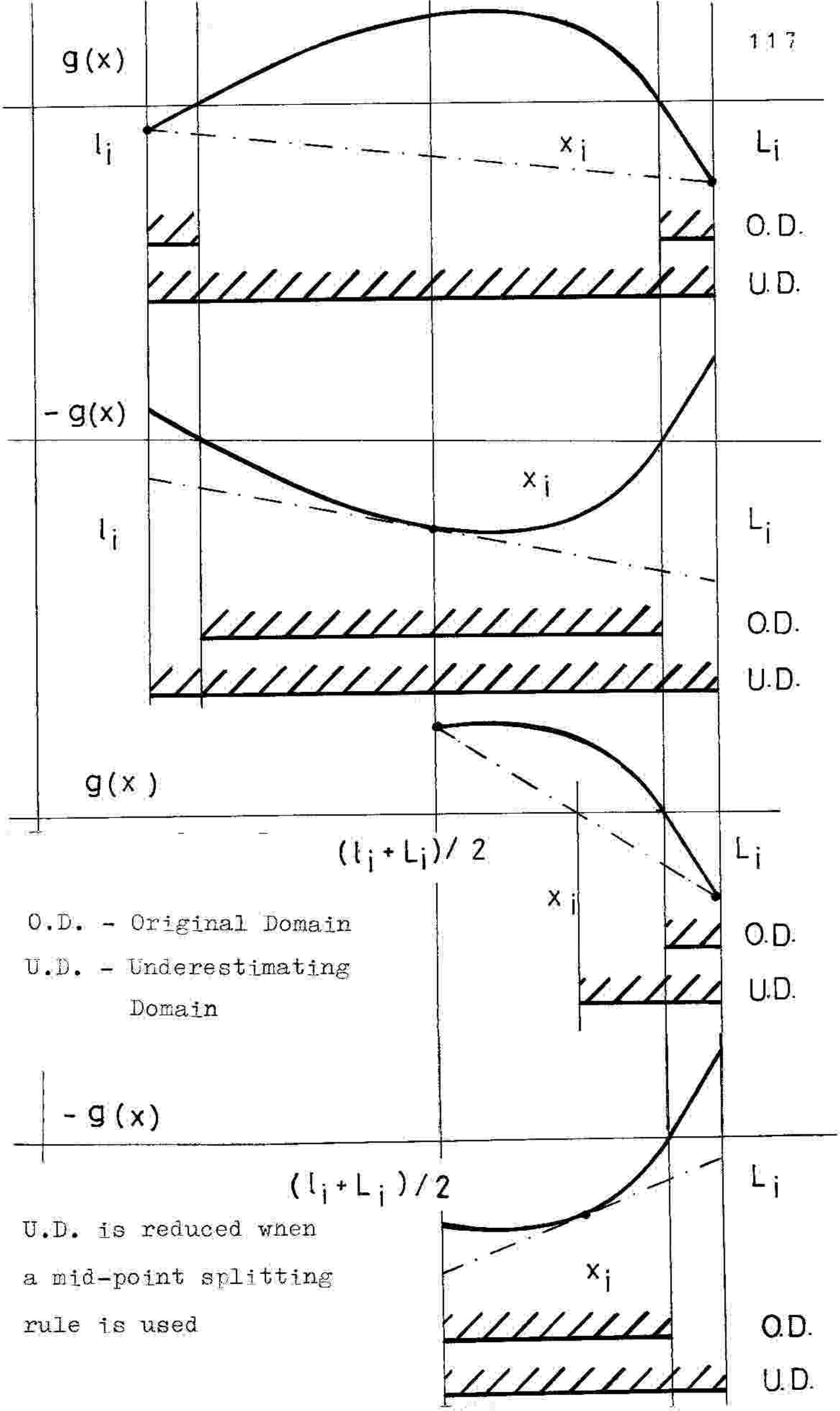


Fig 4.6 If the quadratic terms are approximated by mid-point tangents the maximum constraint violation may lead to the choice of a variable at a limiting value and the algorithm proposed by Soland would get stuck

the start of the procedure is the value given by the stress-ratio formula generally called fully stress design. The subproblems created by the algorithm will be solved by a LP routine which will ensure that the optimum will be approximated from below at all time. By fixing a set of areas a stress/displacement resultant is uniquely determined by matrix inversion. It is possible to scale the design in order to make the state variables feasible. If the scaled design variables are also feasible and their volume produces a better solution than the current incumbent they must be stored as new values for the solution and their volume becomes the new incumbent.

4.5. Enumeration of Suboptimal Solutions

Once the global solution has been identified this strategy can be reapplied to enumerate any other local solutions to the bilinearly constrained problem. By terminating the leaf corresponding to the optimal solution and resetting incumbent with an upper level. All (nearly) feasible solutions possessing a smaller volume than a given value should be enumerated. Since nothing prevents a local/global solution from existing at a boundary of the variable range as can be seen in Fig 4.7 the K-T conditions will be used to test local optimality in any other point than the global optima.

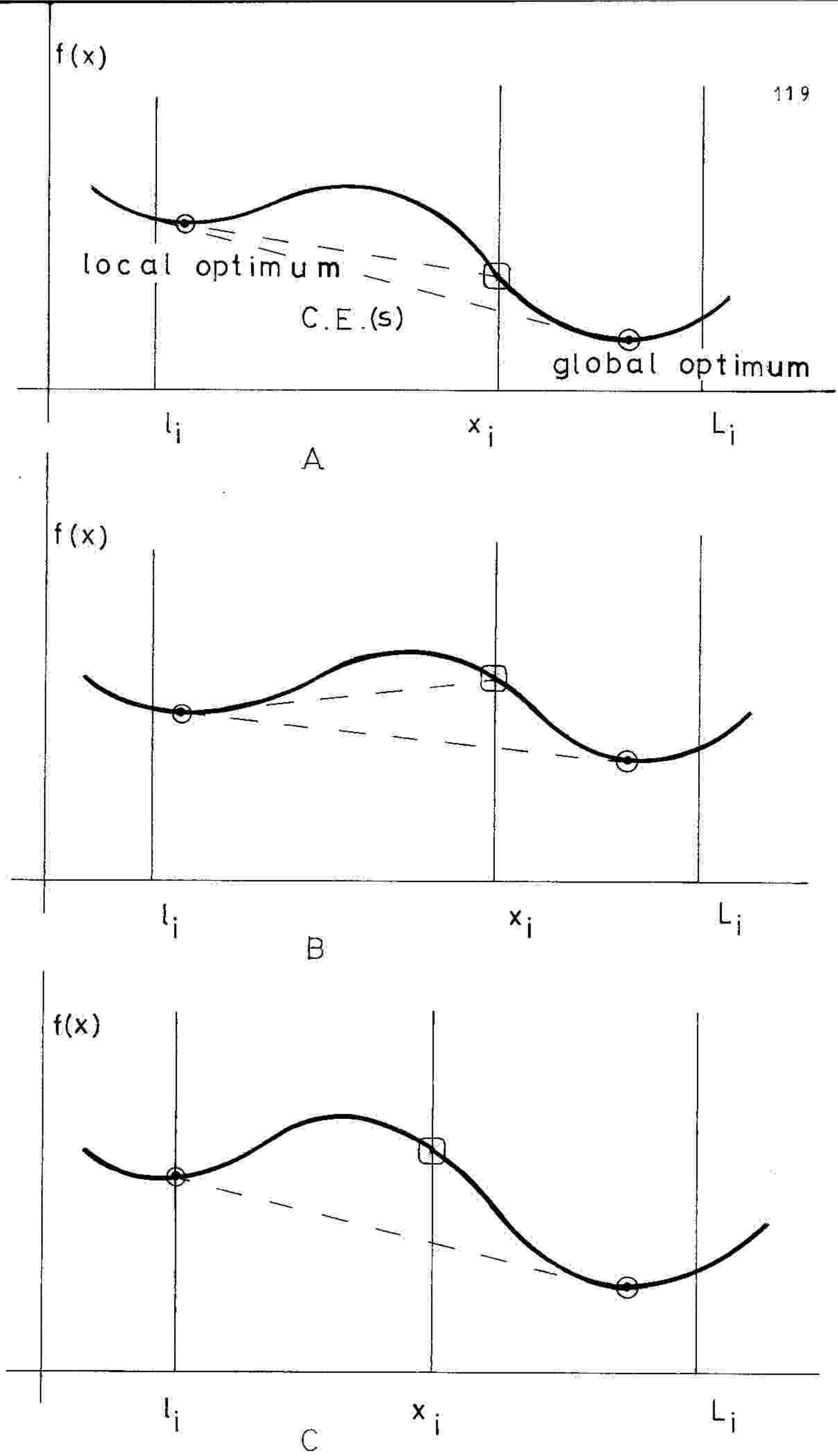


Fig. 4.7 Case A - Local Optimum is an interior point and has an higher value than the feasible point at the boudary of the varible range (what does not happen in Case B)
Case C - Local optima is a boundary point

4.6. Branch and Bound Algorithm - minimum objective function selection

The minimum volume bilinearly constrained subproblem (P)

$$\text{Min } \tilde{c}^T \tilde{a} \tag{4.28}$$

$$\text{st} \quad \tilde{a}^T \tilde{H}_i \tilde{s}^k = \tilde{\lambda}_i^k \qquad \text{Equilibrium} \tag{4.30}$$

$$\tilde{B}^T \tilde{L} \tilde{s}^k = \tilde{\emptyset} \qquad \text{Compatibility} \tag{4.31}$$

$$\tilde{a}_l \leq \tilde{a} \leq \tilde{a}_u \qquad \text{Area Bounds} \tag{4.32}$$

$$\tilde{d}_l^k \leq \tilde{D} \tilde{s}^k \leq \tilde{d}_u^k \qquad \text{Displacement Bounds} \tag{4.33}$$

$$\tilde{s}^k \leq \tilde{s}^k \leq \tilde{s}^k \qquad \text{Stress Bounds} \tag{4.34}$$

is transformed into the equivalent problem (P_T)

$$v = \text{Min } \tilde{c}^T \tilde{x} + \tilde{c}^T \tilde{y} \tag{4.35}$$

$$\text{st} \quad \tilde{A} \tilde{x}^{k^2} - \tilde{A} \tilde{y}^{k^2} = \tilde{\lambda}^k \qquad \text{Equilibrium} \tag{4.36}$$

$$\tilde{B}^T \tilde{L} \tilde{x}^k - \tilde{B}^T \tilde{L} \tilde{y}^k = \tilde{\emptyset} \qquad \text{Compatibility} \tag{4.37}$$

$$\tilde{d}_l^k \leq \tilde{D} \tilde{x}^k - \tilde{D} \tilde{y}^k \leq \tilde{d}_u^k \qquad \text{Displacement Bounds} \tag{4.38}$$

$$\tilde{a}_l \leq \tilde{x}^k + \tilde{y}^k \leq \tilde{a}_u \qquad \text{Area Bounds} \tag{4.39}$$

$$\tilde{x}^\eta + \tilde{y}^\eta = \tilde{x}^\xi + \tilde{y}^\xi \qquad \eta \cap \xi = \emptyset \quad ; \eta, \xi \in \{1, \ldots l\} \tag{4.40}$$

$$\tilde{s}_l^k \leq \tilde{x}^k - \tilde{y}^k \leq \tilde{s}_u^k \qquad \text{Stress Bounds} \tag{4.41}$$

$$\tilde{x}_l^k \leq \tilde{x}^k \leq \tilde{x}_u^k \qquad \text{Bounds on the}$$

$$\tilde{y}_l^k \leq \tilde{y}^k \leq \tilde{y}_u^k \qquad \text{new Variables} \tag{4.42}$$

The linear convex underestimating subproblem (P_u) is in all equal to (P_T) but now the equality bilinear constraints

(4.36) are replaced by linear underestimates

$$\underset{\sim}{Q}\,\underset{\sim}{x} + \underset{\sim}{T}\,\underset{\sim}{y} \leq \underset{\sim}{p} \qquad (4.43)$$

STEP 1

Let UB $=+\infty$ (if no information about any feasible solution is available) and LB $=-\infty$ be the initial upper and lower bounds on the optimal value of the objective function of the program (P_T) . Set a tolerance on the maximum infeasibility of the constraint underestimates ε .

set IT = Ø IB = 1 IN = Ø

STEP 2

Solve (P_u) . Make IT =IT + 1

STEP 3

(1) If (P_u) has no feasible solution and

a) if IT = 1 there is no feasible solution to the problem

b) Go To Step 7

(2) If (P_u) has the optimal value $\bar{v}$ > UB Go To Step 7.

(3) If any $|\underset{\sim}{A}\bar{\underset{\sim}{x}}^{k^2} - \underset{\sim}{A}\bar{\underset{\sim}{y}}^{k^2} - \underset{\sim}{\lambda}^{k}| > \varepsilon$ Go To Step 4

where $(\underset{\sim}{\bar{v}},\bar{\underset{\sim}{x}}^k,\bar{\underset{\sim}{y}}^k)$ is the optimal solution of (P_u) .

Otherwise $(\underset{\sim}{\bar{v}},\bar{\underset{\sim}{x}}^k,\bar{\underset{\sim}{y}}^k)$ is a feasible solution to (P_T) .

If $\bar{v}$ < UB make UB = $\bar{v}$ and store the solution point. The last values recorded will give the global solution to (P_T) when the algorithm ends. Go To Step 7.

STEP 4

Select the index j of the variables $\bar{\underset{\sim}{x}}^k$, $\bar{\underset{\sim}{y}}^k$ that maximizes the difference between each term of the violated constraint in the original problem and its convex underestimate. Increase the number of nodes (IB) by one : store the bounds corresponding to each pending node (VA and VB matrices) and the optimal solution of the corresponding subproblem (vector VE) and the index of the variable that has to be branched (vector ID) . IN = IN +1 ; ID (IN) = J ; V (IN) = V ; VE (IN) = VC = $\bar{\underset{\sim}{x}}^k$ or $\bar{\underset{\sim}{y}}^k$

VB (j,IN) = 1 ; VA (j,IN) = L for all j

STEP 5

(1) If IB = 1 set parameters for RHS branch

IND = ID (IN) ; VF = l (IND) ; l (IND) = VC ; IN = IN - 1 ;

IB = 2

Go To Step 2.

(2) if IB = 2 set parameters for LHS branch

L (IND) = l (IND) ; l (IND) = VF ; IB = 3

Go To Step 2

(3) If IB = 3 Go To Step 6

STEP 6

Choose the minimum value v (l) for l={ 1,...,IN} and set LB = v (l).

If LB > UB Terminate. All the nodes have lower bounds greater than the optimal feasible solution taken as incumbent. Reorder the sets ID,V,VE,VA,VB. Set IB = 1 and Go To Step 5

STEP 7

If IN = Ø and IB = 3 Terminate. The tree has been exhaustively explored . Go To Step 5

IP will record the number of problems that will be solved until the algorithm stops.

4.7. Computational Considerations

The total amount of computation is related to the number of distinct subproblems created and hence to the total number of nodes in the fully developed tree. The amount of temporary storage required is related to the maximum cardinality of the bounding set W^P(ie IN) and hence to the maximum number of leaves at any intermediate stage in its development. In order to try to accelerate convergence several heuristic modifications were tried out. A penality term taking into account the infeasibility of each subproblem by multiplying the value of each constraint violation by the corresponding dual variable available from the solution of the linear subproblem is added to the objective function value $\bar{V}$. This modified solution value will be used in the selection of the next node to be branched. Unfortunatelly it cannot be used as a justification to kill pending nodes since it would correspond to a dual feasible (primal infeasible) procedure. Several strategies were tried out in order to avoid the increase in the amount of computation at the expense of additional storage. A simple side branching followed by backtracking as soon as a terminated leaf was found depends heavily on the availability of a good incumbent at the early stages of the algorithm.

A more involved deep branching srategy has also been run. It requires only the storage of the level of the tree, the index of the variable to be branched and its solution value. At the start (level Ø) we solve a linear underestimating subproblem where all variables are allowed to vary whithin their bounds. If the solution is not feasible to the original problem the interval corresponding to the variable that maximizes infeasibity is split into two parts. Two new subproblems are then generated. If we choose the solution giving the lower value and increase the level of the tree by one we see whether the improvement in the solution value has made it feasible to the original problem. If not another interval will be selected for branching. The algorithm proceeds until a tail is reached. There are three types of tail :(a) feasible solution to the original problem (b) LP solution with OF value at least as high as the best feasible solution found so far (c) infeasible LP.

When a tail is reached the intervals not yet considered corresponding to pending nodes must be investigated. Starting with the level at which the the tail was found the bounds on the remaining branch will be looked at. Any which is no better than the incumbent solution or are infeasible can be deleted and the level of the tree is reduced. But if a level is reached for which the remaining branch still offers the possibility of a better feasible solution to the original problem whithin its set this becomes the main branch and the process of branching is restarted the level will increase by one until a tail is reached. Eventually the backing up investigation will get all the way to level zero indicating that no subset has lower bound than the best feasible solution so far. The incumbent will therefore give the global solution of the nonconvex problem.

4.8. Branch and Bound Algorithm - deep mainbranching strategy

STEP 1

Let UB $=+\infty$ (if no information is available about any feasible solution of (P)) and LB $=-\infty$ be the initial upper and lower bounds on the optimal value of the OF of the program (P_T) . Set a tolerance on the maximum infeasibility of the nonlinear constraints of ε .

Set IT = Ø ; IB = 1 ; LV = Ø . Go To Step 2.

STEP 2

Make IT = IT + 1 . Solve (P_u).

STEP 3

(1) If (P_u) has no feasible solution and

a) IT = 1 there is no feasible solution to the original problem.

b) Go To Step 7

(2) If the optimal value of (P_u) v > UB Go To Step 7

(3) If any $|\underset{\sim}{A}\,\underset{\sim}{\bar{x}}^{k^2} - \underset{\sim}{A}\,\underset{\sim}{\bar{y}}^{k^2} - \underset{\sim}{\lambda}^{k}| > \varepsilon$ Go To Step 4

$(\underset{\sim}{\bar{v}},\underset{\sim}{\bar{x}}^k,\underset{\sim}{\bar{y}}^k)$ is a feasible solution to (P_T)

If $\bar{v}$ < UB make UB = $\bar{v}$ and store $(\underset{\sim}{\bar{v}},\underset{\sim}{\bar{x}}^k,\underset{\sim}{\bar{y}}^k)$. The values recorded in UB when the algorithm stops will give the global solution of the problem (P) . Go To Step 7.

STEP 4

Select the index j of the variable $\underset{\sim}{\bar{x}}^k$ or $\underset{\sim}{\bar{y}}^k$ that maximizes the error derived from taking convex underestimates out of the violated constraints in (P_T).

STEP 5

(1) If IB = 1 Set the bounds for the 1st branch of the node

to be split. IB = 2 . Go To Step 2 .

(2) If IB = 2 Keep a record of the solution corresponding to the 1st branch. Set the bounds for the 2nd branch emanating from the node that is being split. IB = 3 . Go To Step 2

(3) If IB = 3 Keep a record of the solution corresponding to the 2nd branch. Go To Step 6.

STEP 6

Increase the level of the tree (LV) by one.

(1) If any of the branches is not a tail choose the node giving the lower bound. LB = $\bar{v}$; IB = 1 . Go To Step 5.

(2) Backtrack. Reduce the level of the tree by one.

If LV = Ø Terminate. All the tree has been explored.

If v corresponding to the new main branch < UB set LB = $\bar{v}$; IB = 1 . Go To Step 5.

Otherwise backtrack. Restart Step 6 (2)

STEP 7

A tail is reached . Set $\bar{v} = +\infty$ corresponding to either

a) Infeasible subproblem.

b) best feasible solution so far.

c) a solution of a subproblem with function value at least as high as the best solution found so far.

Go To Step 6 (2)

4.9. Parallel Hyperplane Approximations

The bounds on the variables of the transformed problem are based on the bounds on the state and design variables and effectively increase the domain that has to be split and thoroughly explored by the B & B strategy. The consideration of a number of loading cases greater than one is penalized with an increase in the number of constraints and variables in the problem where Soland's algorithm will be applied. The bounds on both state and design variables will not disappear in the formulation of (P_T) and they become range constraints in this problem heavily increasing the computational effort needed to solve each subproblem. After the linearization of (P_T) an inverse rotation was performed in order to obtain a problem (P_E) in terms of the initial state and design variables. The linear expressions that are convex underestimates of each bilinear term when the convex envelope is substituted by the maximum of the two endpoints tangents in terms of $\underset{\sim}{x}^k$ and $\underset{\sim}{y}^k$ are

$$[(\underset{\sim}{a}_l + 3\,\underset{\sim}{s}_l^k + \underset{\sim}{s}_u^k - \underset{\sim}{a}_u)\,\underset{\sim}{a} + (\underset{\sim}{s}_l^k + 3\,\underset{\sim}{a}_l - \underset{\sim}{s}_u^k + \underset{\sim}{a}_u)\,\underset{\sim}{s}^k]/4 + [\underset{\sim}{a}_l\,\underset{\sim}{a}_u - \underset{\sim}{s}_l^k\,\underset{\sim}{a}_u + \underset{\sim}{s}_u^k\,\underset{\sim}{s}_u^k - \underset{\sim}{a}_l^2 - \underset{\sim}{s}_l^{k^2} - 3\,\underset{\sim}{a}_l\,\underset{\sim}{s}_l^k]/4 \le \underset{\sim}{a}\;\underset{\sim}{s}^k \tag{4.44}$$

$$[(\underset{\sim}{a}_u + 3\,\underset{\sim}{s}_u^k + \underset{\sim}{s}_l^k - \underset{\sim}{a}_l)\,\underset{\sim}{a} + (\underset{\sim}{s}_u^k + 3\,\underset{\sim}{a}_u - \underset{\sim}{s}_l^k + \underset{\sim}{a}_l)\,\underset{\sim}{s}^k]/4 + [\underset{\sim}{a}_l\,\underset{\sim}{a}_u - \underset{\sim}{s}_l^k\,\underset{\sim}{a}_l + \underset{\sim}{s}_u^k\,\underset{\sim}{s}_l^k - \underset{\sim}{a}_u^2 - \underset{\sim}{s}_u^{k^2} - 3\,\underset{\sim}{a}_u\,\underset{\sim}{s}_u^k]/4 \le \underset{\sim}{a}\;\underset{\sim}{s}^k \tag{4.45}$$

$$[(\underset{\sim}{a}_l - 3\,\underset{\sim}{s}_u^k - \underset{\sim}{s}_l^k - \underset{\sim}{a}_u)\,\underset{\sim}{a} + (\underset{\sim}{s}_u^k - 3\,\underset{\sim}{a}_l - \underset{\sim}{s}_l^k - \underset{\sim}{a}_u)\,\underset{\sim}{s}^k]/4 + [\underset{\sim}{a}_l\,\underset{\sim}{a}_u + \underset{\sim}{s}_l^k\,\underset{\sim}{a}_u + \underset{\sim}{s}_l^k\,\underset{\sim}{s}_u^k - \underset{\sim}{a}_l^2 - \underset{\sim}{s}_u^{k^2} + 3\,\underset{\sim}{a}_l\,\underset{\sim}{s}_u^k]/4 \le -\underset{\sim}{a}\;\underset{\sim}{s}^k \tag{4.46}$$

$$[(\underset{\sim}{a}_u - 3\,\underset{\sim}{s}_l^k - \underset{\sim}{a}_l - \underset{\sim}{s}_u^k)\,\underset{\sim}{a} + (\underset{\sim}{s}_l^k - 3\,\underset{\sim}{a}_u - \underset{\sim}{a}_l - \underset{\sim}{s}_u^k)\,\underset{\sim}{s}^k]/4 + [\underset{\sim}{a}_l\,\underset{\sim}{a}_u + \underset{\sim}{a}_l\,\underset{\sim}{s}_u^k + \underset{\sim}{s}_l^k\,\underset{\sim}{s}_u^k - \underset{\sim}{a}_u^2 - \underset{\sim}{s}_l^k$$

$$+ 3\, a_u\, s_l^k]/4 \leq -a\, s^k \tag{4.47}$$

If the mid point tangent are used as convex approximation in x^k and y^k the corresponding linear expressions in terms of a and s^k have the form

$$[(s_l^k + s_u^k)\, a + (a_l + a_u)\, s^k]/2 - [(a_u - a_l)^2 + (s_u^k - s_l^k)^2]/16 - [3\, a_l\, s_l^k + a_l\, s_u^k + a_u\, s_l^k + 3\, a_u\, s_u^k]/8 \leq a\, s^k \tag{4.48}$$

$$[(s_l^k + s_u^k)\, a + (a_l + a_u)\, s^k]/2 - [(a_u - a_l)^2 + (s_u^k - s_l^k)^2]/16 + [3\, a_l\, s_l^k + a_l\, s_u^k + a_u\, s_l^k + 3\, a_u\, s_u^k]/8 \leq -a\, s^k \tag{4.49}$$

In this approximation each bilinear term defining a hyperbolic line is substituted by the surface limited by the variables range and by two parallel lines separated by a width given by the sum of the squares of the range of these variables divided by a constant [Fig 4.8] .

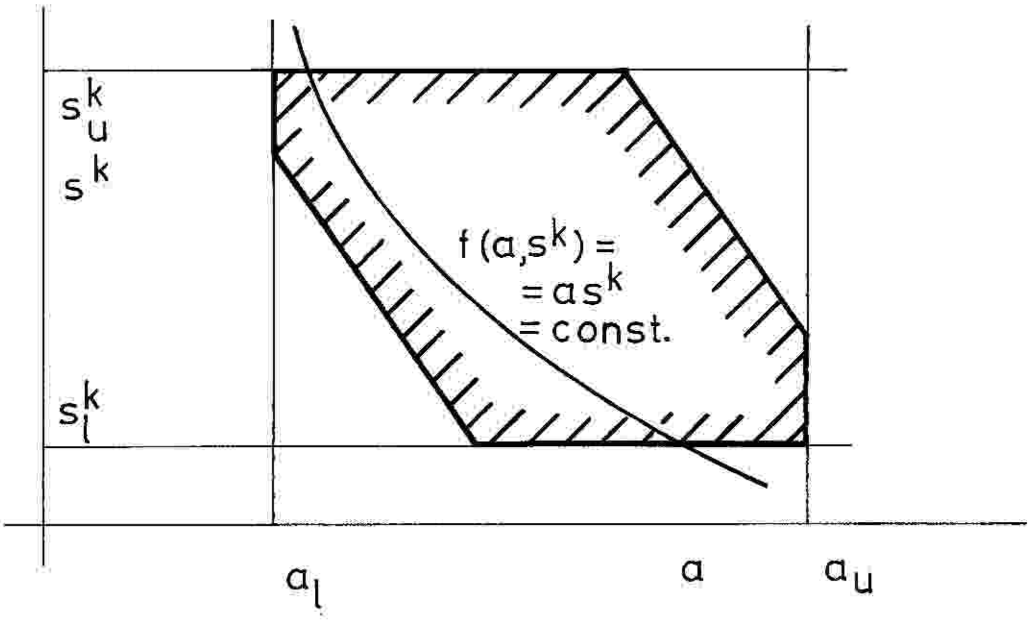


Fig 4.8 Tangent Hyperplane Approximation

4.10. Factorable Programming

Bilinear expressions of the form $p x + q y + k x y$ are particular cases of the more general class of problems termed Factorable Programming. Since the convex underestimate of the linear part of the expression is trivial (it is the linear function itself) we have to determine the envelope of the product term. Let the function $f(x,y) = xy$ defined in the rectangle of bounds

$$a \leq x \leq b \qquad c \leq y \leq d \tag{4.50}$$

We first note that the function values at the corners must coincide with the underestimate taken. Since three points are sufficient to define a plane in 3-D the convex underestimate will be taken as the z coordinate on the highest of the two planes defining a ridge through the two intermediate function values [Fig 4.9] .

$$z_1 = c\,(x - a) + a\,(y - c) + a\,c$$

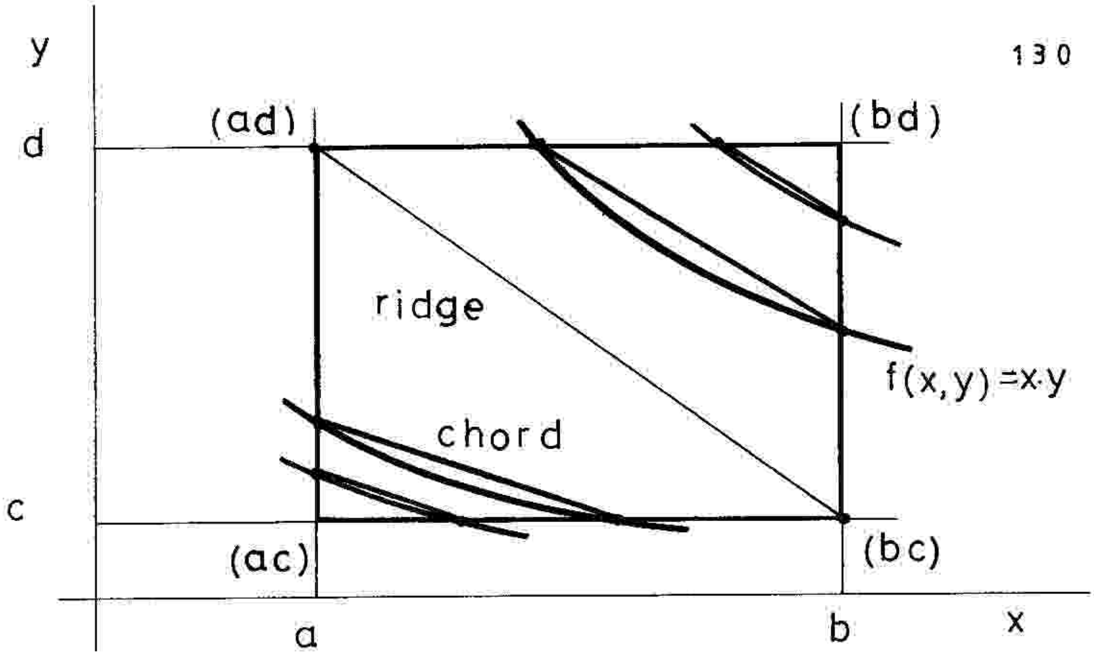
$$= c\,x + a\,y - a\,c \tag{4.51}$$

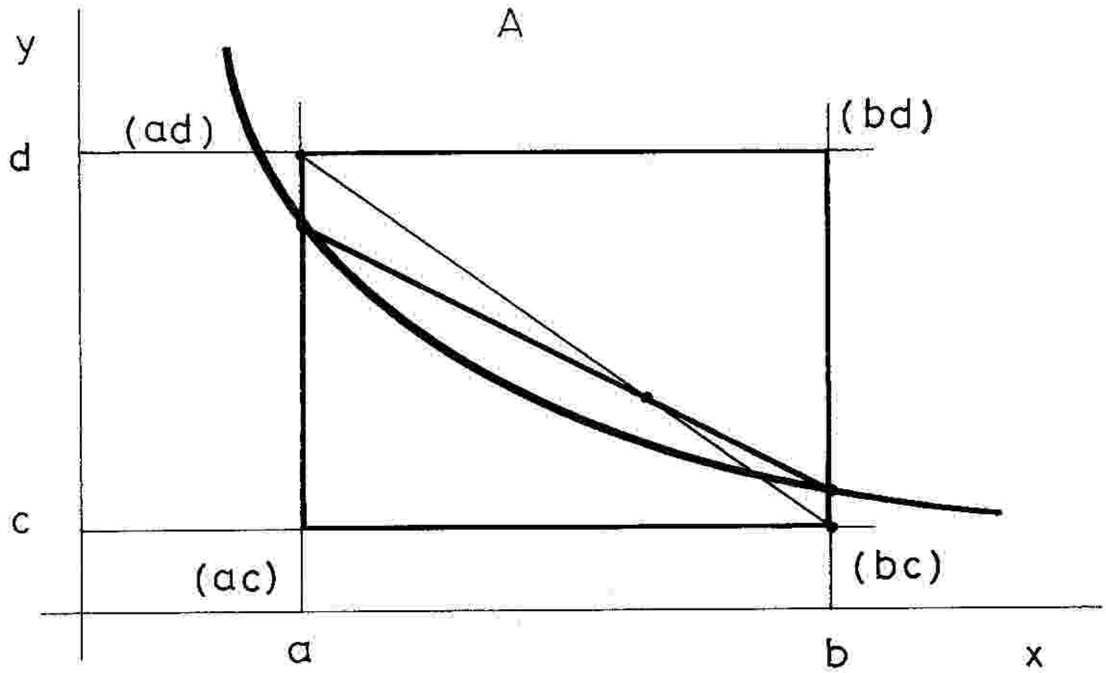
$$z_2 = d\,(x - b) + b\,(y - d) + b\,d$$

$$= d\,x + b\,y - b\,d \tag{4.52}$$

$$z = \mathrm{Max}\,\{ z_1, z_2 \} \leq f(x,y) = x\,y \tag{4.53}$$

These convex underestimate functions give constant slope when either z_1 or z_2 are set constant.





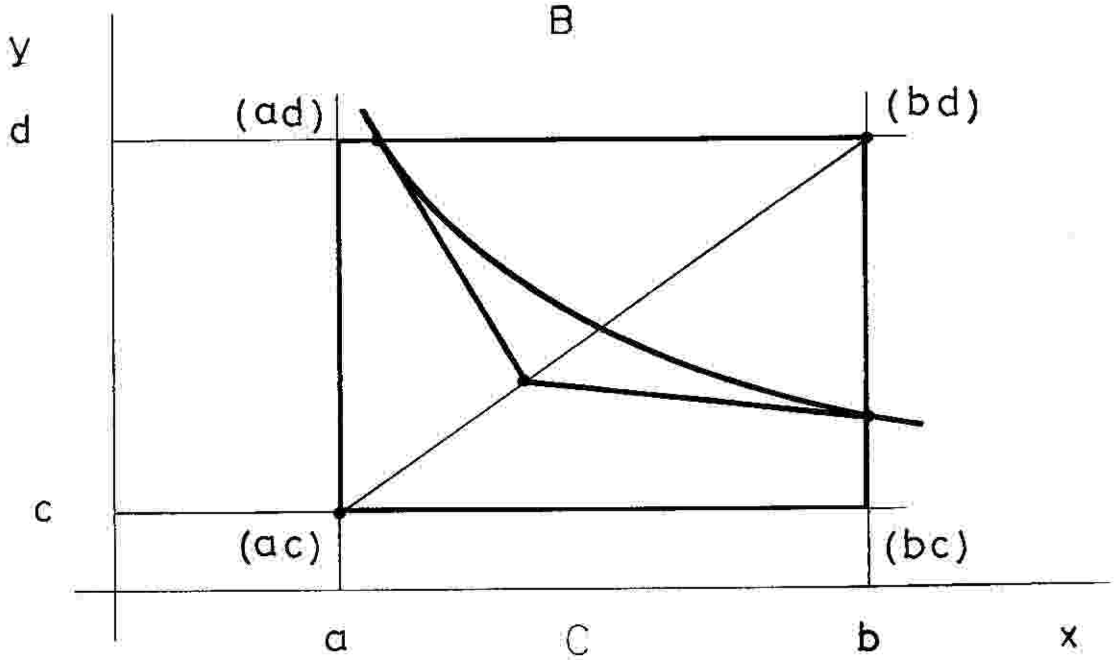


Fig. 4.9 Cases A and B - Convex Envelopes of the product term x.y
Case C - Concave Overestimate of the same product function

It is also possible to show that the chords of $f(x,y)$ match the level curves of the planes defined since they define parallel lines that intersect along the diagonal ridge. The concave overestimating functions are accordingly defined as the z coordinate at the lowest of the two planes defining a ridge through the corners having the higher and lower function values

$$z_1^o = d\,(x - a) + a\,(y - d) + a\,d$$

$$= d\,x + a\,y - a\,d \qquad (4.54)$$

$$z_2^o = c\,(x - b) + b\,(y - c) + b\,c$$

$$= c\,x + b\,y - b\,c \qquad (4.55)$$

$$z^o = \text{Min}\,\{ z_1^o, z_2^o \} \geq x\,y \qquad (4.56)$$

When these underestimating functions are switched around (ie when they belong to another quadrant) one may check how these approximations are modified. Supposing $f(x,y)=-xy$ is the function whose convex underestimate is required defined over the rectangle of bounds

$$a \leq x \leq b \qquad -d \leq -y \leq -c \qquad (4.57)$$

The two planes determined are

$$z_1^- = -b\,y - c\,x + b\,c \qquad (4.58)$$

$$z_2^- = -d\,x - a\,y + a\,d \qquad (4.59)$$

Let $z^- = \min\{ z_1^-, z_2^- \}$ be the envelope of $-xy$.

Therefore

$$z^- \leq -x\ y \qquad (4.60)$$

or

$$-z^- \geq x\ y \qquad (4.61)$$

But

$$z^- = \text{Max}\ \{ z_1^-, z_2^- \} \qquad (4.62)$$

$$-z^- = \text{Min}\ \{ -z_1^-, -z_2^- \}$$

$$= \text{Min}\ \{ z_1^o, z_2^o \} = z^o \qquad (4.63)$$

The convex underestimate of -f(x,y) is the symmetric of the concave overestimate of f (x,y).

We remark that the underestimating function z (and overestimating z^o) is not differentiable everywhere. There are several ways of handling this by altering the problem in order to create an equivalent linear programming problem. The simplest way involves the addition of some extra inequality constraints and variables . Considering the problem

$$\text{Min}\ \underset{\sim}{c}^T \underset{\sim}{x} \qquad (4.64)$$

$$\underline{\text{st}}\quad \underset{\sim}{A}\ \underset{\sim}{x} \geq \underset{\sim}{b} \qquad (4.65)$$

$$\underset{\sim}{f}^T \underset{\sim}{x} + \max\ \{ \underset{\sim}{g}_1^T \underset{\sim}{x}, \underset{\sim}{g}_2^T \underset{\sim}{x} \} \leq \underset{\sim}{h} \qquad (4.66)$$

$$\underset{\sim}{p}^T \underset{\sim}{x} + \min\ \{ \underset{\sim}{q}_1^T \underset{\sim}{x}, \underset{\sim}{q}_2^T \underset{\sim}{x} \} \geq \underset{\sim}{r} \qquad (4.67)$$

That is equivalent to

$$\text{Min} \quad c^T x \tag{4.68}$$

$$\text{st} \quad A x \geq b \tag{4.69}$$

$$f^T x + g_1^T x \leq h \qquad f^T x + g_2^T x \leq h \tag{4.70}$$

$$p^T x + q_1^T x \geq r \qquad p^T x + q_2^T x \geq r \tag{4.71}$$

and can also be written

$$\text{Min} \quad c^T x \tag{4.72}$$

$$\text{st} \quad A x \geq b \tag{4.73}$$

$$f^T x + u \leq h \qquad p^T x + v \geq r \tag{4.74}$$

$$u \geq g_1^T x \qquad u \geq g_2^T x \tag{4.75}$$

$$v \leq q_1^T x \qquad v \leq q_2^T x \tag{4.76}$$

Both formulations may be employed. The introduction of more variables is recommendable when the same nondifferentiable term will appear in a number of constraints. The globally convergent algorithms described in 4.6. and 4.8. can also be used in conjunction with the convex envelope described in this section. The solution of the bilinearly constrained structural synthesis problem is also obtained by solving a sequence of linear underestimating subproblems. The efficiency of the method depends upon the tightness of the underestimating problems. This is helpfull in eliminating regions which cannot contain the global minimizer. The convergence to the global solution is ensured by the fact that the rules for splitting the regions when a node is branched leads to a nondecreasing monotone sequence of lower bounds on the final solution. In fact the feasible constrained set for a linear subproblem is larger than the feasible region of the original subproblem.

If the optimal solution of the subproblem is feasible in the original domain and has an OF equal to the global solution it indicates that it is global minimizer. If not the enlarged domain will be divided by choosing the product term out of the violated constraints that maximizes the difference with respect to its envelope and then selecting the variable with larger range. The interval corresponding to the selected variable will be split and the optimal variables of the subproblem will define the new interval bounds. This partition never fails to be interior when using convex envelopes since if it would be at a boundary no difference would occur. These partitions ensure that the new feasible regions defining new convex envelopes will be brought up closer to the value of the optimal solution variable.

4.11. Applications

4.11.1. Determination of the global solution [Appendix A]

The small scale problem in 6 variables described in 2.3 will be used to illustrate the strategy described in this Chapter. When factorable functions are replaced by their convex underestimates the problem becomes

$$\text{Min} \quad x_1 + x_2 + x_3$$

$$\underline{\text{st}} \quad x_1 x_4 \qquad\qquad x_3 x_6 = 0$$

$$3\, x_1 x_4 + 1.2\, x_2 x_5 - x_3 x_6 = 10$$

$$5\, x_1 + x_2 + x_3 \leq 2.5$$

$$(x_l)^T = [.1 \quad .1 \quad .1 \quad 0. \quad 0. \quad -2.5]$$

$$(x_u)^T = [5. \quad 5. \quad 5. \quad 2.5 \quad 2.5 \quad 0.]$$

The convex envelope of $x_1\, x_4$ in the rectangle of bounds

$$x_{1l} \leq x_1 \leq x_{1u} \qquad x_{4l} \leq x_4 \leq x_{4u}$$

is the maximum of two affine quantities and is defined by introducing the new variable $x_7 \leq x_1 x_4$

$$x_7 \leq x_{4u} x_1 + x_{1u} x_4 - x_{1u} x_{4u}$$

$$x_7 \leq x_{4l} x_1 + x_{1l} x_4 - x_{1l} x_{4l}$$

Similarly the convex envelope corresponding to $-x_1 x_4$ is represented by $x_8 \leq -x_1 x_4$

$$x_8 \leq -x_{4l} x_1 - x_{1u} x_4 + x_{4l} x_{1u}$$

$$x_8 \leq -x_{4u} x_1 - x_{1l} x_4 + x_{4u} x_{1l}$$

The underestimate of $x_3 x_6$ and $-x_3 x_6$ are respectively

$$x_9 \leq x_{6u} x_3 + x_{3u} x_6 - x_{6u} x_{3u}$$

$$x_9 \leq x_{6l} x_3 + x_{3l} x_6 - x_{6l} x_{3l}$$

and

$$x_{10} \leq -x_{6u}x_3 - x_{3l}x_6 + x_{6u}x_{3l}$$

$$x_{10} \leq -x_{6l}x_3 - x_{3u}x_6 + x_{6l}x_{3u}$$

The first bilinear constraint is an equality and will be changed in the standard fashion by a pair of inequalities

$$x_1 x_4 + x_3 x_6 \leq 0$$

$$x_1 x_4 + x_3 x_6 \geq 0 \quad \Rightarrow \quad -x_1 x_4 - x_3 x_6 \leq 0$$

to which correspond the inequalities

$$x_7 + x_9 \leq 0$$

$$x_8 + x_{10} \leq 0$$

The constraint

$$3\, x_1 x_4 + 1.2\, x_2 x_5 - x_3 x_6 \leq 10$$

is replaced by a pair of linear expressions

$$3\, x_7 + 1.2\, (x_{5u}x_2 + x_{2u}x_5 - x_{5u}x_{2u}) + x_{10} \leq 10$$

$$3\, x_7 + 1.2\, (x_{5l}x_2 + x_{2l}x_5 - x_{5l}x_{5u}) + x_{10} \leq 10$$

The same procedure is adopted for

$$-3\, x_1 x_4 - 1.2\, x_2 x_5 + x_3 x_6 \leq -10$$

$$3\, x_8 + 1.2\, (-\, x_{5l}x_2 - x_{2u}x_5 + x_{2u}x_{5l}) + x_9 \leq -10$$

$$3\, x_8 + 1.2\, (-\, x_{5u}x_2 - x_{2l}x_5 + x_{2l}x_{5u}) + x_9 \leq -10$$

The remaining expressions of the problem are linear and therefore the determination of their envelopes is trivial . Thus this simple example is reformulated as a linear underestimating problem in 10 variables and 15 inequality constraints. For each set of bounds on the variables x_1 to x_6 a node is defined. We may remark that the structural matrix corresponding to each subproblem has to be up-dated whenever a new node is created. The results of the initial iterations of the algorithm known as breadth first (choose the node with lower bound) are reported next.

INITIALIZATION

STEP1

Let $UB = +\infty; LB = -\infty$

IT = 0   IB =1   IN = 0 $\varepsilon = 0.01$

Set x_l and x_u at the values defining the hyper-rectangle of bounds of the original bilinear problem

STEP 2

IT = 1

Solve the subproblem corresponding to the 1st. node :

Optimal solution ...1.545

$x_1 = 1.34; x_2 = .1; x_3 = .1; x_4 = .64; x_5 = 1.82; x_6 = -2.5$

STEP 3   (3)

Max infeasibility ... 2.32 > .01

STEP 4

Index of the worse approximation ... 1

IN = 1 ID(1) = 1 V(1) = 1.34

VB(J,1) = X_{J1} VA(J,1) = X_{Ju} VE(1) = 1.34 VC = 1.34

J=1,...6

STEP 5

VF = VB(1,1) X_{11} = VE(1) IND = 1 IN = Ø IB = 2

STEP 2 IT = 2

Optimal solution ... 2.502

$x_1 = 1.96; x_2 = .1; x_3 = .54; x_4 = 1.; x_5 = 0.; x_6 = -2.5$

STEP 3 (3)

Maximum infeasibility ... 1.98 > .Ø1

STEP 4

Index of the worse approximation ... 1

IN = 1 ID(1) = 1 U(1) = 2.599

VB(J,1) = X_{J1} VA(J,1) = X_{Ju} VC(1) = 1.96

STEP 5 (2)

$X_{1u} = X_{11}; X_{11} =$ VF IB = 3

STEP 2 IT = 3

Optimal solution ... 2.846

$x_1 = .49; x_2 = 2.26; x_3 = .1; x_4 = 1.; x_5 = 0.; x_6 = -2.5$

STEP 3 (3)

Maximum infeasibility ... 3.01 > .01

STEP 4

Index of the worse approximation ... 2

IN = 2   ID(2) = 2   V(2) = 2.85

VE(2) = 2.26   VB(J,2) = X_{Jl} VA(J,2) = X_{Ju}

STEP 5 (3)

V(1) = MIN V(1), V(2)   IN = 1   IND = ID(1)

LB = V(IV)   VC = VE(IV) X_{Jl} = VB(J,IV) X_{Ju} = VA(J,IV)

Reorder the remaining data corresponding to pending nodes

ID(1) = ID(2)   VA(J,1) = VA(J,2)   VB (J,1) = VB(J,2)

VE(1) = VE(2)   IB = 1

STEP 5 (1)

VF = $X_{IND1} = X_{11}; X_{IND1} = VC$

IN = 1   IB = 2

STEP 2   IT = 4

Optimal solution ...3.13

$x_1 = 2.25; x_2 = .1; x_3 = .78; x_4 = 1.; x_5 = .0; x_6 = -2.5$

STEP 3 (3)

Maximum infeasibility ... .43 > .01

STEP 4

Index of the worse approximation ... 1

IN = 2 ID(2) = 1 V(2) = 3.13 VE(2) = 2.25

VC = 2.25 VB(J,2) = X_{J1} VA(J,2) = X_{Ju}

STEP 5 (2)

$X_{1u} = X_{11}$ $X_{11} = VF$

and the algorithm proceeds until the combinatorial tree is completely explored. The global solution is found to be

$x_1 = .1$ $x_2 = 3.33$ $x_3 = .1$ $x_4 = 0.$ $x_5 = 2.5$ $x_6 = 0.$

In Fig 4.10 the combinatorial tree concerning this problem is represented.

4.11.2. Multiple optimal solutions

Multiple solutions can be located by using the following strategy. The combinatorial tree used in the last problem is restarted by fathoming the branch corresponding to the local optimum and keeping a record of all feasible nodes until a certain level (3.60) and it is drawn in Fig 4.11 .

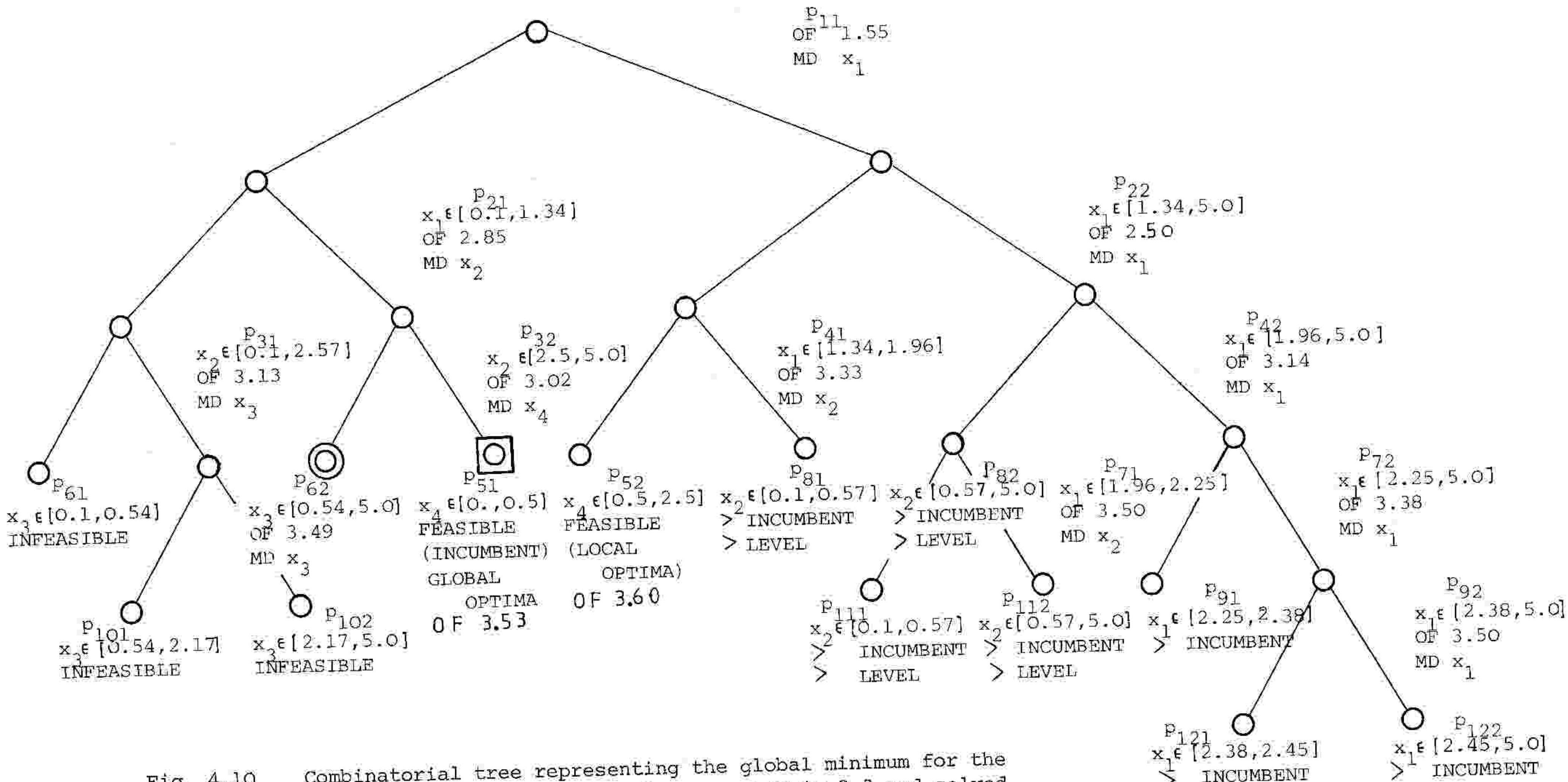


Fig. 4.10 Combinatorial tree representing the global minimum for the bilinearly constrained example presented in 2.3 and solved by using the envelopes of the factorable terms.

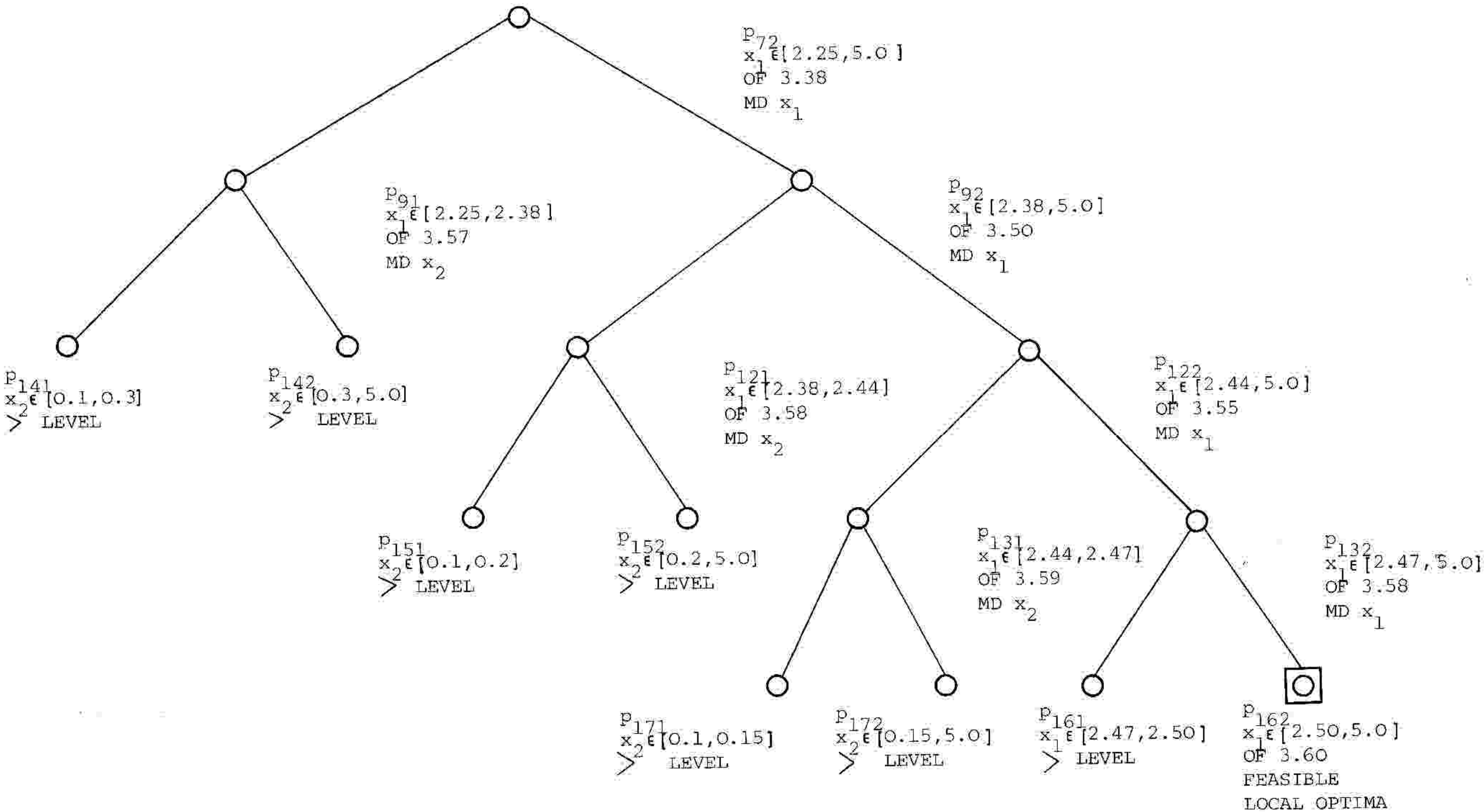


Fig. 4.11 Continuation of the combinatorial tree enumerating the several local solutions of example 2.3

For all such points the K-T necessary conditions will be used to discard all solutions that are not extrema.

4.12. Computational Experience

A three bar truss subjected to two alternative loading conditions referred in 3.6.1 was solved in order to compare relative efficiency of the several underestimating functions and algorithms. We start by reporting the results corresponding to a search strategy known as breadth first (always choose the node with lower bound) . The termination criteria was determined by limiting maximum infeasibility of the constraint functions.

W1 Using a NAG LP routine the underestimating function as given by the separable programming techniques and the strictly convex terms are approximated by mid-point tangents. The branching rule will consists of dividing the interval corresponding to the variable selected into half. Each subproblem has 12 variables and is subject to 31 linear inequality constraints.

CPU 593 ieu (instruction execution units)

W2 As W1 but using Land and Powell (1973) subroutine to solve each suproblem . It is assumed that no knowledge of any basis for each LP is available.

CPU 275 ieu 137 stored nodes 925 problems solved

W3 As W2 but restarting each subproblem with a basis not necessarly feasible.

CPU 230 ieu 137 stored nodes 925 problems solved

W4 As W3 but no accuracy checks are made for the LP solutions

CPU 206 ieu 137 stored nodes 912 problems solved

W5 As W4 but the branching rule requiring that the index of the variable to be partitioned has its choice governed by the terms belonging to the more violated constraint.

CPU 156 ieu 131 stored nodes 728 problems solved

W6 As W5 but uses a feasibility test applied to the constraints whose coefficients are modified by the branching operations.

CPU 143 ieu 129 stored nodes 697 problems solved

W7 As W6 but including a pseudo-cost in the objective function that attempts to give measure of the infeasibility of the approximation. This penalty term is the sum of the products of the Lagrange multipliers of the subproblem corresponding to the node to be branched times the infeasibility of the violated constraints.

CPU 133 ieu 42 stored nodes 740 problems solved

(33 ieu and 170 subproblems until the first feasible solution available)

Using an alternative search strategy such as right-hand depth first (pick up the right hand sucessor of the current node otherwise backtrack to the predecessor of the current node and reapply the rule) tends to require less storage space. The number of levels of the tree (partitions of the intervals) until the termination criterion is met is 35. In all other aspects the following results refer to the peviously mentioned problems.

B2 As W2 CPU 373 ieu 1367 problems solved

B3 As W3 CPU 248 ieu 1351 problems solved

B4 As W4 CPU 234 ieu 1351 problems solved

B5 As W5 CPU 174 ieu 1016 problems solved

B6 As W6 CPU 160 ieu 993 problems solved

Consider now the convex underestimate corresponding to the strictly convex terms made up by the maximum value out of the two endpoints tangents . This penalizes the effort to solve each subproblem that is now subjected to a larger number of constraints (61) . The results will refer to breadth first strategy.

E1 As W7

CPU 213 ieu 56 nodes stored 1035 problems solved.

(55 ieu 26 nodes stored and 169 problems solved until the first feasible solution is found)

We remark that in this case a mid-point splitting rule was used after since if we were to use Soland's splitting rule in order to find the first feasible point we would obtain

CPU 324 ieu 209 nodes stored 1037 problems solved

The smallest underestimating subproblem was obtained by using the tangent hyperplanes approximation : L P in 9 variables and subject to 21 constraints

H1 As W7

CPU 21 ieu 20 nodes stored 373 problems solved

(14 ieu 7 nodes stored and 71 subproblems solved until the first feasible point is found)

These results show a marked improvement in the approximation made although the number of subproblems to be solved is still high. Besides the difference might be maximal when a variable is at its endpoints.

These drawbacks are overcome by the use of factorable underestimating functions. The main disadvantage associated with this approach lies in the higher dimensionality of each subproblem : 13 variables and 26 constraints.

F1 As W7 but partitioning the intervals according to Soland's rule

CPU 11 ieu 7 nodes stored 50 problems solved

F2 As W7 but using mid-point splitting rule

CPU 18 ieu 9 stored nodes 89 problems solved

F3 As F1 but instead of nearly feasible values taken as the incumbent a feasible set of areas is found by scaling the nearly feasible design defining a new upper bound

CPU 14 ieu 7 nodes stored 61 problems solved

F4 As F3 but now the search strategy is of the type depth first: choose the best sucessor when available otherwise backtrack to the predecessor and apply the rule.

CPU 15 ieu 6 levels of branching 67 problems solved

F5 As F3 but the dimensionality of each subproblem is increased to 21 variables and 34 constraints

CPU 42 ieu 9 nodes stored 79 problems solved

The small scale example of multiple solution behaviour referred in 4.11.1 was solved by the approach thought of being more efficient. The optimum was located after solving 16 subproblems each having 12 variables and 17 constraints ellapsing 3 ieu and occupying a storage space corresponding to a maximum of 4 nodes. The multiple optima solution until level 3.6 were located after solving 22 subproblems and using 4 ieu since the beggining of the tree.

The ten bar truss was also studied. Convergence to the global solution is very slow when the displacement constraints are active even if small intervals of variation for each variable are considered . This is due to the fact that the problem is ill-conditioned in the sense that small variations in the design variables will lead to big changes in the state variables and vice versa. It is therefore necessary to guarantee the obtaining of a sufficientl y close value to the true optimum by setting the maximimum allowable constraint violation to a very small number. This would lead to both an increase in the number of problems to be solved and of pending nodes.

The conclusions about the main components of the Branch and Bound algorithm will be drawn next : The search strategy or node selection rule is the criterion needed to select a subproblem ie: node of the tree to be examined. The best strategy seems to be between the extremes of depth first and breadth first. Choose the best sucessor node when available but do not perform automatically backtracking when no sucessor is available. Instead in such cases one chooses the node with the best evaluation. In fact the node should take into account besides the lower bound given by the solution of the LP the distance from a feasible solution measured by the product of the dual multipliers by the constraint infeasibility.

The branching rule is the device used for breaking up a subproblem ie : For generating sucessors of a node of the search tree. Soland's weak branching rule was adopted. By contrast to the strong rule where 2^n nodes are created at each iteration in this instance only two problems are originated. Since in all problems the maximum infeasibility has been defined with respect to a single variable the branching rule subsumes the criterion for selecting that variable. Its choice may be governed by various heuristic rules like the largest difference out of all violated constraints or the largest difference in the most violated constraint. The latter scheme seems to be superior to the usual one since the original domain is subject to bilinear equalities. Perhaps the single most important component of all B & B methods from the point of view of its influence on overall efficiency is the lower bounding procedure. The envelopes given for factorable terms within a rectangle of bounds substantially reduce the gap between the nonconvex and underestimating functions.

4.13. "Inside-Out" Approach

In this section an alternative branch and bound algorithm for the minimization of problems that may be reduced to factorable form is presented. Each iteration takes place over a subinterval of the original domain given by the bounds on the variables and consists of three basic steps. The first step of each iteration is to determine a base point from which to branch and bound. Highly desirable points are local optima of the minimum volume design although it is possible to start the algorithm with points that may not be local minima or even feasible. If this algorithm is used as a verification procedure then the local solution obtained by convex programming techniques is an ideal starting point. Take as incumbent solution the OF value that corresponds to a base point if feasible or to the best feasible solution found so far.

Once a base point is obtained the second step consists of eliminating an interval surrounding it. For a feasible point $\tilde{x}^V$ where V represents the iteration number of the algorithm an interval is eliminated for which $\tilde{x}^V$ is global for the least volume problem. It is composed of three basic substeps. First we divide the interval under consideration into subintervals around its base point. Next we define a region of each subinterval over which the base point is global to the original problem. Finally we form a total elimination interval from the union of the regions eliminated over the individual subintervals. For infeasible $\tilde{x}^V$ we find and eliminate an interval surrounding it for which the linear convex envelope problem (LCE) and hence also the least volume problem has no feasible solution.

In the last step we enter the branch and bound section. Uneliminated regions are partitioned into subintervals and a LCE problem is solved for each of them. These lower bounds are compared to the value of the best upper bound (incumbent solution). All previously uneliminated intervals with bounds which equal or exceed the incumbent are eliminated from further consideration. If any subset of the original domain remain a new iteration of the algorithm is initiated over the uneliminated subinterval with the lower bound around a new base point.

Soland's method partitions the original domain into two (or more) m-dim subintervals and the LCE are constructed and solved over each subinterval and their respective minima are recorded. At each subsquent stage partitioning is carried out on the subinterval with the smallest recorded minimum and the process continues. In Fig 4.12 it is represented the difference in strategies between both B & B methods.

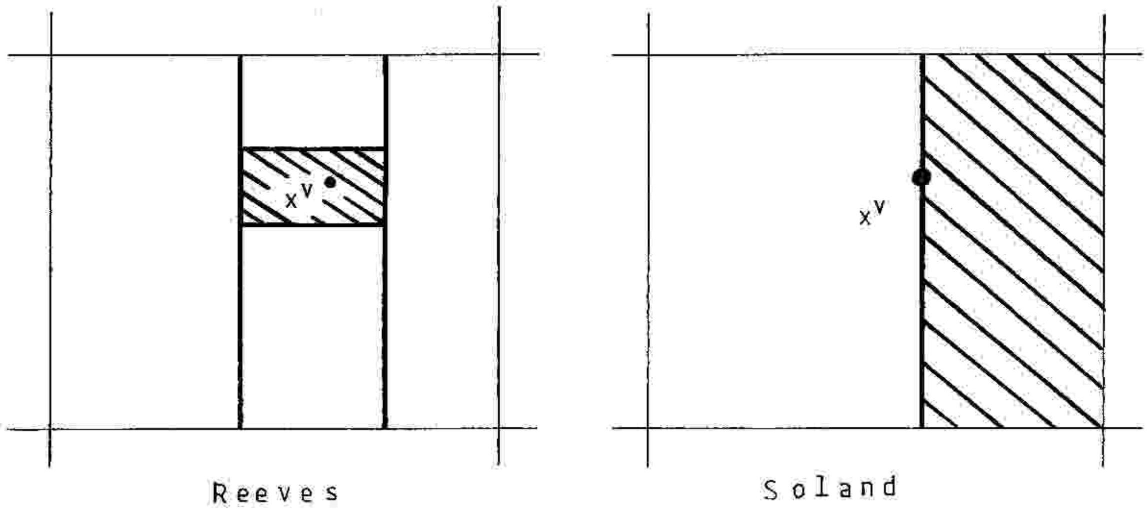


Fig 4.12 Comparison of Reeves and Soland Strategies

The termination criteria for verifying a global minimum of the minimum volume design for a fixed set of bounds requires that the solution point $\tilde{y}^V$ given by LCE on the same interval should also be feasible to the original problem. This implies that all design variables are endpoints ie:

$$\text{either } y_j^V = a_{jl} \quad \text{or } y_j^V = a_{ju} \tag{4.77}$$

We remind the reader that in our case the OF is linear and for any interior point the convex envelope of the constraint functions would give a lower bound on the true values.

Since y_j^V is one endpoint of the region of each subinterval for problem r the other endpoint is a distance Δ_j^r away we want to find Δ_j^r for a particular r. By adding the artificial variables q_i the problem $(V\Delta^r)$ of choosing $\Delta^r, \gamma^r, \mu_a^r, \mu_d^r$ can be written

$$\text{Min} \sum_{i=1}^{m} q_i \tag{4.78}$$

$$\underline{\text{st}} \quad L(\tilde{d},\tilde{\Delta}^r)\tilde{\gamma}^r - \tilde{c}_a \tilde{\mu}_a^r - \tilde{l} - \tilde{q} = \tilde{\emptyset} \tag{4.79}$$

$$K(\tilde{a}\,\tilde{\Delta}^r)\tilde{\gamma}^r - \tilde{c}_d \tilde{\mu}_d^r \quad - \tilde{q} = \tilde{\emptyset} \tag{4.80}$$

$$\tilde{\gamma}, \tilde{\mu}_a, \tilde{\mu}_d, \tilde{q} \geq \emptyset$$

where $\tilde{l}, \tilde{c}_d$ are the vector of member length and active displacements at the local solution respectively. c_{ia} is 1 is $y_j^V = a_{ju}$ and -1 if $y_j^V = a_{jl}$

There is always a feasible set of Lagrange multipliers that solve $(V\Delta^r)$. In particular when $\Delta^r = \emptyset$ the problem $(V\Delta^r)$

reduces to the K-T equalities for the minimum volume design. These subproblems are similar in structure over the various subintervals.

Each problem $(V\Delta^r)$ is nonlinear since terms of the form $\gamma_i^r \cdot \Delta_j^r$ are present but they become linear in $\underset{\sim}{\gamma}^r, \underset{\sim}{\mu}_a^r, \underset{\sim}{\mu}_d^r$ for a particular $\underset{\sim}{\Delta}^r$. We should not attempt to solve them directly. Instead this problem is replaced by a sequence of $(V\Delta^r)$ with small adjustments in previous iterations. As we move from a subproblem to another we never attempt to eliminate more of either the upper or lower range of $\underset{\sim}{a}$ that was eliminated over previous subintervals. While this decision tends to reduce the size of the overall elimination interval it makes the total elimination interval completely rectangular. This reduces the number of (LCE) solved at Step 3.

The algorithm will be briefly stated in the sequel

INITIALIZE

Set V = Ø and the bounds on the variables corresponding to the supplied range .

STEP 1

Find a local minima to the bilinearly constrained problem over the hyper-rectangle of bounds on the variables under consideration.

V = V + 1

STEP 2

Eliminate an interval surrounding $\underset{\sim}{x}^V$ over which $\underset{\sim}{x}^V$ is global to the least volume design. We will consider only the areas as endpoints.

(1) For feasible $\underset{\sim}{x}$

a) Subdivide

Since each factorable term consists of the product of a design variable times a state variable and the volume of the material depends solely on the values assumed for the areas .

For each component $a_{jl} < a_j < a_{ju}$ subdivide the interval $[a^V_{jl}, a^V_{ju}]$ into $[a^V_{jl}, x^V_j]$ and $[x^V_j, a^V_{ju}]$. If There are r such elements this operation results in 2^r subintervals.

b) Investigate individual subintervals

Initialize the bounds for the total elimination interval

If

$$x^V_j = a^V_{jl} \quad \Rightarrow e^V_{jl} = x^V_j \qquad e^V_{ju} = x^V_j + \Delta^V_j$$

$$a^V_{jl} < x^V_j < a^V_{ju} \Rightarrow e^V_{jl} = x^V_j - \Delta^V_j \qquad e^V_{ju} = x^V_j + \Delta^V_j$$

$$x^V_j = a^V_{ju} \quad \Rightarrow e^V_{jl} = x^V_j - \Delta^V_j \qquad e^V_{ju} = x^V_j$$

Find $\underset{\sim}{\Delta}^r$ by solving $(V\Delta^r)$ corresponding to each of the subintervals defined. After all the subproblems are solved for each j = 1,...m choose that corresponds to the minimal value of the range eliminated by all relevant subproblems.

c) Form total elimination interval

The total elimination interval is given by $[e^V_{jl}, e^V_{ju}]$,j=1,...n .

(2) For infeasible $\underset{\sim}{x}$ this procedure is simplified. There

is no subdivision phase and therefore only one elimination interval to consider

Set the rectangle

$$x_j^V = a_{jl}^V \Rightarrow e_{jl}^V = x_j^V \qquad e_{ju}^V = x_j^V + \Delta_j^V$$

$$a_{jl}^V < x_j^V < a_{ju}^V \Rightarrow e_{jl}^V = x_j^V - \Delta_j^V \qquad e_{ju}^V = x_j^V + \Delta_j^V$$

$$x_j^V = a_{ju}^V \Rightarrow e_{jl}^V = x_j^V - \Delta_j^V \qquad e_{ju}^V = x_j^V$$

defining the interval of variation of the design variables

and $e_{n+j,l}^V = d_l$ $\quad e_{n+j,u}^V = d_u$ are the bounds on the displacements.

Construct and solve the LCE over this elimination interval. If it has no feasible solution increase each Δ_j^r by some amount and adjust the bounds accordingly. Repeat this iterative process until either the entire feasible region of the original problem is eliminated or a feasible solution is obtained. In this latter instance then the last interval with no feasible solution will be taken as the elimination interval.

STEP 3

Branch and bound

(1) Partition the remaining regions into subintervals. Initialize the bounds on the remaining intervals $[\underset{\sim}{a}_l^V, \underset{\sim}{a}_u^V]$ to $[\underset{\sim}{a}_l, \underset{\sim}{a}_u]$. Take a region remaining after the elimination interval has been removed from the rectangle of bounds on the areas and for each j=1,...m in turn (where m is the number of members) partition it in intervals

1. If $\underline{a}_{jl}^{V} < e_{jl}^{V}$ then form subinterval $\underline{a}_{jl}^{V} \leq x_j \leq e_{ju}^{V}$ and $\underline{a}_{jl}^{V} \leq x_j \leq \underline{a}_{ju}^{V}$ for the remaining area bounds.
   Set $\underline{a}_{jl}^{V} = e_{jl}^{V}$
2. If $e_{ju}^{V} < \underline{a}_{ju}^{V}$ then form the subinterval $e_{ju}^{V} \leq x_j \leq \underline{a}_{ju}^{V}$ and $\underline{a}_{jl}^{V} \leq x_j \leq \underline{a}_{ju}^{V}$ for the remaining area bounds.
   Set $\underline{a}_{ju}^{V} = e_{ju}^{V}$

(2) Construct and solve LCE over each subinterval.

(3) Examine LCE bounds.

Eliminate those subintervals that either make LCE infeasible or with bounds greater than the incumbent solution. Collect the uneliminated intervals together with those remaining from previous iterations.

Terminate if there are no more pending nodes. Otherwise

(4) Branch to the interval with lowest bound.

Reset the bounds on the design variables to those of this range. Go To Step 1

4.1 3 .1. Applications

Reeves' algorithm was initially intended to solve all-quadratic programming problems. All bilinear forms need to be separated by an appropriate transformation involving the addition of variables. In this paragraph an application of these concepts will be made regarding the three bar truss design problem acted upon by two alternative loading conditions. In order to eliminate an interval surrounding the local optimum at Step 2 up to 2^{12} subproblems would have to be solved at each iteration.

It has been mentioned in 4.12. that a factor of prime importance in view of overall efficiency is the lower bounding procedure adopted. Therefore we will approximate the nonconvex functions with the envelopes obtained from the factorable programming approach. Now only $2^3 = 8$ subproblems $(V\Delta^r)$ need to be solved at Step 2. This is one method for eliminating an interval surrounding a base point. There are obviously many ways of adjusting the upper and lower bounds on such an interval. For instance using the factorable envelopes set the optimal design areas as one of the bounds in each (LCE). This avoids the elimination of a small ε-interval corresponding to two infeasible $(V\Delta^r)$ subproblems defined for opposite $\underset{\sim}{\Delta}^r$.

V = 1

STEP 1

The first step of each iteration is to determine a local optimum from which to branch and bound. This base point may be determined by a convex minimizer routine. Suppose we start with the optimal value

$a_1 = 7.024 \quad a_2 = 2.138 \quad a_3 = 2.756 \quad OF = 15.969$

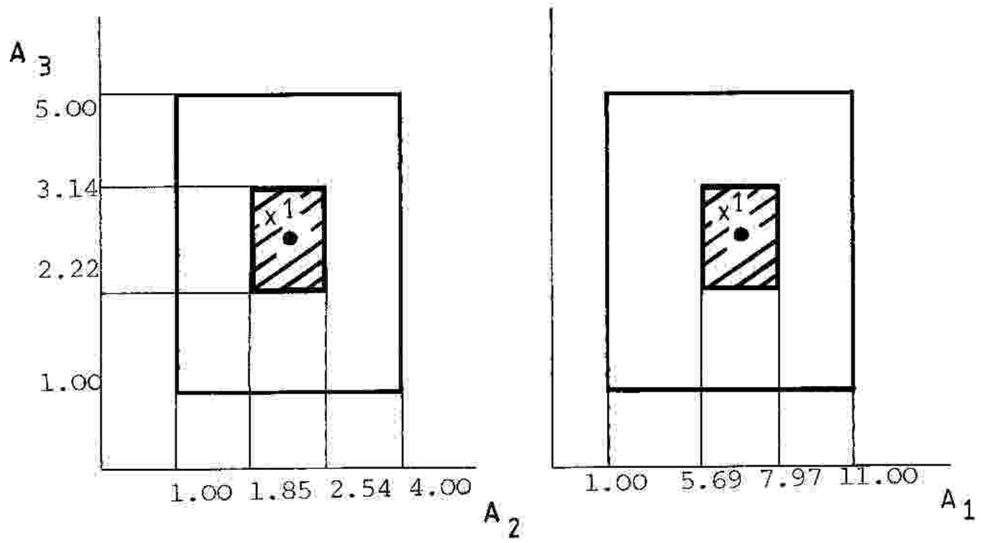
$s_1^1 = 7.071 \quad s_2^1 = 4.566 \quad s_3^1 = -2.505$

$s_1^2 = -1.252 \quad s_2^2 = 5.819 \quad s_3^2 = 7.071$

STEP 2

First we try to eliminate the entire range of a_j ie: $[a_{jl}, a_{ju}]$ for as many j as possible .The purpose of this is to determine whether or not we can reduce branching and bounding in Step 3 that is to minimize the total number of (LCE).

The elimination interval was obtained after two runs of 8 linear subproblems where the maximum infeasibility was limited to an upper value for solutions not coinciding with the base point. The scaled feasible points corresponding to those solutions would have a volume that exceeds the incumbent. Thus the intervals they define can be eliminated



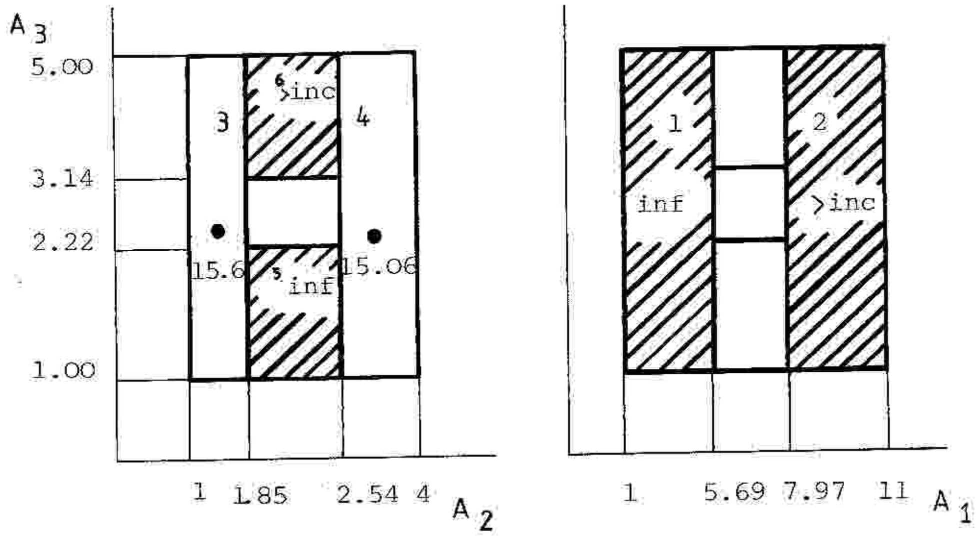
The total elimination interval is given by

$$e_l^1 = [\ 5.69\ 1.85\ 2.22]$$

$$e_u^1 = [7.97\ 2.54\ 3.14]$$

STEP 3

We now take the region remaining after the elimination interval has been removed from the entire range and partition it into 3.2 = 6 regions. Subintervals 1,2,5,6 can be eliminated (1 and 5 because their LCE has no feasible solution 2 and 6 because their lower bound will exceed the incumbent).



For the next iteration of the algorithm we branch to interval 4 with the lowest lower bound.

V = 2

STEP 2

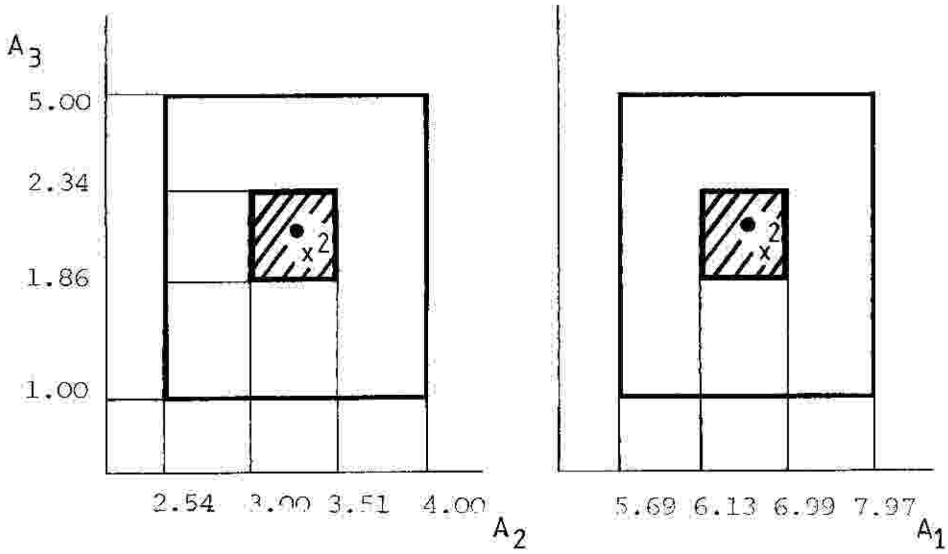
The base point

$a_1 = 6.281$ $a_2 = 3.155$ $a_3 = 2.140$ $OF = 15.064$

$s_1^1 = 7.071$ $s_2^1 = 5.056$ $s_3^1 = -2.015$

$s_1^2 = -1.461$ $s_2^2 = 5.610$ $s_3^2 = 7.071$

is infeasible. The procedure is therefore simplified: There is no subdivision phase and only one interval should be considered. If we use the underestimate given by factorable underestimates we find the total elimination interval by trial and error. We will also limit the maximum infeasibility and eliminate any interval giving a scaled feasible design greater than the best local solution.



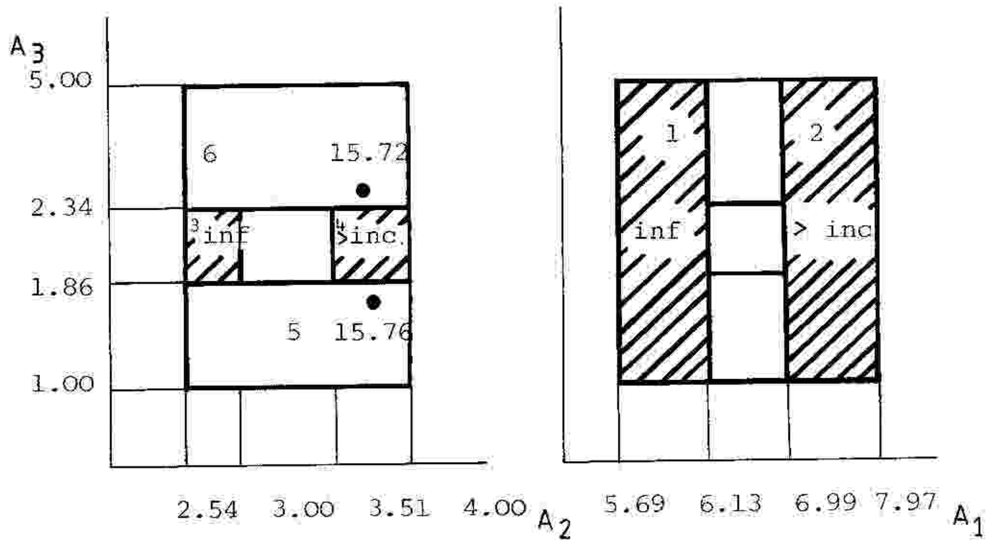
Total elimination interval

$$e_l^2 = [\ 6.13\ \ 3.00\ \ 1.86\ ]$$

$$e_u^2 = [\ 6.99\ \ 3.51\ \ 2.34\ ]$$

STEP 3

The branching and bounding section is entered. The procedure is equivalent to the one described in Step 3 of the previous iteration.



We eliminate the regions corresponding to 3,4,5,6 and choose the bound 15.600 obtained at Step 3 of the previous iteration

V = 3

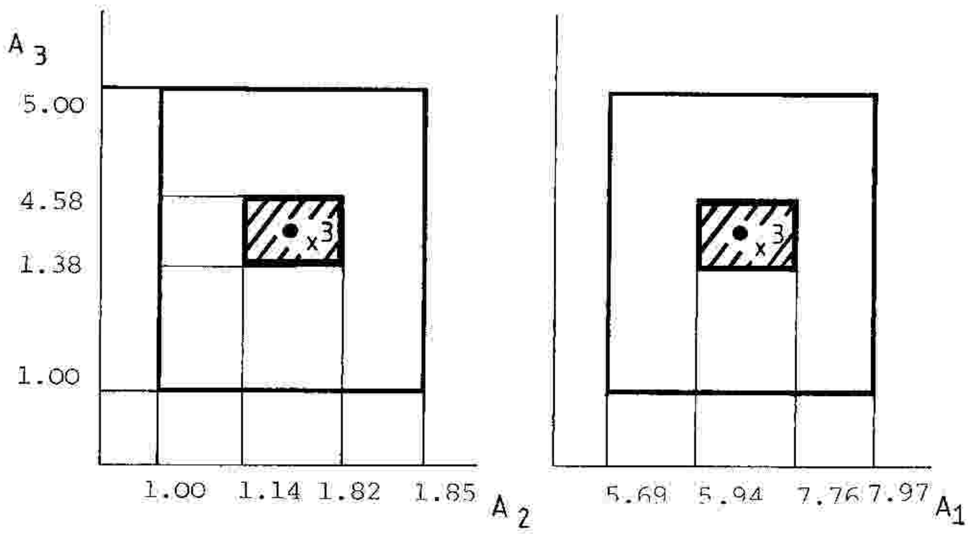
STEP 2

Consider the infeasible base point

$a_1 = 6.940 \quad a_2 = 1.689 \quad a_3 = 2.897 \qquad OF = 15.600$

$s_1^1 = 7.071 \quad s_2^1 = 5.732 \quad s_3^1 = -1.339$

$s_1^2 = -0.910 \quad s_2^2 = 6.161 \quad s_3^2 = 7.071$

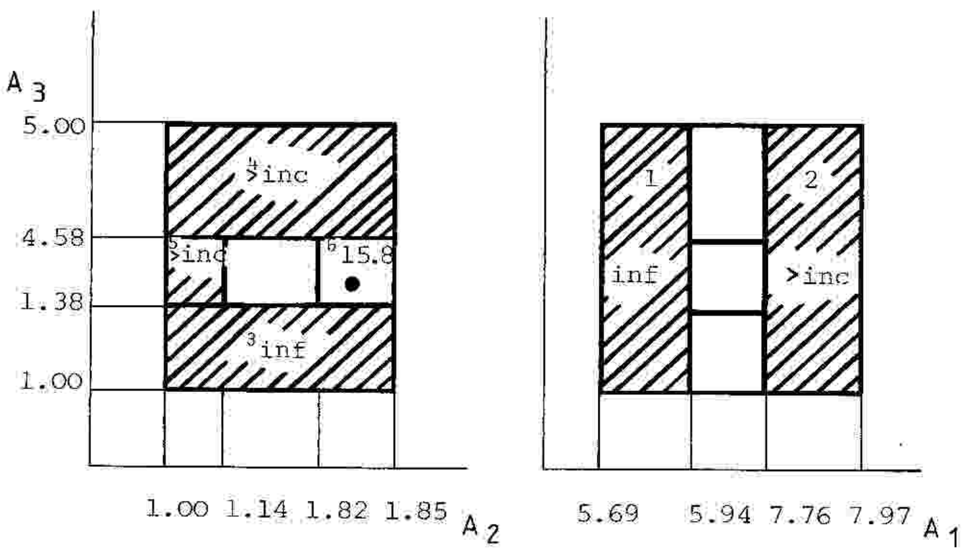


The total elimination interval is given by

$$e_l^3 = [\ 5.94\ 1.14\ 1.38\ ]$$

$$e_u^3 = [\ 7.76\ 1.82\ 4.58\ ]$$

STEP 3



The algorithm would proceed. It terminates after 8 cycles comprising 78 linear underestimating subproblems and elapsing 18 ieu. This verification procedure does not include the effort the local minimization routine has taken to find out the first base point. The problem refered in 4.12. as F3 serves to compare the efficiency of this strategy with Soland's solution method.

4.13.2. Further discussion

The main advantage of this approach lies in the convergence to a ϵ-optima ensured in a finite number of steps [Reeves (1973)] . Soland's method generally does not converge in a finite number of iterations and intermediate solution points generated are not necessarly feasible (although they may be scaled in the least volume design). Also the local minima provide good intermediate solution points. In many instances the first found optimum will be global. They also tend to accelerate the remainder of the algorithm in which we either verify that the local is global or determine a better local and repeat the process. Using local minima as base points tends to eliminate larger intervals. Local minima provide a good basis for comparison for the elimination of entire intervals with greater LCE bounds . Further obtaining a local minimum $\underset{\sim}{x}^*$ and eliminating a region around it increases the likelihood that the LCE will be either infeasible or with a solution value that will exceed the incumbent.

Theoretically in verifying a global minimum the number of elimination subintervals determined at Step 2 grows exponentially with problem size while the number of LCE grows linearly with the number of variables. This appears to be a major drawback to the application of this philosophy. A factor which tends to lessen the importance of the growth rate in the number of elimination subintervals is that as the problem size increases it is increasingly unlikely that all such subintervals will be investigated (we recall that as soon as it is not possible to eliminate any nontrivial interval for a particular subinterval any remaining intervals are not investigated). Another criticism of this method is related to the trial and error procedure for determining each set of Δ^r. A exponentially increasing number of problems $(V\Delta^r)$ may need to be solved a number of times.

There are many possible variations of the algorithm concerning elimination interval stategies (in the formation of the total elimination interval) and partitioning strategies at the last step (use of an hybrid method to avoid the exponential number of subproblems corresponding to increasingly smaller intervals).

CHAPTER FIVE

RESOURCE-DECOMPOSITION APPROACH

5.1. Introduction

The publication of the Dantzig-Wolfe decomposition principle [Dantzig and Wolfe (1960)] initiated the work on large scale mathematical programming that has followed. It operates by forming an equivalent master program (MP) with a smaller number of rows that link blocks of equations but with very many columns that can be generated without being tabulated. The resulting algorithm involves iteration between a set of independent subproblems whose OF contain variable parameters and the MP. The subproblems receive a set of Lagrange multipliers from the MP and they send their solutions back to the MP which connects them with previous results in an optimal way and computes new multipliers. These are again sent to the subproblems and the iteration proceeds until an optimality test is passed.

It can be viewed as an instance of the GLP whose columns are drawn freely from given convex sets. Such a problem can be studied by an appropriate generalization of the duality theorem of LP which presents a sharp distinction to be made between those constraints that pertain only to a part of the problem and those that connects its parts. The decomposition principle has an economic interpretation based upon viewing the Lagrange multipliers as shadow prices. In this decision making scheme that is not truly de c entralized the central agency makes the final decisions by assigning optimal weights for subsystem proposals. Old offers are never forgotten by the coordinating unit.

Consider the action of the ith subsystem and the way they affect the overall objective z viewed here as the total cost to be minimized. If subsystem j chooses an activity vector x_j it incurs a direct cost $c_j x_j$. It also uses the amount $A_j x_j$ of the shared items thus denying them to other subsystems and possibly increase their costs. In order to make subsystem j take this indirect contribution to cost into account a set of shadow prices is announced for the shared items. The subsystem is then forced to pay for whatever the quantities of the resources they use. If a particularly valuable item is assigned a high price then this should discourage the subsystems from using excessive quantities of it as they might if no penalty were imposed.

Bender's resource-decomposition algorithm initially used to solve convex and partially convex problems involving two types of variables was shown to be a dual pair of the Dantzig-Wolfe decomposition principle when applied to the solution of a large scale LP [Lasdon (1970)] : If we have a LP in which the variables are divided into two groups to solve the primal problem by Bender's decomposition is equivalent to apply the Dantzig-Wolfe principle to the dual LP. Of course when one considers nonlinear programs there are important differences between both procedures. The most important is that Bender' s algorithm can handle a much wider variety of programmes than any extension of the decompositon principle.

If one set of the variables $\underset{\sim}{y}$ in the bilinearly constrained optimization is held fixed the resulting problem in the $\underset{\sim}{x}$ set of variables becomes a much easier linear optimization task. Although the former problem is not convex in the $\underset{\sim}{x}$ and $\underset{\sim}{y}$ variables jointly by fixing $\underset{\sim}{y}$ renders it so in $\underset{\sim}{x}$ (LP). It is evident that there are substantial opportunities for achieving computational economy by somehow looking at the y-space rather than in the x.y-space. We expect the nonconvexities to be treated separately from the convex portion of the problem.
The key idea that enables the problem (P)

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{x} \tag{5.1}$$

$$\underline{\text{st}} \quad g_i^T \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} \geq b_i \qquad i=1,\dots m \tag{5.2}$$

$$\underset{\sim}{\emptyset} \leq \underset{\sim}{y} \leq \underset{\sim}{y}_{max} \qquad \text{ie} \qquad y \in Y \tag{5.3}$$

$$\underset{\sim}{x}_l \leq \underset{\sim}{x} \leq \underset{\sim}{x}_u \qquad \text{ie} \qquad x \in X \tag{5.4}$$

to be viewed as a problem in the x-space is the concept of projection (sometimes also known as partitioning)

$$\underset{\underset{\sim}{x}}{\text{Min}} \quad v(\underset{\sim}{x}) \tag{5.5}$$

$$\underline{\text{st}} \quad \underset{\sim}{x} \in X \cap V \tag{5.6}$$

where

$$v(\underset{\sim}{x}) = \text{infimum} \quad \underset{\sim}{c}^T \underset{\sim}{y} \tag{5.7}$$

$$\underline{\text{st}} \quad g_i^T \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i \geq \emptyset \tag{5.8}$$

$$\underset{\sim}{0} \le \underset{\sim}{y} \le \underset{\sim}{y}_{max} \tag{5.9}$$

and

$$V = \{ \underset{\sim}{x} : g_i^T \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i \ge 0 \text{ for some } y \in Y \} \tag{5.10}$$

Note that $v(\underset{\sim}{x})$ is the optimal value of (P) for fixed $\underset{\sim}{x}$ and evaluating $v(\underset{\sim}{x})$ is much easier than to solve the bilinearly constrained problem itself.

Denote $(P(\underset{\sim}{\bar{x}}))$ the optimization problem (5.11)-(5.13)

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{y} \tag{5.11}$$

$$\underline{\text{st}} \quad g_i^T \underset{\sim}{\bar{x}} + \underset{\sim}{\bar{x}}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i \ge 0 \tag{5.12}$$

$$\underset{\sim}{0} \le \underset{\sim}{y} \le \underset{\sim}{y}_{max} \tag{5.13}$$

The set V consists of those values of $\underset{\sim}{x}$ for which $P(\underset{\sim}{\bar{x}})$ is feasible and $X \cap V$ can be thought of as the projection of the feasible region of (P) onto X-space. It will be shown that the projected problem (5.5)-(5.10) is equivalent to the original problem. The difficulty with the use of the latter as a route for solving (P) is that the function v and the set V are only known implicitly via their definitions.

In order to overcome this difficulty a cutting plane method is devised that builds up approximations to v and V. The central idea is to use linear duality theory applied to v and V after projecting the original problem. The master problem will be solved via a process of relaxation that generates dominating approximations to v and V. This is accomplished by obtaining the optimal multiplier vectors for $P(\bar{x})$ corresponding to various trial values of x and adding new cuts to the relaxed master problem as needed.

5.2. Formulation of the master problem

The master problem which is equivalent to (P) is originated by a sequence of three manipulations.

(A) Project (P) onto x resulting in $(P(\bar{x}))$

(B) Invoke the natural dual representation of V in terms of the intersection of a collection of regions that contain it.

(C) Invoke the natural dual representation of v in terms of the pointwise infimum of a collection of functions that dominate it.

(A) is based on the following projection theorem: [Geoffrion (1972)]

Problem (P) is infeasible iff the same is true of (5.5)-(5.10). If x^* is optimal in (5.5)-(5.10) and y^* achieves its minimum in $(P(x))$ with $x=x^*$ then (x^*,y^*) is optimal in (P). This theorem can be extended to cover ε-optimal solutions.

(B) Assuming that the set

$$Z_y = \{ z \in R^m : g_i^T x + x^T H_i y - b_i \geq z_i \}$$

$$0 \leq y \leq y_{max} \qquad i=1,\ldots m \qquad (5.14)$$

is closed for each fixed $x \in X$.

The V-representation theorem states that a point $\bar{x} \in X$ is also in the set V iff it satisfies the system

$$[\inf_{y \in Y} (-\sum_{i=1}^{m} u_i (g_i^T x + x^T H_i y - b_i))] \leq 0$$

$$\underline{\text{all}}\ u \in U \qquad (5.15)$$

where

$$U = \{ u \in R_+^m : \sum_{i=1}^{m} u_i = 1 \} \qquad (5.16)$$

Proof (1) =>

Let $\bar{x}$ be an arbitrary point in X. It is trivial to verify that if

$$\bar{x} \in V = \{ x : g_i^T x + x^T H_i y - b_i \geq 0 \} \qquad (5.17)$$

$$\text{for some } 0 \leq y \leq y_{max} ;\ i=1,...m$$

and for all

$$u_i \geq 0 \ ; \sum_{i=1}^{m} u_i = 1 \qquad \text{ie } u \quad U$$

$$[\sum_{i=1}^{m} (-u_i (g_i^T \bar{x} + \bar{x}^T H_i y - b_i))] \leq 0 \qquad (5.18)$$

$$\text{for some } 0 \leq y \leq y_{max}$$

Hence

$$[\inf_{y \in Y} (-\sum_{i=1}^{m} u_i (g_i^T \bar{x} + \bar{x}^T H_i y - b_i))] \leq 0$$

$$\text{for all } u \in U \qquad (5.19)$$

(2) <=

Suppose that $\bar{x} \in X$ such that

$$[\inf_{y \in Y} (-\sum_{i=1}^{m} u_i (g_i^T \bar{x} + \bar{x}^T H_i y - b_i))] \leq 0$$

for all $u \in U$ (5.20)

Then

$$\sup_{u \in U} [\inf_{y \in Y} (-\sum_{i=1}^{m} u_i (g_i^T \bar{x} + \bar{x}^T H_i y - b_i))] \leq 0 \qquad (5.21)$$

It follows that

$$\sup_{u \in U} [\inf_{y \in Y} (-\sum_{i=1}^{m} u_i (g_i^T \bar{x} + \bar{x}^T H_i y - b_i))] = 0 \qquad (5.22)$$

since scaling does not affect the sign of the bracketed expression.

But this last expression is the dual with respect to the bilinear constraints (for a fixed value of x) of the linear program

$$\min_{y \in Y} \quad 0^T y \qquad (5.23)$$

$$\underline{\text{st}} \quad g_i^T \bar{x} + \bar{x}^T H_i y - b_i \geq 0 \qquad (5.24)$$

having an optimal value of 0. This condition together with the verifiable assumption that Z_y is closed for each $x \in X$ (since it is affine) implies that the primal problem is feasible and hence $\bar{x} \in X$.

QED

REMARK

If the primal problem obtained by fixing $\underset{\sim}{x} = \bar{\underset{\sim}{x}}$ is <u>not</u> feasible

$$\text{Min} \quad \underset{\sim}{c}^T \underset{\sim}{y} \tag{5.25}$$

$$\underline{\text{st}} \quad g_i^T \bar{\underset{\sim}{x}} + \bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i \geq \emptyset \tag{5.26}$$

$$\underset{\sim}{\emptyset} \leq \underset{\sim}{y} \leq \underset{\sim}{y}_{max} \tag{5.27}$$

then there is an extreme ray such that the dual OF increases infinitely along a certain direction. This only happens if the OF of the dual problem (D_r) is positive when $x \in X$ is infeasible

$$\text{Max} \sum_{i=1}^{m} u_i (b_i - g_i^T \bar{\underset{\sim}{x}}) - \sum_{j=1}^{n} v_j \, y_{jmax} \tag{5.28}$$

$$\underline{\text{st}} \quad \sum_{i=1}^{m} u_i (\bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i) - v_j \leq \emptyset \tag{5.29}$$

$$v_j \geq \emptyset \qquad \underset{\sim}{u} \in U \tag{5.30}$$

or

$$\underset{\underset{\sim}{u},\underset{\sim}{v}}{\text{Max}} \sum_{i=1}^{m} u_i (b_i - g_i^T \bar{\underset{\sim}{x}}) - \sum_{j=1}^{n} v_j \, y_{jmax} \tag{5.31}$$

$$\underline{\text{st}} \quad v_j \geq \sum_{i=1}^{m} u_i (\underset{\sim}{x}^T \underset{\sim}{H}_i) \tag{5.32}$$

$$v_j \geq \emptyset \qquad \underset{\sim}{u} \in U \tag{5.33}$$

We have

$$\max_{\underset{\sim}{u}} \sum_{i=1}^{m} u_i (b_i - g_i^T \bar{\underset{\sim}{x}}) - \sum_{j=1}^{n} \left[\sum_{i=1}^{m} u_i (\bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i) \right]^+ y_{jmax}$$

$$\underset{\tilde{u} \in U}{} \tag{5.34}$$

where the notation $[(\ldots)]^+$ means max $[\emptyset,(\ldots)]$

Wolsey (1981) arrived at these cuts by using dual functions in a general duality scheme. Let $\bar{\tilde{u}}$ be the solution to problem (5.34). A cut function is defined such that

If $\tilde{x}$ is infeasible

$$-\sum_{i=1}^{m} \bar{u}_i (b_i - g_i^T \tilde{x}) + \sum_{j=1}^{n} [\sum_{i=1}^{m} \bar{u}_i (\tilde{x}^T \tilde{H}_i)]_j^+ \; y_{jmax} < \emptyset \tag{5.35}$$

For all feasible $\tilde{x} \in X$

$$-\sum_{i=1}^{m} \bar{u}_i (b_i - g_i^T \tilde{x}) + \sum_{j=1}^{n} [\sum_{i=1}^{m} \bar{u}_i (\tilde{x}^T \tilde{H}_i)]_j^+ \; y_{jmax} \geq \emptyset \tag{5.36}$$

(C) Assuming that $v(\bar{\tilde{x}})$ is finite and $P(\bar{\tilde{x}})$ possesses an optimal vector for each fixed $\bar{\tilde{x}} \in X \cap V$ the v-representation theorem states that the optimal value of $P(\bar{\tilde{x}})$ equals that of its dual on $X \cap V$ ie:

$$v(\tilde{x}) = \sup_{u \geq 0} [\inf_{y \in Y} c^T \tilde{y} - \sum_{i=1}^{m} u_i (g_i^T \bar{\tilde{x}} + \bar{\tilde{x}}^T \tilde{H}_i \tilde{y} - b_i)]$$
$$\text{for all } \bar{\tilde{x}} \in X \cap V \tag{5.37}$$

Proof

The proof is a result of LP duality. Let $\bar{\tilde{u}}$ be a optimal multiplier vector for the primal problem; then $\bar{\tilde{u}}$ is also an optimal solution of the dual and the optimal values of the primal and dual problems are equal e.g.

For each fixed $\bar{x} \in X \cap V$ the LP

$$\min_{y \in Y} c^T y \tag{5.38}$$

$$\text{st} \quad g_i^T \bar{x} + \bar{x}^T H_i y - b_i \geq 0 \tag{5.39}$$

has the following dual

$$\max_{u \geq 0} \left[\inf_{y \in Y} c^T y - \sum_{i=1}^{m} u_i \left(g_i^T \bar{x} + \bar{x}^T H_i y - b_i\right)\right] \tag{5.40}$$

REMARK

For each feasible $\bar{x} \in X \cap V$ consider the problem $P(\bar{x})$

$$\text{Min} \quad c^T y \tag{5.41}$$

$$\text{st} \quad g_i^T \bar{x} + \bar{x}^T H_i y - b_i \geq 0 \tag{5.42}$$

$$0 \leq y \leq y_{max} \tag{5.43}$$

ie

$$y \geq 0 \text{ and } -y \geq -y_{max}$$

Using duality we obtain the following LP

$$\text{Max} \sum_{i=1}^{m} u_i \left(b_i - g_i^T \bar{x}\right) - \sum_{j=1}^{n} v_j \, y_{jmax} \tag{5.44}$$

$$\text{st} \quad \sum_{i=1}^{m} u_i \left(\bar{x}^T H_i\right) - v_j \leq c_j \tag{5.45}$$

$$u, v \geq 0$$

or

$$\text{Max} \sum_{i=1}^{m} u_i \left(b_i - g_i^T \bar{x}\right) - \sum_{j=1}^{n} v_j \, y_{jmax} \tag{5.46}$$

$$\text{st} \quad v_j \geq -c_j + \sum_{i=1}^{m} u_i \left(\bar{x}^T H_i\right) \tag{5.47}$$

$$v_j \geq 0$$

$$\underset{\sim}{u}, \underset{\sim}{v} \geq \emptyset$$

Thus

$$\underset{\underset{\sim}{u} \geq 0}{\text{Max}} \sum_{i=1}^{m} u_i (b_i - g_i^T \bar{\underset{\sim}{x}}) - \sum_{j=1}^{n} [-c_j + \sum_{i=1}^{m} u_i (\bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i)]_j^+ \; y_{jmax}$$

(5.48)

Let the optimal multiplier vector (dual variables) of this LP be $\bar{u}$

A support function is defined for each feasible $\underset{\sim}{x}$ such that will give a lower bound on the solution of (P) ie:

$$\sum_{i=1}^{m} \bar{u}_i (b_i - g_i^T \underset{\sim}{x}) - \sum_{j=1}^{n} [-c_j + \sum_{i=1}^{m} \bar{u}_i (\underset{\sim}{x}^T \underset{\sim}{H}_i)]_j^+ \; y_{jmax} \leq v(x)$$

for all $\underset{\sim}{x} \in X$ (5.49)

and

$$\sum_{i=1}^{m} \bar{u}_i (b_i - g_i^T \bar{\underset{\sim}{x}}) - \sum_{j=1}^{n} [-c_j + \sum_{i=1}^{m} \bar{u}_i (\bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i)]_j^+ \; y_{jmax} = v(\bar{x})$$

(5.50)

By using the three theorems it is possible to define the following master problem :

$$\underset{\underset{\sim}{x} \in X}{\min} \; [\; \underset{\underset{\sim}{u} \geq 0}{\sup} \; [\; \underset{\underset{\sim}{y} \in Y}{\inf} \; \underset{\sim}{c}^T \underset{\sim}{y} - \sum_{i=1}^{m} u_i^s (g_i^T \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i)]]$$

for all $\underset{\sim}{u}^s \geq \emptyset$ (5.51)

$$\underline{\text{st}} \quad [\; \underset{\underset{\sim}{y} \in Y}{\inf} - \sum_{i=1}^{m} u_i^c (g_i^T \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{H}_i \underset{\sim}{y} - b_i)] \geq \emptyset \quad (5.52)$$

for all $\underset{\sim}{u}^c \geq \emptyset$ and $\sum_{i=1}^{m} u_i^c = 1$ (5.53) or using the definition of supremum as the smallest upper bound

$$\text{Min } \eta \tag{5.54}$$

st (i) $\eta - \sum_{i=1}^{m} u_i^s (b_i - g_i^T \underset{\sim}{x})$

$$+ \sum_{j=1}^{n} [-c_j + \sum_{i=1}^{m} u_i^s (\underset{\sim}{x}^T \underset{\sim}{H}_i)_j]^+ \; y_{jmax} \geq \emptyset \tag{5.55}$$

(ii) $- \sum_{i=1}^{m} u_i^c (b_i - g_i^T \underset{\sim}{x})$

$$+ \sum_{j=1}^{n} [\sum_{i=1}^{m} u_i^c (\underset{\sim}{x}^T \underset{\sim}{H}_i)_j]^+ \; y_{jmax} \geq \emptyset \tag{5.56}$$

$$\underset{\sim}{u}^S \geq \emptyset \quad \underset{\sim}{u}^C \geq \emptyset \quad \text{and} \quad \sum_{i=1}^{m} u_i^c = 1 \tag{5.57}$$

This master problem (MP) is therefore equivalent to the original problem (P). However (MP) is of theoretical interest only since it has an enormous number of constraints. But it can be solved via a series of subproblems. At each iteration a relaxed version of the (MP) containing only few of the constraints of type (i) and (ii) is solved. The solution $(\bar{\eta}, \bar{\underset{\sim}{x}})$ will be tested for feasibility in the initially unrelaxed master problem by solving the subproblem $(P(\bar{\underset{\sim}{x}}))$ or its dual) and either new cuts or support functions will be added until a termination criteria shows that a solution of acceptable accuracy has been obtained.

Both the cut and support functions define a piecewise concave region and each relaxed master problem (RMP) will consist of a minimization over a piecewise concave region ie: nonconvex programming. The disjuntive terms in both support and cut functions can be reformulated by introducing binary variables so that (RMP) becomes a standard mixed 0-1 LP. Let L and U be the lower and upper bounds on the affine expression in each term

$$L \leq \underset{\sim}{f}^T \underset{\sim}{x} + e \leq U \tag{5.58}$$

$$0 \leq \delta \leq 1 \tag{5.59}$$

is the interval of variation for any binary variable $\delta \in B = \{0,1\}$. By introducing the new variable r the disjunction can be linearized

$$r \leq \delta U \tag{5.60}$$

$$r \leq (\delta - 1) L + (\underset{\sim}{f}^T \underset{\sim}{x} + e) \tag{5.61}$$

A major drawback of this substitution is the increase in problem size at each iteration due to the introduction of a number of new variables and constraints.

Alternatively (RMP) can be written as a complementarity programming problem by introducing two real variables r and q and a constraint for each of the terms $[\ \underset{\sim}{f}^T \underset{\sim}{x} + e\ ]$ such that

$$q = r - \underset{\sim}{f}^T \underset{\sim}{x} - e \tag{5.62}$$

and the complementarity condition

$$r^T q = 0 \tag{5.63}$$

thus saving the number of constraints and avoiding the use of integer variables.

5.3. Relaxation process

In the following relaxed master problem only few of the constraints (i) and (ii) of the master problem are included

$$\text{Min } \eta \tag{5.64}$$

$$\underline{\text{st}} \quad \text{(i)} \quad \eta - \sum_{i=1}^{m} \bar{u}_i^t (b_i - g_i^T x) + \sum_{j=1}^{n} [-c_j + \sum_{i=1}^{m} \bar{u}_i^t (x^T H_i)_j]^+ y_{jmax} \geq 0 \tag{5.65}$$

$$\text{(ii)} \quad -\sum_{i=1}^{m} u_i (b_i - g_i^T x) + \sum_{j=1}^{n} [\sum_{i=1}^{m} u_i (x^T H_i)_j]^+ y_{jmax} \geq 0 \tag{5.66}$$

$$x \in X \qquad t=1,\ldots t_f \qquad s=1,\ldots s_f \tag{5.67}$$

An optimal solution $(\bar{\eta},\bar{x})$ of (RMP) is also optimal for the unrelaxed (MP) and therefore the original program (P) iff $(\bar{\eta},\bar{x})$ is feasible for (MP). Furthermore subproblem $(P(x))$ is used to test $(\bar{\eta},\bar{x})$ for feasibility in (MP).

A- $\bar{x}$ satisfies the constraint set (ii) of (MP) iff $(P(\bar{x}))$ is feasible. The feasibility of $(P(\bar{x}))$ implies that $\bar{x} \in X \cap V$ and hence satisfies the constraints (ii).

B- if $(P(\bar{x}))$ is feasible then $(\bar{\eta},\bar{x})$ satisfies the constraint set (i) iff $\bar{\eta} \geq v(\bar{x})$.

Thus $(\bar{\eta},\bar{x})$ is feasible for (MP) iff A- $(P(\bar{x}))$ is feasible and B- $\bar{\eta} \geq v(\bar{x})$.

First suppose $(P(\bar{x}))$ is not feasible. This means that the its dual problem $(D(\bar{x}))$ increases along a certain direction ie: a combination of constraints that has no solution in Y. Hence $\bar{x}$ does not satisfy some of the constraints of type (ii) in (MP). To eliminate this inadmissible point $\bar{x}$ the multiplier vectors corresponding to the dual of

$$\inf \{ \emptyset^T y : \bar{x}^T H_i y \leq b_i - g_i^T \bar{x} \quad ; i=1,..m \; ; y \in Y \}$$

are used to build up a cut function (constraint of type (ii)) that is added to the (RMP).

Next suppose that $(P(\bar{x}))$ is feasible but $\bar{\eta} < v(\bar{x})$. This implies that some of the constraints of type (i) are violated. In this event the dual variables $\bar{u}$ come close enough to be optimal in the dual of $(P(\bar{x}))$. Since $v(\bar{x}) > \bar{\eta}$, these u are called near optimal vectors. In order to satisfy the violated constraints we save $\bar{u}$ and add a support function (constraint of type (i)) to the (RMP).

The relaxed master problem is solved by successively cutting off solutions which do not satisfy all constraints. Both the cut and the support functions are piecewise linear convex but because of the signs of the inequalities involved they give rise to nonconvex features. In Fig 5.1 the several subdomains on the x-space represent the feasible region left after a number of such cuts have been introduced and the function over these regions is the lower edge of η that is supported by a number of nonconvex "kinks".

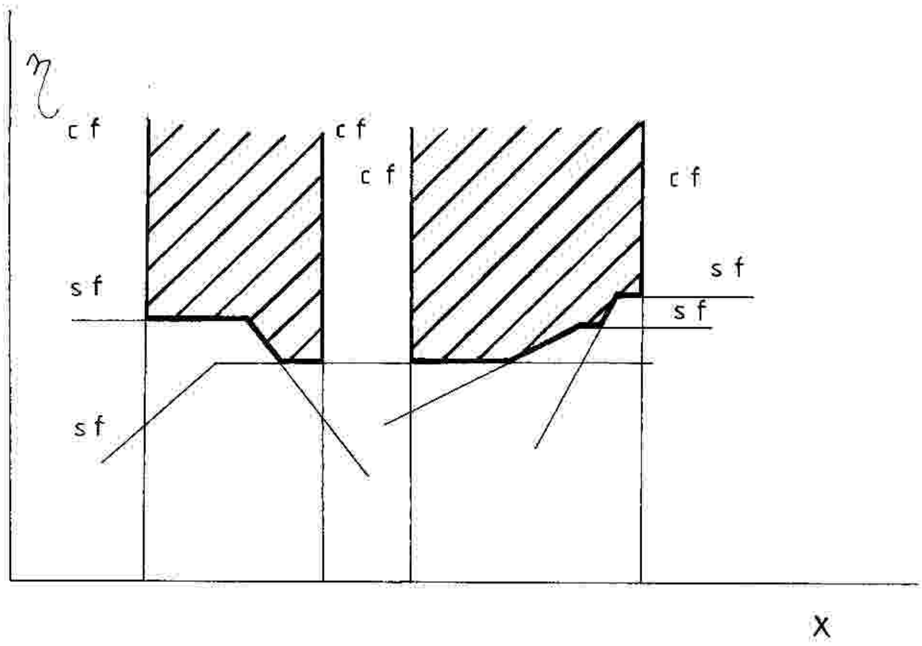


Fig 5.1 Resource-decomposition Algorithm

5.4. The R-D algorithm

The Resource-Decomposition algorithm is stated below:

STEP 1

Let $UB = +\infty$ and $LB = -\infty$ be the initial upper and lower bounds respectively on the optimal value of the OF $\underset{\sim}{c}^T \underset{\sim}{y}$ of program (P). Set a tolerance value ε and $k = 1 = \emptyset$. Generate $\underset{\sim}{x} \in X$ and Go To Step 3

STEP 2

Solve the (RMP)

(1) If (RMP) has no feasible solution there is no feasible solution to (P).

(2) let $(\bar{\eta}, \bar{\underset{\sim}{x}})$ be optimal solution to (RMP). Put $LB = \bar{\eta}$

If $UB - LB < \varepsilon$ Terminate. $(\bar{\tilde{x}}, \bar{\tilde{y}})$ is the optimal solution to (P).

STEP 3

Solve the dual $D(\bar{\tilde{x}})$ of the projected problem $(P(\bar{\tilde{x}}))$

(1) If the dual is unbounded Go To Step 5 .

(2) Let $\bar{\tilde{u}}$ be the optimal solution to $(D(\bar{\tilde{x}}))$. $v(\bar{\tilde{x}})$ is given by $\tilde{c}^T \bar{\tilde{y}}$ where $\bar{\tilde{y}}$ is the set of multipliers of $(D(\bar{\tilde{x}}))$. If $v(\bar{\tilde{x}}) - LB \leq \varepsilon$ Terminate : $(\bar{\tilde{x}}, \bar{\tilde{y}})$ is the optimal solution to (P)

STEP 4

If $v(\bar{\tilde{x}}) < UB$ set $UB = v(\bar{\tilde{x}})$

Let $t_f = t_f + 1$ and $\bar{\tilde{u}}^{t_f} = \bar{\tilde{u}}$

Add a constraint of type (i) to the (RMP) and Go To Step 2

STEP 5

Generate a dual ray via the LP (D_r)

Let $s_f = s_f + 1$ and $\bar{\tilde{u}}^{s_f} = \bar{\tilde{u}}$

Add a constraint of type (ii) to the (RMP) and Go To Step 2

5.5. Structural Synthesis Problem

The application of the algorithm described above to our structural optimization problem is straight forward.

Let

$\tilde{x} = \tilde{s}$ vector of member stresses

$\tilde{y} = \tilde{a} - \tilde{a}_1$ corresponding to a translation of the axis corresponding to member areas

$\underset{\sim}{c} = \underset{\sim}{l}$ member lengths defined in the OF of P(x)

$$\underset{\sim}{H}_i^k = \begin{bmatrix} -\underset{\sim}{H}_i^k \\ -\underset{\sim}{H}_i^k \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \end{bmatrix} \qquad g_i = \begin{bmatrix} -\underset{\sim}{H}_i^k \underset{\sim}{a}_l \\ \underset{\sim}{H}_i^k \underset{\sim}{a}_l \\ (\underset{\sim}{L} \;\; \underset{\sim}{B})^T \\ \underset{\sim}{D}^T \\ -\underset{\sim}{D}^T \end{bmatrix} \qquad b = \begin{bmatrix} \underset{\sim}{\lambda}^k \\ -\underset{\sim}{\lambda}^k \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{d}_u \\ -\underset{\sim}{d}_l \end{bmatrix}$$

are the matrix and vectors corresponding to bilinear linear terms and the right hand side of the constraints respectively.

5.6. Global optimality

If the algorithm is terminated with a feasible solution the global optimum is reached.

Proof

When $\bar{\underset{\sim}{u}}$ is an optimal multiplier vector it indexes a constraint of type (i) that is most violated at $(\bar{\eta},\bar{\underset{\sim}{x}})$. When no such optimal multiplier exists how near $\bar{\underset{\sim}{u}}^t$ comes to indexing a most violated constraint depends solely on how nearly it solves the dual problem. How close $\bar{\underset{\sim}{u}}^S$ comes indexing a most violated constraint of type (ii) depends only on how close it solves the dual problem of

$$\inf \{ \underset{\sim}{\emptyset}^T \underset{\sim}{y} : g_i^T \bar{\underset{\sim}{x}} + \bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i \underset{\sim}{y} \geq b_i \; ; \; i=1,\ldots m \; ; \underset{\sim}{y} \in Y \}$$

(5.68)

Therefore $(\bar{\eta}, \bar{\underset{\sim}{x}})$ will be infeasible at the next iteration of the algorithm.

(1) If $\bar{\underset{\sim}{x}}$ is feasible the new constraint of type (i) will cut-off $(\bar{\eta}, \bar{\underset{\sim}{x}})$.

(2) If $\bar{\underset{\sim}{x}}$ is infeasible the new constraint of the type (ii) will cut-off $\bar{\underset{\sim}{x}}$.

To prove optimality it suffices to note that since (RMP) is a relaxation of (MP) will always be a lower bound on $v(\bar{\underset{\sim}{x}})$.

QED

REMARK :

Although the sequence of LB is monotonic nondecreasing the sequence of values for $v(\bar{\underset{\sim}{x}})$ needs not be nonincreasing.

5.7. Acceleration Algorithm for Structural Optimization

One of the most basic properties of a trussed type of structure is the scaling invariance of the stress resultant vector . The internal forces in a statically indeterminate structure are a function of the cross section . However if all areas are multiplied by a positive scaling factor ρ the member forces remain unchanged and all member stresses would be multiplied by $1/\rho$. The nodal displacements would also be affected by a factor of $1/\rho$ since they can be represented by linear combinations of the member stresses. These scaling properties will be used to devise a simplified version of the R-D algorithm. The algorithm previously presented requires that each $\underset{\sim}{y}$ variable be situated in the nonnegative half space and the corresponding nonnegative piecewise concave terms will be multiplied by the upper bound on $\underset{\sim}{y}$, which has caused $\underset{\sim}{x}$ to represent the vector of member stresses.

It is evident that a bar is subjected to either compression or tension but not to both at the same time. Each member stress may be restricted to vary in the half-space that will be nonpositive if the member is compressed or nonnegative if the member is tensioned. Now if we let $\underset{\sim}{y}$ represent the state variables and $\underset{\sim}{x}$ the design variables

$$[\, \bar{\underset{\sim}{u}}^T \underset{\sim}{H}\, x_j - c\,]_j^+ \; y_{j\max} \qquad \underset{\sim}{\emptyset} \leq \underset{\sim}{y} \leq \underset{\sim}{y}_{\max} \tag{5.69}$$

can be transformed into

$$[\, (\bar{\underset{\sim}{u}}^T \underset{\sim}{H}_j x - c)_j \; (y_{j\min} + y_{j\max})\,]^+$$

$$\underset{\sim}{y}_{\min} \leq \underset{\sim}{y} \leq \underset{\sim}{y}_{\max} \tag{5.70}$$

where either $y_{j\min}$ or $y_{j\max}$ will be zero. The nature of each piecewise concave term would therefore be maintained eg: in the case of $y_{j\max} = \emptyset$ we would consider only the nonpositive factor $(\bar{\underset{\sim}{u}}^T \underset{\sim}{H}\, x_j - c)_j$

The projected problem $(P(\bar{\underset{\sim}{x}}))$ onto $\underset{\sim}{y}$ is

$$\text{Min} \quad \underset{\sim}{1}^T \bar{\underset{\sim}{x}} \tag{5.71}$$

$$\underline{\text{st}} \quad \bar{\underset{\sim}{x}}^T \underset{\sim}{H}_i^k \underset{\sim}{y}^k = \lambda_i^k \qquad i=1,\ldots\beta \tag{5.72}$$

$$d_l^k \leq D\, y^k \leq d_u^k \tag{5.73}$$

$$B^T L\, y^k = 0 \tag{5.74}$$

$$y_l^k \leq y^k \leq y_u^k \tag{5.75}$$

Therefore by fixing a set of areas $\bar{x}$ a unique set of stresses can be determined by matrix inversion from the equality constraints corresponding to equilibrium and the compatibility equations. Assuming that a feasible set of state variables was determined by using a scaling factor of $0 < \rho < 1$ the volume of the structure could be reduced until it eventually touches the boundary of the space of stress/displacements. Alternatively if the stress/displacements are outside their rectangle of bounds the design variables could be multiplied by $\rho > 1$ and the stress resultant vector would be linearly reduced until it fits within its bounds.

The scaled vector of areas $\rho\bar{x}$ to be considered in this version has at least one member fully stressed or a displacement at a boundary. The solution of $(P(\rho\bar{x}))$ is unique. To avoid degeneracy a set of multipliers could be determined by finding the product of the matrix used to find the set of stresses by a unit vector containing the Lagrange multipliers of the stress/displacement at their boundary . The dual ray corresponding to the previously infeasible set of areas $\bar{x}$ can be obtained by mutiplying by ρ the dual variables corresponding to the equilibrium equations while the multipliers related to the linear active constraints would remain unchanged.

Since the (RMP) gives a nondecreasing lower bound on the final solution at each iteration the state variables resultant will be located outside their bounds unless the design is optimal. A scaling factor $\rho > 1$ would be determined in order to make $(P(\rho \underset{\sim}{x}))$ feasible. The infeasible set of areas define the lower bound as $\underset{\sim}{1}^T \underset{\sim}{x}$.

$$\text{Min} \quad \eta = \underset{\sim}{1}^T x \qquad (5.76)$$

$$\underline{\text{st}} \quad -\sum_{i=1}^{m} \bar{u}_i^s b_i + \sum_{j=1}^{n} \left[\sum_{i=1}^{m} \bar{u}_i^s (\underset{\sim}{x}^T \underset{\sim}{H}_i + f_i)_j \right.$$

$$\left. (y_{jl} + y_{ju}) \right]^+ \geq 0 \qquad (5.77)$$

$$\underset{\sim}{x} \in X \qquad s=1,\dots s_f$$

This relaxed accelerated master problem (RAP) can also be converted either into a standard 0-1 mixed LP or a CCP that has to be solved at each iteration of the following algorithm:

STEP 1

Let $UB = +\infty$ and $LB = \underset{\sim}{1}^T \underset{\sim}{x}_1$ be the initial upper and lower bounds on the opimal solution of problem (P).

Set the tolerance values ε . Let $s_f = 0$ and $\underset{\sim}{x} = \underset{\sim}{x}_1$

STEP 2

Find a scaled feasible stress/displacement vector and the corresponding factor ρ using $\bar{x}$.

(1) $\rho < 1 + \varepsilon$ Terminate : $(\bar{x},\bar{y})$ is the optimal solution to (P)

(2) Find a set of Lagrange multipliers corresponding to the scaled vector of design variables $\rho \bar{x}$.

Let $s_f = s_f + 1$ and $\bar{u}_i^{s_f} = \bar{u}_i$ for i corresponding to the constraints not involving $\rho \bar{x}$ and $\bar{u}_i^S = \rho \bar{u}_i$ for the remaining variables.

If $\underset{\sim}{1}^T \underset{\sim}{\bar{x}} < UB$ and $\underset{\sim}{\bar{x}} \in X$ set $UB = \rho \underset{\sim}{1}^T \underset{\sim}{\bar{x}}$. Add a constraint to the (RAP).

STEP 3

Solve the (RAP)

(1) If (RAP) has no feasible solution there is no feasible solution to (P).

(2) Let $(\bar{\eta}, \underset{\sim}{\bar{x}})$ be the optimal solution to (RAP). Put $LB = \bar{\eta}$

If $UB - LB < \varepsilon$ Terminate : $(\underset{\sim}{\bar{x}}, \underset{\sim}{\bar{y}})$ is the optimal solution to (P). Otherwise

Go To Step 2

In this new formulation $\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{l}$ represent vectors of member areas,stresses and length repectively

$$X = \{ \underset{\sim}{a} : \underset{\sim}{a}_l \le \underset{\sim}{a} \le \underset{\sim}{a}_u \}$$

is the hyper rectangle of bounds on the design variables.

$$\underset{\sim}{H}_i^k = \begin{bmatrix} \underset{\sim}{H}_i^k \\ -\underset{\sim}{H}_i^k \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \end{bmatrix} \quad \underset{\sim}{f} = \begin{bmatrix} \underset{\sim}{\emptyset} \\ \underset{\sim}{\emptyset} \\ \underset{\sim}{D} \\ -\underset{\sim}{D} \\ \underset{\sim}{I} \\ -\underset{\sim}{I} \end{bmatrix} \quad \underset{\sim}{b} = \begin{bmatrix} \underset{\sim}{\lambda}^k \\ -\underset{\sim}{\lambda}^k \\ \underset{\sim}{d}_l^k \\ -\underset{\sim}{d}_u^k \\ \underset{\sim}{y}_l^k \\ -\underset{\sim}{y}_u^k \end{bmatrix}$$

$$\underset{\sim}{y}_l = \underset{\sim}{y}_{min} = \underset{\sim}{s}_1^k \qquad \underset{\sim}{y}_u = \underset{\sim}{y}_{max} = \underset{\sim}{s}_1^k$$

5.8. Computational Considerations

The reformulation of bilinearly constrained problems although leading to a nonconvex master problem may be of some computational interest since it is possible to convert this problem into a standard 0-1 mixed LP for which Land and Powell (1963) routines are applicable. This nonconvex problem grows into a considerable size if the number of iterations until the optimal solution is reached is high.

Alternative procedures to solve this problem are Branch and Bound strategy based on Lagrangian relaxation and implicit enumeration methods. Some results for the standard problems tested will be reported in a forthcoming section. The 0-1 mixed LP can also be converted into a real quadratic concave minimization linearly constrained by using the methods described in (2.12) This concave quadratic problem can be solved by piecewise linearization. A translation to the mid point along the axes corresponding to the 0-1 variables (which would be allowed to vary from 0 to 1) lead to a minimization of a piecewise linear concave function over a linear domain.

5.9. Applications

The bilinearly constrained problem possessing multiple optima already solved by the Branch and Bound method will be the subject of our study in this section. Since we want the y-variables to vary within the nonnegative half-space but including the axes a translation becomes necessary

$$\text{Min} \quad y_1 + y_2 + y_3 + .3$$

$$\underline{\text{st}} \quad x_1 y_1 + x_3 y_3 + .1 x_1 + .1 x_3 = 0$$

$$3 x_1 y_1 + 1.2 x_2 y_2 - x_3 y_3 + .3 x_1 + .12 x_2 - .1 x_3 = 10$$

$$5 x_1 + x_2 + x_3 \le 2.5$$

$$\underset{\sim}{x}_l^T = [0. \quad 0. \quad -2.5] \qquad \underset{\sim}{x}_u^T = [2.5 \quad 2.5 \quad 0.]$$

$$\underset{\sim}{y}_l^T = [0. \quad 0. \quad 0.] \qquad \underset{\sim}{y}_u^T = [4.9 \quad 4.9 \quad 4.9]$$

INITIALIZATION

STEP 1

Let UB $= +\infty$ LB $= -\infty$ be the initial upper and lower bounds on the optimal solution.

Set $\varepsilon = .001$ and $t_f = 1$. Let $x \in X$ be

$x_1 = .7 \quad x_2 = 1.5 \quad x_3 = -2.5$

STEP 3

$$\text{Min } y_1 + y_2 + y_3 + .3$$

$$\underline{\text{st}} \quad .7 y_1 - 2.5 y_3 = .13$$

$$2.1 y_1 + 1.8 y_2 + 2.5 y_3 = 9.36$$

$$-y_i \ge -4.9 \qquad y_i \ge 0 \quad i=1,\dots 3$$

has the dual

$$\text{Max } .3 + .18\, u_1 + 9.36\, u_2 - 4.9\, u_6 - 4.9\, u_7 - 4.9\, u_8$$

<u>st</u> $.7\, u_1 + 2.1\, u_2 + u_3 - u_6 \leq 1$

$1.8\, u_2 + u_4 - u_7 \leq 1$

$-2.5\, u_1 + 2.5\, u_2 + u_5 - u_8 \leq 1$

u_1, u_2 unrestricted and u_3 to $u_8 \geq 0$

The solution of the dual is $4.588 = v(\bar{\underset{\sim}{x}})$. This means that the primal problem is feasible and the given vector $\bar{\underset{\sim}{x}}$ has a solution point.

(2) $\bar{\underset{\sim}{u}}^T = [.057\ .457\ 0.\ 0.\ .177\ 0.\ 0.\ 0.]$

$v(\bar{\underset{\sim}{x}}) - LB > \varepsilon$

STEP 4

Set $UB = 4.588$ $s_f = 1$ $\bar{\underset{\sim}{u}}^1 = \bar{\underset{\sim}{u}}$

$\underset{\sim}{c}^T = [1.\ 1.\ 1.]$ $(\underset{\sim}{y}_l + \underset{\sim}{y}_u)^T = [4.9\ 4.9\ 4.9]$

$(\underset{\sim}{x}^T \underset{\sim}{H})_1 = [x_1\ 3x_1\ 1.\ 0.\ 0.\ -1.\ 0.\ 0.]$

$(\underset{\sim}{x}^T \underset{\sim}{H})_2 = [0.\ 1.2x_2\ 0.\ 1.\ 0.\ 0.\ -1.\ 0.$

$(\underset{\sim}{x}^T \underset{\sim}{H})_3 = [x_3\ -x_3\ 0.\ 0.\ 1.\ 0.\ 0.\ -1.]$

$$(\underset{\sim}{b} - \underset{\sim}{g}^T \underset{\sim}{x}) = \begin{bmatrix} (-.1x_1 - .1x_3) \\ (10. - .1x_2 - .12x_2 + .1x_3) \\ 0. \\ 0. \\ 0. \\ -4.9 \\ -4.9 \\ -4.9 \end{bmatrix}$$

We have to add a constraint to the (RMP) such as

$$-.3 - [.057\ .457\ 0.\ .177\ 0.\ 0.\ 0.]\ [(-.1x_1 - .1x_3)$$

$$(10. - .3x_1 - .12x_2 + .1x_3)\ .\ 0.\ 0.\ -4.9\ -4.9\ -4.9]^T$$

$$+ [-1 + (.057\ .457\ 0.\ .177\ 0.\ 0.\ 0.\ 0.)$$

$$(x_1\ 3x_1\ 1.\ 0.\ 0.\ -1.\ 0.\ 0.)^T]^+\ 4.9$$

$$+ [-1 + (.057\ .457\ 0.\ .177\ 0.\ 0.\ 0.\ 0.)$$

$$(0.\ 1.2x_2\ 0.\ 1.\ 0.\ 0.\ -1.\ 0.)^T]^+\ 4.9$$

$$+ [-1 + (.057\ .457\ 0.\ .177\ 0.\ 0.\ 0.\ 0.)$$

$$(x_3\ -x_3\ 0.\ 0.\ 1.\ 0.\ 0.\ -1.)^T]^+\ 4.9$$

obtaining the following (RMP)

$$\text{Min } \eta$$

$$\underline{\text{st}}\ \eta - 4.87 - .04x_3 + .055x_2 + .143x_1 + [1.428x_1 - 1]^+ 4.9$$

$$+ [.548x_2 - .823]^+ 4.9 + [-.4x_3 - 1]^+ 4.9 \geq 0$$

$$0 \leq x_1 \leq 2.5$$

$$0 \leq x_2 \leq 2.5$$

$$-2.5 \leq x_3 \leq 0.$$

The first constraint makes the problem (RMP) a nonconvex problem. Each term of the type $[\ldots]^+$ can be linearized by introducing both a binary and a real variable and adding two constraints to the problem

Let

$$x_{4u} = 1.428\ x_{1u} - 1. \qquad x_{4l} = 1.428\ x_{4l} - 1.$$

$$x_{6u} = .548\ x_{2u} - .823 \qquad x_{6l} = .548x_{2l} - .823$$

$$x_{8u} = -.4x_{3u} - 1. \qquad x_{8l} = -.4x_{3l} - 1.$$

(RMP) is transformed into a 0-1 mixed linear program

$$\text{Min } \eta$$

$$\underline{\text{st}}\ \eta \geq 4.87 - .143\ x_1 - .055\ x_2 + .04x_3$$

$$-4.9\ x_4 - 4.9\ x_5 - 4.9\ x_6$$

$$x_4 \leq x_5\ x_{4u}$$

$$x_4 \leq (1 - x_5)\ x_{4l} + 1.428\ x_1 - 1.$$

$x_6 \leq x_7 \, x_{6u}$

$x_6 \leq (1-x_7) \, x_{6l} + .548 \, x_2 - .823$

$x_8 \leq x_9 \, x_{8u}$

$x_8 \leq (1-x_9) \, x_{8l} - .4 \, x_3 - 1.$

$\underset{\sim}{x} \geq \underset{\sim}{x}_l \qquad \underset{\sim}{x} \leq \underset{\sim}{x}_u$

x_5, x_7, x_9 are binary variables.

The optimal solution to the (RMP) is 1.868

LB = 1.868 $\quad \bar{\bar{x}}_1 = .5 \quad \bar{x}_2 = 2.5 \quad \bar{x}_3 = -2.5$

UB-LB $> \varepsilon$

STEP 3

Solve the dual of

$$\text{Min} \quad y_1 + y_2 + y_3 + .3$$

<u>st</u> $\quad .5 \, y_1 - 2.5 \, y_3 = .2$

$$1.5 \, y_1 + 3 \, y_2 + 2.5 \, y_3 = 9.3$$

The dual solution is finite and gives 3.60 corresponding to

$$\underset{\sim}{u}^T = [1. \ .333 \ 0. \ 0. \ 2.667 \ 0. \ 0. \ 0.]$$

The primal problem is therefore feasible.

Since $\quad 3.60 - \text{LB} > \varepsilon$

STEP 4

$3.60 < 4.59 = UB$. Make $UB = 3.60$

Add a constraint of the type (i) to (RMP) .

$s_f = 2$ and $\bar{\underset{\sim}{u}}^2 = \bar{\underset{\sim}{u}}$

STEP 2

Solve (RMP)

$$\text{Min } \eta$$

$$\text{st} \quad \eta \geq 4.87 - .143x_1 - .055x_2 + .04x_3 - [1.428x_1 - 1]^+ 4.9$$

$$- [.548x_2 - .823]^+ 4.9 - [-.4x_3 - 1]^+ 4.9$$

$$\eta \geq 3.33 - .2x_1 - .04x_2 - .067x_3 - [2x_1 - 1]^+ 4.9$$

$$- [.4x_2 - 1]^+ 4.9 - .66x_3 - 1.667]^+ 4.9$$

We obtain after linearization

(2) $\eta = 1.912 \quad \bar{x}_1 = .395 \quad \bar{x}_2 = 2.5 \quad \bar{x}_3 = -1.972$

$LB = 1.912 \qquad 3.6 - 1.912 > \varepsilon$

STEP 3

Solve the dual problem

$$\text{Max } .3 + .158\, u_1 + 9.384\, u_2 - 4.9\, u_6 - 4.9u_7 - 4.9u_8$$

$$\text{st} \quad .395u_1 + 1.185\, u_2 + u_3 - u_6 \leq 1$$

$$3\, u_2 + u_4 - u_7 \leq 1$$

$$-1.972\,u_1 + 1.972\,u_2 + u_5 - u_6 \leq 1$$

u_1, u_2 unrestricted and u_3 to $u_6 \geq 0$

The dual solution is 3.67

$$u^T = [1.537\ .333\ 0.\ 0.\ 3.364\ 0.\ 0.\ 0.]$$

$$v(\bar{x}) = 3.67 \qquad 3.67 - LB > \varepsilon$$

STEP 4

$$3.67 > UB$$

Let $s_f = 3$ and $\bar{u}^3 = \bar{u}$ in order to define the constraint to be added to the relaxed master problem

The algorithm would proceed. As we have mentioned η form a nondecreasing sequence of lower bounds while the upper bounds obtained from finite solutions to the dual problem need not decrease. In fact at iteration 5 the η obtained by solving (RMP) would give the smallest upper bound recorded. After three more iterations the algorithm came to an end since the value of the lower bound that coincide with η is higher than the incumbent upper bound . the final solution is 3.533 corresponding to the point

$$y_1 = 0.\ y_2 = 3.233\ y_3 = 0.\ x_1 = 0.\ x_2 = 2.5\ x_3 = 0.$$

By applying a translation inverse to the initial transformation we get $y_1 = .1 \quad y_2 = 3.33 \quad y_3 = .1$

5.10. Computational Experience

As in any implementation of Bender's method the computational performance of the procedure described above is highly dependent upon the structure of the relaxed master problem and upon the 0-1 mixed LP algorithm to solve it. We have used two alternative classes of methods for solving the ((RMP)). One is typified by the branch and bound algorithm aimed at solving the general mixed integer programming and uses LP as its main vehicle. Another is an implicit enumeration algorithm that solves an LP for each mutually exclusive set of 0-1 variables deleting from further consideration the combinations that would lead either to infeasibility or to a solution value that would not be able to improve the incumbent.

Our considerations will refer only to the former routine that was used in conjunction with the R-D algorithm in order to find the global solution of the small scale problem possessing multiple optima. Three new 0-1 variables the same number of real variables and seven constraints would have to be added after each iteration in order to replace the nonconvex terms that are a part of either support or cut functions. After 7 iterations and ellapsing 73 ieu the algorithm terminated giving the global solution stored as an upper bound. The same algorithm and same routine was used to attempt to solve the three bar truss subjected to two alternative loading conditions. The convergence turned out to be very slow and due to the number of variables and constraints generated by the sucessive cuts the method failed.

The same problem was sucessfully solved after only 6 iterations and using 12 ieu when the (RAP) algorithm was implemented in conjunction with Land and Powell's routine (1973) for mixed LP. We note that in this case only two binary variables and two real variables plus five constraints were added after each iteration. The implicit enumeration routine was used in the same problem consuming 36 ieu. If a smaller tolerance was set the (RAP) algorithm would proceed and the CPU would increase accordingly :

	6 iterations	7 iterations	10 iterations
Land & Powell	12 ieu	82 ieu	260 ieu
Implicit Enumeration	36 ieu	165 ieu	

The transformation of the 0-1 mixed linear programming into a piecewise concave minimization in real variables lead to an exponential growth in the CPU ellapsed and was aborted after three cuts

1st iteration	2nd iteration	3rd iteration
1 ieu	11 ieu	113 ieu

This seems to indicate that although the Land and Powell (1973) general purpose subroutine for mixed integer programming while sophisticated in its LP counterpart is more efficient it is time consuming and may not be used if the number of cuts is high. We also remark that early master problems have too little information to be worth optimizing very strictly. It takes several Wolsey's cuts in order to give accurate information concerning minimum volume design. This suggests the MP should be suboptimized particularly when s_f+t_f is small. There is a common contention that cutting plane based methods (or their extensions) have been a failure in practice. In fact there have been enough sucess enough untried avenues and sufficientely many new directions to ensure that the contention is disputable. The developments of methods more appropriate to the solution of the relaxed master problem deserve separate research of its own.

The study of complementarity in relation to mathematical programming has been a prolific source of research in the last fifteen years or so. The feasible region of the CPP is nonconvex and even disconnected and this causes complications when solving this problem. If the MP is cast in CPP format both branch and bound and cutting plane based methods could be employed. Balas (1975) may be cited exemplifying the school following the latter approach.

In general the idea of strengthening the LP relaxation of a Ø-1 mixed LP is thought to be very important for B & B based methods. However a straight forward implementation need not work well. In general a stronger LP is important. Cuts could be used in the OF in a Lagrangian fashion but the need for preserving the structure of the relaxed LP should be considered.

Bender's conventional decomposition procedure could be used to solve each relaxed master problem. This method has the advantage of splitting the size of the problem. Whenever the binary variables are held fixed the remaining LP has the number of variables and constraints considerably reduced. Geoffrion and Graves (1974) reported that an optimal solution could be found after a small number of cuts. One of the advantages of Bender's conventional approach to solve the RMP (or RAP) is that it offers the possibility of making sequences of related runs in much less computer time as compared with doing each run independently. The reoptimization capability is due to the fact that cuts devised to solve one problem can often be revised with little or no work so as to be valid in a modified version of the same problem.

5.11. B & B versus R-D

In many cases the solution of the structural synthesis problem with continuous variables found by using an algorithm which obtain local minimizers can be seen to be the global solution. Rather than an algorithm to solve the problem one should regard these strategies as verification procedures. Then its use which is longer in computer time than an algorithm which directly tries to obtain local solutions can be made greater or lesser by those who formulate the problem to be solved. The improvement of R-D based methods is highly dependent on the development of routines that will take advantage of the special nature of the 0-1 mixed LP relaxed master problem.

CHAPTER SIX

STRUCTURAL SYNTHESIS PROBLEMS WITH DISCRETE AREA VARIABLES

6.1. Problem Definition

Work in structural optimization with continuous system variables do not reflect well the real circumstances which practical engineers have to face. In order to produce a safe structure with minimum investment they may not be concerned with an indirect measure such as minimum volume. In addition to this daily design work consists of selection and allocation of appropriate cross sections from the catalogue supplied by the material makers and not detailed information on design variables. The minimum cost design is obtained simply by assigning an available cross-section to each independent structural member. If the fabrication cost is not included the objective function will only consists of the material needed to carry the loads.

Every available cross section will be assumed to have two independent quantities defining an element of the set Q_j.

(i) cross sectional area c_j

(ii) cost per unit length p_j

The original truss optimization problem would now have additional constraints

$$a_j \in Q_j = \{ Q_{1j}, \ldots Q_{r_j j} \}$$

$$= \{ (c_{1j}, p_{1j}), \ldots (c_{r_j j}, p_{r_j j}) \}$$

$$j=1,..n \qquad (6.1)$$

These constraints can be substituted by the equivalent formulation

$$a_j = \sum_{t=1}^{r_j} c_{tj} \delta_{tj} \qquad (6.2)$$

$$p_j = \sum_{t=1}^{r_j} p_{tj} \delta_{tj} \qquad (6.3)$$

$$\sum_{t=1}^{r_j} \delta_{tj} = 1 \qquad j=1,\ldots n \qquad (6.4)$$

where $\delta_{tj}=\emptyset$ or 1 ($\in$ B) and r_j is the number of sections available for the current member from the set Q_j. The constraints representing bounds on the design variables would no longer be required.

In order to reflect the minimum cost criteria the OF would become

$$\text{Min} \quad \underset{\sim}{1}^T \underset{\sim}{p} \qquad (6.5)$$

ie: a linear function of the cost of each section. The optimal truss design with fixed topology and discrete variables is a Ø-1 mixed bilinearly constrained problem. It is a problem having a set of disjoint domains defined for each set of δ_{tj} and is therefore liable to possess multiple optima. The only bilinear equations of the system are the equilibrium relations. The Ø-1 variables will appear as linear constraints. Fixing δ_{tj} $(t=1,\ldots r, j=1,\ldots n)$ each

a_j and p_j can be determined and the bilinear equations would become linearly dependent upon the state variables. The resulting problem is therefore an LP. In order to avoid an enumeration of all possible solutions some strategies will be developed in the forthcoming sections that will lead to the global solution of the original problem.

But first we will develop the standard formulation in order to find out some special properties the problem may have. Due to its special uncoupling nature it can be transformed into another solvable by a standard programming code. From

$$a_j = \sum_{t=1}^{r_j} c_{tj} \delta_{tj} \quad \text{and} \quad \sum_{t=1}^{r_j} \delta_{tj} = 1 \qquad (6.6)$$

where $\delta_{tj} \in B = \{ \emptyset, 1 \}$

a typical product term of the equilibrium equations $a_j s_j^k$ can be substituted by

$$\left(\sum_{t=1}^{r_j} c_{tj} \delta_{tj} \right) s_j^k \qquad (6.7)$$

Each term of this sum can be replaced by its convex envelope. Under each loading condition the bars of the structure will be submitted to either compression or tension varying along one of the half-spaces. Assuming the stresses are nonnegative the convex envelope over the rectangle of bounds

$$\emptyset \leq \delta_{tj} \leq 1 \qquad \emptyset \leq s_j^k \leq s_{ju}^k \qquad (6.8)$$

is the maximum of the two expressions

$$V_{tj}^{k} = \text{Max}\, \{ 0, s_{ju}^{k} \delta_{tj} + s_{j}^{k} - s_{ju}^{k} \}$$

$$\leq \delta_{tj} s_{j}^{k} \qquad (6.9)$$

or

$$V_{tj}^{k} \geq 0$$

$$V_{tj}^{k} \geq s_{ju}^{k} \delta_{tj} + s_{j}^{k} - s_{ju}^{k} \qquad (6.10)$$

Since the variable δ_{tj} will only take binary values and the convex envelope of V_{tj}^{k} will coincide with the product term $\delta_{tj} s_{j}^{k}$ at its endpoints. When the convex underestimate of the symmetric of the product function is required (ie: concave overestimate of the product function) in the rectangle

$$0 \leq \delta_{tj} \leq 1 \qquad -s_{ju}^{k} \leq -s_{j}^{k} \leq 0 \qquad (6.11)$$

We have

$$\bar{V}_{tj}^{k} \geq -s_{j}^{k} \qquad (6.12)$$

$$\bar{V}_{tj}^{k} \geq -\delta_{tj} s_{ju}^{k} \qquad (6.13)$$

where $\delta_{tj} \in B = \{0, 1\}$

By analogy we have

$$\bar{V}_{tj}^{k} = -\delta_{tj} s_{j}^{k} = -V_{tj}^{k} \qquad (6.14)$$

so that $\bar{V}_{tj}^{k}$ has a value symmetric to the variable previously considered V_{tj}^{k}

The nonnegativity requirement for the variables is a standard feature of most LP codes so that the product term $\delta_{tj} s_{j}^{k}$ can be substituted by the variable V_{tj}^{k} and the constraint

$$V_{tj}^{k} \geq s_{ju}^{k}\, \delta_{tj} + s_{j}^{k} - s_{ju}^{k} \qquad (6.15)$$

By symmetry the convex envelope on the rectangle of bounds

$$\emptyset \leq \delta_{tj} \leq 1 \qquad s_{jl}^{k} \leq s_{j}^{k} \leq \emptyset \qquad (6.16)$$

for the factorable term $\delta_{tj} s_{j}^{k}$ is given by

$$V_{tj}^{k} \geq s_{j}^{k} \qquad (6.17)$$

$$V_{tj}^{k} \geq \delta_{tj}\, s_{jl}^{k} \qquad (6.18)$$

But this new variable can be related with the nonnegative variable

$$\emptyset \leq \delta_{tj} \leq 1 \qquad \emptyset \leq -s_{j}^{k} \leq -s_{jl}^{k} \qquad (6.19)$$

$$\underline{V}_{tj}^{k} \geq s_{jl}^{k} - s_{j}^{k} - \delta_{tj}\, s_{jl}^{k} \qquad (6.2\emptyset)$$

$$\underline{V}_{tj}^{k} \geq \emptyset \qquad (6.21)$$

Thus

$$\underline{V}_{tj}^{k} = -V_{tj}^{k} \qquad (6.22)$$

Each term of the equilibrium equation can be written

$$a_{j}\, s_{j}^{k} = \sum_{t=1}^{r_j} c_{tj}\, \delta_{tj} s_{j}^{k} = \sum_{t=1}^{r_j} c_{tj}\, V_{tj}^{k} \qquad (6.23)$$

Assembling for the whole equation

$$\sum_{j=1}^{n} h_{ij}\, a_{j}\, s_{j}^{k} = \sum_{j=1}^{n} \sum_{t=1}^{r_j} h_{ij}\, c_{tj}\, V_{tj}^{k} = \lambda_{i}^{k} \qquad (6.24)$$

that is a linear expression on the variables V_{tj}^{k} $j=1,\ldots n; k=1,\ldots l$; $i=1,\ldots \beta$

Therefore the initial problem can be transformed into another of increased dimensionality [(6.25)-(6.34)] in which the equations are linear and a set of V^k_{tj} variables will substitute every area variable. All the functions involved in this problem are linear. Since some of the variables are binary this problem falls into the class of 0-1 mixed LP. It is of no practical use because of its size. In subsequent sections methods that also solve this problem will be described and a comparison will be drawn between both formulations.

$$\text{Min} \sum_{j=1}^{n} l_j \sum_{t=1}^{r_j} p_{tj} \delta_{tj} \tag{6.25}$$

$$\underline{\text{st}} \quad \sum_{t=1}^{r_j} {}_{tj} = 1 \tag{6.26}$$

$$\sum_{j=1}^{n} h_{ij} \sum_{t=1}^{r_j} c_{tj} V^k_{tj} = \lambda^k_i \tag{6.27}$$

$$\underset{\sim}{B}^T \underset{\sim}{L} \underset{\sim}{s}^k = 0 \tag{6.28}$$

$$\underset{\sim}{s}^k_l \le \underset{\sim}{s}^k \le \underset{\sim}{s}^k_u \tag{6.29}$$

$$\underset{\sim}{d}^k_l \le \underset{\sim}{D} \underset{\sim}{s}^k \le \underset{\sim}{d}^k_u \tag{6.30}$$

For

$$0 \le s^k_j \le s^k_{ju} \tag{6.31}$$

$$V^k_{tj} \ge s^k_{ju} \delta_{tj} + s^k_j - s^k_{ju} ; V^k_{tj} \ge 0 \tag{6.32}$$

For

$$s^k_{jl} \le s^k_j \le 0 \tag{6.33}$$

$$V^k_{tj} \le \delta_{tj} s^k_{jl} + s^k_j - s^k_{jl} ; V^k_{tj} \le 0 \tag{6.34}$$

and

$$\delta_{tj} \in B = \{ 0, 1 \}$$

Some simplifications to the original problem will occur when the areas are only allowed to take integer values with constant step size and the cost is considered proportional to the cross section. Supposing the step size unitary (otherwise the transformation of the area variables to the ratio of these variables by the step size yields new variables differing by an integer along the feasible domain) the synthesis problem for discrete variables is in every way the same as for continuous variables but with the additional constraint that the $\underset{\sim}{a}$ variables should be integer.

$$\text{Min} \quad \underset{\sim}{1}^T \underset{\sim}{a} \qquad (6.35)$$

$$\underline{\text{st}} \quad \underset{\sim}{a}^T \underset{\sim}{H}_i \underset{\sim}{s}^k = \lambda_i^k \qquad i=1,\ldots\beta \qquad (6.36)$$

$$\underset{\sim}{B}^T \underset{\sim}{L} \underset{\sim}{s}^k = 0 \qquad (6.37)$$

$$\underset{\sim}{a}_l \leq \underset{\sim}{a} \leq \underset{\sim}{a}_u \qquad (6.38)$$

$$\underset{\sim}{s}_l^k \leq \underset{\sim}{s}^k \leq \underset{\sim}{s}_u^k \qquad (6.39)$$

$$\underset{\sim}{d}_l^k \leq \underset{\sim}{d}^k \leq \underset{\sim}{d}_u^k \qquad (6.40)$$

$\underset{\sim}{a}$ is integer

The discrete nature of the set of design variables compels the existence of multiple optima solutions. Any discrete set of areas producing a feasible design with respect to the stress resultant vector is a solution locally optimal. It is more reasonable in the light of the current state of the knowledge to adopt methods for the solution of nonconvex programming than to adopt alternative stance and commit considerable time only with a minimal expectation of success.

6.2. Branch And Bound Methods

6.2.1. Sequence of LP

The Branch and Bound procedure previously developed based on Soland's algorithm for separable problems can be applied to this problem. Now the design variables can only take a finite number of values and it is possible to use as an additional information the knowledge that no such variable will have a value not coinciding with any available section at the optimal solution. The splitting rule is therefore modified to acommodate this fact. The design variables are eligible to define the interval that will be partitioned and instead of the solution coordinate the available section immediately above and below this real value will be used to define the new lower and upper bounds corresponding to the two branches. A test on the integrality on all design variables has to be introduced to perclude noninteger solutions.

6.2.2. Sequence of LIP

When the available sections differ by a constant step size the minimum volume design becomes subjected to integrality requirements in the design variables and possess bilinear constraints joining state and design variables. The algorithm mentioned above can be used again to determine the global solution of this nonconvex problem. If we enforce the integrality requirements on the areas a sequence of LIP instead of LP has to be soved. The advantage of using a more sophisticated LIP code lies in the fact that the solutions of the problem would be brought closer to the set of feasible solutions to the original problem and less problems with greater difficulty woud have to be solved.

6.2.3 Analogy with an algorithm for piecewise convex functions

The mathematical problem of finding the minimum of a separable piecewise convex function has been solved by Soland's B & B algorithm. The global optimum is obtained by solving a finite sequence of convex programming problems corresponding to sucessive partitions of the set defined by the bounds on the variables on the basis of piecewise convexity of the problem functions. The algorithm is almost identical to those proposed for solving problems with nonconvex constraints but in the latter case the convergence can only be established at the limit of an infinite sequence. The interval of definition of each variable is therefore partitioned into a finite number of subintervals over each of which all the functions are convex. They will form a disjoint union of the initial interval. When each variable is restricted to a particular partition the constraint set will become convex and so does the OF. Thus any local solution of each subproblem is also the global solution over that partition.

In order to avoid solving all such problems explicitly a B & B strategy is employed. The algorithm considers subsets of the feasible solution of the original problem. Find lower bounds on the optimal solution values in such subsets by solving convex programming problems and identifies an optimal solution. The piecewise convex problem is relaxed by using convex underestimates of the original functions and a lower bound on the optimal solution is obtained. If this relaxed solution is not feasible in the original problem the index of the variable that presents the maximum difference between the original function and its piecewise underestimate out of the violated constraints, is picked up in much the same way as previously described. The interval of variation is now ready to be partitioned. In the next stage a number of subproblems corresponding to the number of subsets splitting the interval corresponding to the variable chosen has to be solved (as opposed to the branching in two new nodes that is a feature of all B & B algorithms previously mentioned) .

The algorithm is finite. By the nth. level of the tree where n represents the number of variables the lower bound determined for an intermediate node will equal or exceed the solution vector found at that node. If the application to stuctural synthesis problems is considered, we have to bear in mind that each interval corresponding to a design variable can be partitioned into a finite number of subsets (available sections) over each of which it is possible to find a convex envelope to the original problem. The convex underestimating subproblem (that could employ the factorable envelopes referred) can be solved by a LP code. Each node would then be branched into n new nodes where n would represent the set of available sections of the selected design variable. This corresponds to a stronger branching rule than the procedure described in 4.2.

6.3. Resource-Decomposition Methods

The acceleration algorithm (RAP) described in the previous Chapter seems particularly suited to the structural synthesis problems with discrete design variables. (RAP) becomes either a Ø-1 mixed linear programming or a Ø-1 mixed linear integer programming for the general Ø-1 mixed bilinearly constrained programming and the simplified version where the available sizes differ by a constant step size respectively. The injective relationship between areas and their stress/displacement resultant will be used again to define a design feasible with respect to the state variable space. The dual variables $\tilde{u}$ obtained from the same matrix used to find the state variables $\tilde{s}$ are appropriately scaled and used to define the constraints added at each iteration.

The discrete nature of the problem makes the scaled problem generally infeasible in the design variables space. The upper bounding stopping criteria is now meaningless. By simple variable substitution (RAP) becomes

$$\text{Min} \sum_{j=1}^{n} l_j \; p_j = \sum_{j=1}^{n} l_j \sum_{t=1}^{r_j} p_{tj} \, \delta_{tj} \tag{6.41}$$

$$\underline{\text{st}} \quad \sum_{t=1}^{r_j} \delta_{tj} = 1 \qquad i=1,\ldots m \tag{6.42}$$

$$a_j = \sum_{t=1}^{r_j} c_{tj} \, \delta_{tj} \tag{6.43}$$

$$\underset{\sim}{a}^T \underset{\sim}{H}_i \underset{\sim}{s}^k = \lambda_i^k \qquad j=1,\ldots n \tag{6.44}$$

$$\underset{\sim}{d}_l^k \le \underset{\sim}{D} \, \underset{\sim}{s}^k \le \underset{\sim}{d}_u^k \tag{6.45}$$

$$\underset{\sim}{s}_l^k \le \underset{\sim}{s}^k \le \underset{\sim}{s}_u^k \tag{6.46}$$

where $\delta_{tj} \in B = \{ \emptyset, 1 \}$ which

can be reformulated as the following relaxed master program (RAD)

$$\text{Min} \quad \eta = \underset{\sim}{1}^T \underset{\sim}{p} \tag{6.47}$$

$$\underline{\text{st}} \quad \sum_{i=1}^{m} u_i^e \lambda_i^k + \sum_{j=1}^{n} \Big\{ [-\sum_{i=1}^{m} u_i^e \, (\underset{\sim}{a}^T \underset{\sim}{H}_i + \underset{\sim}{D})]_j$$

$$[s_{jl}^k + s_{ju}^k] \Big\}^+ \ge \emptyset \tag{6.48}$$

$$a_j = \sum_{t=1}^{r_j} c_{tj} \, \delta_{tj} \tag{6.49}$$

$$p_j = \sum_{t=1}^{r_j} p_{tj} \, \delta_{tj} \tag{6.50}$$

$$\sum_{t=1}^{r_j} \delta_{tj} = 1 \qquad j=1,\ldots n \qquad e=1,\ldots e_f \tag{6.51}$$

By using $\emptyset,1$ varibles δ_{je} each of the terms $[\ldots]^+$ can be linearized.

Supposing

$$L_{je} \leq \left\{ [-\sum_{i=1}^{m} u_i^e (\underset{\sim}{a}^T \underset{\sim}{H}_i + \underset{\sim}{D})] \quad [s_{jl}^k + s_{ju}^k] \right\} \leq U_{je} \qquad (6.52)$$

$$V_{je} \leq \delta_{je} V_{je} \qquad (6.53)$$

$$V_{je} \leq (\delta_{je} - 1) L_{je} + \left\{ [-\sum_{i=1}^{m} u_i^e (\underset{\sim}{a}^T \underset{\sim}{H}_i + \underset{\sim}{D})_j \right.$$

$$\left. (\underset{\sim}{s}_l^k + \underset{\sim}{s}_u^k)]_j \right\} \qquad (6.54)$$

The solution of the problem having discrete member sizes can be obtained by solving a sequence of Ø-1 mixed LP each creating a cut function until the optimal solution is obtained. We remark that the discrete nature of the cross sections makes (RAD) more efficient than (RAP) due to the sluggishness of the algorithm for continuous variables near a feasible solution. The problem of higher dimensionality mentioned in 6.1. was obtained directly from the Ø-1 mixed bilinearly constrained problem. Although it has a similar structure to (RAD) both problems are generated in a different way. In the former the convex envelope of the factorable bilinear terms was taken to rewrite the constraints of the primal problem; the optimal solution of that Ø-1 mixed LP would be the optimal solution of the problem with discrete member sizes since the envelope coincides with the value of the function at the boundaries.

When the member sizes differ by a constant step size and the cost is proportional to the cross sectional area the (RAD) will be defined as a Ø-1 mixed linear integer programming (RIM)

$$\text{Min } \eta = \underset{\sim}{1}^T \underset{\sim}{a} \tag{6.55}$$

$$\underline{\text{st}} \quad \sum_{i=1}^{m} u_i^e \, {}_i^k + \sum_{j=1}^{n} \Big\{ [-\sum_{i=1}^{m} u_i^e (\underset{\sim}{a}^T \underset{\sim}{H}_i + \underset{\sim}{D})_j] \; [s_{jl}^k + s_{ju}^k] \Big\}^+ \geq \emptyset \tag{6.56}$$

$$\underset{\sim}{a}_l \leq \underset{\sim}{a} \leq \underset{\sim}{a}_u \qquad e = 1, \dots e_f \tag{6.57}$$

Some results conce rning the computational implementation of (RIM) will be given latter in this Chapter.

6.4 Applications

6.4.1. Resource-decomposition algorithm -accelerated convergence [Appendix B]

The ultimate goal of any theory is its application to practical problems. A three bar truss subject to stress constraints integrality requirements on the member areas and undergoing two alternative loading conditions is solved in this section by using a generalization of Bender's algorithm . The nodal stiffness formulation of this problem is as follows

$$\min \sqrt{2}\, a_1 + a_2 + \sqrt{2}\, a_3$$

<u>st</u>

$$\sqrt{2}/2\, a_1 d_1^1 + 1/2\, a_2 (d_1^1 + d_2^1) = 40.$$

$$1/2\, a_2 (d_1^1 + d_2^1) + \sqrt{2}/2\, a_3 d_2^1 = 0$$

$$\sqrt{2}/2\, a_1 d_1^2 + 1/2\, a_2 (d_1^2 + d_2^2) = 0.$$

$$1/2\, a_2 (d_1^2 + d_2^2) + \sqrt{2}/2\, a_3 d_2^2 = 20.$$

$$0 \leq \sqrt{2}/2\, d_1^1 \leq 5.$$

$$-5. \leq \sqrt{2}/2\, d_1^2 \leq 0.$$

$$0 \leq \sqrt{2}/2\, (d_1^1 + d_2^1) \leq 5.$$

$$0 \leq \sqrt{2}/2\, (d_1^2 + d_2^2) \leq 5.$$

$$-5. \leq \sqrt{2}/2\, d_2^1 \leq 0.$$

$$0 \leq \sqrt{2}/2\, d_2^2 \leq 5.$$

The equilibrium equations can be written in the matrix form as

$$\underset{\sim}{K}\, \underset{\sim}{d}^k = \underset{\sim}{\lambda}^k \qquad k=1,2$$

where $\underset{\sim}{K}$ is the assembled stiffness matrix of the structure. By differentiating the Lagrangian of this problem with respect to the displacement variables we get the following set of equations

$$(\sqrt{2}/2\, a_1 + 1/2\, a_2)\, \gamma_1^1 + 1/2\, a_2 \gamma_2^1 - \sqrt{2}/2\, \mu_1^1 - \sqrt{2}/2\, \mu_2^1 = 0$$

$$1/2\, a_2\, \gamma_1^1 + (\sqrt{2}/2\, a_3 + 1/2\, a_2)\, \gamma_2^1 - \sqrt{2}/2\, \mu_2^1 + \sqrt{2}/2\, \mu_3^1 = 0$$

$$(\sqrt{2}/2\, a_1 + 1/2\, a_2)\, \gamma_1^2 + 1/2\, a_2 \gamma_2^2 + \sqrt{2}/2\, \mu_1^2 - \sqrt{2}/2\, \mu_2^2 = 0$$

$$1/2\ a_2 \gamma_1^2 + (\sqrt{2}/2\ a_3 + 1/2\ a_2)\ \gamma_2^2 - \sqrt{2}/2\ \mu_2^2 - \sqrt{2}/2\ \mu_3^2 = 0$$

and γ_i^k, μ_i^k are Lagrange multipliers refering to bilinear and stress limit constraints respectively. These equations can also be written as

$$\underset{\sim}{K}\ \underset{\sim}{\gamma}^k = \underset{\sim}{\mu}^k$$

If the set of areas give a feasible design in the stress space we would have

$$\sum_{k=1}^{l} \left(\underset{\sim}{\lambda}^{k^T} \underset{\sim}{\gamma}^k + \underset{\sim}{s}^{k^T} \underset{\sim}{\mu}^k \right) = 0.$$

Otherwise in order to obtain a dual ray one may either scale the Lagrange multipliers or assume a fixed value for μ^k corresponding to the stress that is further apart from the rectangle of bounds. In the latter case the vector γ^k is uniquely determined eg: supposing that the design has member 2 fully stressed under loading condition 1 we have

$$\mu^1 = [0\ \mu_2\ 0]^T\ ;\ \mu^2 = [0\ 0\ 0]^T;\ \mu_2 > 0$$

The relaxed master program (RIM) may be restated here for convenience

$$\text{Min} \quad \eta = \underset{\sim}{1}^T \underset{\sim}{x} \tag{6.58}$$

$$\underline{\text{st}} \quad -\sum_{i=1}^{m} u_i^e b_i + \sum_{j=1}^{n} \left\{ \left[\sum_{i=1}^{m} u_i^e (\underset{\sim}{x}^T \underset{\sim}{H}_i + f_i)_j \right. \right.$$

$$\left. \left. (y_{jl} + y_{ju}) \right] \right\}^+ \geq 0 \tag{6.59}$$

$$\underset{\sim}{x} \in X \qquad e = 1, \ldots e_f \tag{6.60}$$

Let $x_1 = a_1; x_2 = a_2; x_3 = a_3$ represent the design variables

$y_1 = d_1^1 ; y_2 = d_2^1 ; y_3 = d_1^2 ; y_4 = d_2^2$

INITIALIZATION

STEP 1

Let UB $= +\infty$; LB = 3.828 corresponding to

$x_1 = 1 \quad x_2 = 1 \quad x_3 = 1$

and ε = .000 (since we will end up with a exact solution).

$$\underset{\sim}{K} = \begin{bmatrix} 1.201 & .5 \\ .5 & 1.201 \end{bmatrix} ; \underset{\sim}{K}^{-1} = \begin{bmatrix} 1.00 & -.414 \\ -.414 & 1.00 \end{bmatrix}$$

$$\underset{\sim}{d}^k = \underset{\sim}{K}^{-1} \underset{\sim}{\lambda}^k \qquad k=1,2$$

$$\underset{\sim}{d}^1 = [40. \ -16.569]^T ; \underset{\sim}{d}^2 = [-8.282 \ 20.]^T$$

In order to find a feasible set of nodal displacements the scaling factor

$$\rho = \max \left\{ 40./5\sqrt{2}, 24.43/5\sqrt{2}, -16.57/-5\sqrt{2}, -8.28/-5\sqrt{2}, 11.72/5\sqrt{2}, 20/5\sqrt{2} \right\}$$

$$=5.657$$

would increase the design variables until the set of stresses would fit within their bounds. The consideration of this scaling factor corresponds to having member 1 at its upper bound (fully stressed) under loading condition 1.

Let

$$\tilde{\mu}^1 = [\sqrt{2}\ 0\ 0]^T \qquad \tilde{\mu}^2 = [0\ 0\ 0]^T$$

We get

$$\tilde{\gamma}^1 = [1.\ -.414]^T \qquad \tilde{\gamma}^2 = [0\ 0]^T$$

$e_f=1$

$$\tilde{u}^{1^T} = [1.\ -.414\ 0\ 0\ \sqrt{2}\ 0\ 0\ 0\ 0\ 0]$$

$$\tilde{b}^T = [40.\ 0.\ 0.\ 20.\ -5.\ -5.\ 5.\ 5.\ -5.\ -5.]$$

$$\tilde{x}^T \tilde{H}_1 = [(\sqrt{2}/2\ x_1 + 1/2\ x_2)\ 1/2\ x_2\ 0\ 0]$$

$$\tilde{x}^T \tilde{H}_2 = [\ 1/2\ x_2\ (\sqrt{2}/2\ x_3 + 1/2\ x_2)\ 0.\ 0.]$$

$$\tilde{x}^T \tilde{H}_3 = [\ 0.\ 0.\ (\sqrt{2}/2\ x_1 + 1/2\ x_2)\ 1/2\ x_2]$$

$$\tilde{x}^T \tilde{H}_4 = [\ 0.\ 0.\ 1/2\ x_2\ (\sqrt{2}/2\ x_3 + 1/2\ x_2)]$$

$$\tilde{f}_5 = [-\sqrt{2}/2\ 0.\ 0.\ 0.]$$

$$\tilde{f}_6 = [-\sqrt{2}/2\ -\sqrt{2}/2\ 0.\ 0.]$$

$$\tilde{f}_7 = [0.\ -\sqrt{2}/2\ 0.\ 0.]$$

$$\tilde{f}_8 = [0.\ 0.\ -\sqrt{2}/2\ 0.]$$

$$\tilde{f}_9 = [0.\ 0.\ -\sqrt{2}/2\ -\sqrt{2}/2]$$

$$\tilde{f}_{10} = [0.\ 0.\ 0.\ -\sqrt{2}/2]$$

$$y_{1l} + y_{1u} = 5\sqrt{2} \qquad y_{2l} + y_{2u} = -5\sqrt{2}$$

$$y_{3l} + y_{3u} = -5\sqrt{2} \qquad y_{4l} + y_{4u} = 5\sqrt{2}$$

The constraint that cuts off the initial solution point is

$$-40. + 5\sqrt{2} + [\sqrt{2}/2\ x_1 + 1/2\ x_2$$

$$- .414 \times 1/2\ x_2 - \sqrt{2} \times \sqrt{2}/2]^{+} 5\sqrt{2} + [1/2\ a_2$$

$$- .414\ (\sqrt{2}/2\ x_3 + 1/2\ x_2)]^{+} [-5\sqrt{2}] \geq 0$$

or

$$-32.929 + (5\ x_1 + 2.071\ x_2 - 7.071)^{+}$$

$$+ (-2.071\ x_2 + 2.071\ x_3)^{+} \geq 0$$

The nonnegative terms are linearized by introducing two variables per term . We have

$$x_{4u} = 5x_{1u} + 2.071x_{2u} - 7.071 \quad ; x_{4l} = 5\ x_{1l} + 2.071\ x_{2l} - 7.071$$

$$x_{6u} = -2.071x_{1l} + 2.071x_{3u} ; x_{6l} = -2.071\ x_{1u} + 2.071\ x_{3l}$$

This new constraint can be represented by

$$x_4 + x_6 \geq 32.929$$

$$x_4 - x_5\ x_{4u} \leq 0$$

$$x_4 - x_5\ x_{4u} - 5\ x_1 - 2.071\ x_2 \leq 7.071 - x_{41}$$

$$x_6 - x_7\ x_{6u} \leq 0$$

$$x_6 - x_7\ x_{61} + 2.071\ x_1 - 2.071\ x_2 \leq -x_{61}$$

where x_5, x_7 are binary and x_4, x_6 are real variables. The nonconvex constraint is therefore replaced by a 4 new variables and 5 linear constraints

STEP 3 IT = 1 (2)

The optimal solution of (RIM) is 13.728

LB =13.728 UB-LB > ε

$\bar{x}_1 = 8$; $\bar{x}_2 = 1$; $\bar{x}_3 = 1$

STEP 2 (2)

$$\underset{\sim}{K} = \begin{bmatrix} 6.157 & .5 \\ .5 & 1.207 \end{bmatrix} \quad \underset{\sim}{K}^{-1} = \begin{bmatrix} .168 & -.070 \\ -.070 & .857 \end{bmatrix}$$

$$\underset{\sim}{d}^1 = [6.724 \ -2.784]^T; \ \underset{\sim}{d}^2 = [-1.392 \ 17.148]^T$$

so that ρ = 2.42 would correspond to a displacement d_2^2 at its upper bound. We have

$e_f = 2$

$$\underset{\sim}{\bar{u}} = [0. \ 0. \ -.07 \ .857 \ 0. \ 0. \ 0. \ 0. \ 0. \ 2]^T$$

The new constraint is given by

$$20. \ (-.857) + 5\sqrt{2} + [-.07(1/2x_2 + \sqrt{2}/2x_1)$$

$$+ \ .857 \ 1/2x_2] [-5\sqrt{2}]^+ + [-.07$$

$$1/2x_2 + .857(1/2x_2 + \sqrt{2}/2x_3)$$

$$- \ 2 \ \sqrt{2}/2] \ [5\sqrt{2}]^+ \geq 0$$

or

$$-10.069 + [.35 \ x_1 - 2.784x_2]^+$$

$$+ [4.235x_3 - 2.784x_2 - 7.071]^+ \geq 0$$

STEP 3 IT = 2 (2)

$(\bar{\eta}, \bar{x}) = (16.142, 7, 2, 3)$ LB = 16.142

STEP 2 (2)

$\rho = 1.005$ corresponds to an upper bound on d_1^1.

$e_f = 3$ and the new constraint that is added to RIM is

$$-.048 + [.84\ x_1 + .427\ x_2 - 7.071]^+$$
$$+ [-.427\ x_2 + .285\ x_3]^+ \geq 0$$

STEP 3 IT = 3 (2)

$(\bar{\eta}, \bar{x}) = (16.556, 6, 1, 5)$ LB = 16.556

STEP 2

$\rho = 1.2086$ will correspond to an upper bound on d_1^1. In order to cut off this point we will add a 4th constraint ($e_f = 4$)

$$-1.4289 + [1.07\ x_1 + .665\ x_2 - 7.071]^+$$
$$+ [-.665\ x_2 + .13\ x_3]^+ \geq 0$$

STEP 3 IT = 4 (2)

$(\bar{\eta}, \bar{x}) = (16.728, 7, 4, 2)$; LB = 16.728

STEP 2 (1)

$\rho = .996 < 1$ TERMINATE

$$\underset{\sim}{d}^{k} = \underset{\sim}{K}^{-1} \underset{\sim}{\lambda}^{k} \qquad k=1,2$$

$$a_1 = x_1 = 7 \; ; a_2 = x_2 = 4 \; ; a_3 = x_3 = 2$$

is the optimal volume design with integer design variables.

We observe that some of the constraints used in the early stages of the algorithm will not be active at the optimal point since the previous solutions were creating constraints requiring a volume higher each time. Although nonactive constraints could have been dispensed with it is not possible to know in advance which will be needed to find the correct solution to the relaxed master program.

6.4.2. Branch and bound trees

For the same problem we have taken as convex envelopes of the bilinear expressions the underestimates of the product terms. The problem is defined in stress and member area variables in order to minimize the number of functions that are bilinear. Each node is again defined by the upper and lower bounds in both state and design variables. These bounds are required to determine both the structural matrix and the RHS of the linear constraints that approximate the nonconvex terms. By taking underestimates we are penalized by an increase in the number of constraints and variables enlarging the subproblem representing any node. The three bar truss subjected to two alternative loading conditions and bounds on the stress and area variables is a problem in 9 variables restricted by 4 bilinear constraints and two linear equalities. It is transformed into a problem in 13 variables and subject to 24 linear inequalities plus two linear equalities that are not changed.

$$\text{Min } \sqrt{2}\, a_1 + a_2 + \sqrt{2}\, a_3$$

$$\underline{\text{st}} \quad a_1 s_1^1 + \sqrt{2}/2\, a_2 s_2^1 = 40.$$

$$\sqrt{2}/2\, a_2^1 s_2 + a_3 s_3^1 = 0.$$

$$a_1 s_1^2 + \sqrt{2}/2\, a_2 s_2^2 = 0.$$

$$\sqrt{2}/2\, a_2 s_2^2 + a_3 s_3^2 = 20.$$

$$-s_1^1 + s_2^1 - s_3^1 = 0.$$

$$-s_1^2 + s_2^2 - s_3^3 = 0.$$

$$\underset{\sim}{a}_l = [1.\ 1.\ 1.]^T \qquad \underset{\sim}{a}_u = [11.\ 4.\ 5.]^T$$

$$\underset{\sim}{s}_l^1 = [0.\ 0.\ -5.]^T \qquad \underset{\sim}{s}_u^1 = [5.\ 5.\ 0.]^T$$

$$\underset{\sim}{s}_l^2 = [-5.\ 0.\ 0.]^T \qquad \underset{\sim}{s}_u^2 = [0.\ 5.\ 5.]^T$$

Let

$x_1 = a_1$; $x_2 = a_2$; $x_3 = a_3$

$x_4 = s_1^1 \quad x_5 = s_2^1 \quad x_6 = s_3^1 \quad x_7 = s_1^2 \quad x_8 = s_2^2 \quad x_9 = s_3^2$

The linear problem obtained when using the concept of factorable functions is given by

$$\text{Min } \sqrt{2}\, x_1 + x_2 + \sqrt{2}\, x_3$$

$$\underline{\text{st}} \quad x_{10} \geq x_2 x_{5u} + x_5 x_{2u} - x_{2u} x_{5u}$$

$$x_{10} \geq x_2 x_{5l} + x_5 x_{2l} - x_{2l} x_{5l}$$

$$x_{11} \geq -x_2 x_{5u} - x_5 x_{2l} + x_{5u} x_{2l}$$

$$x_{11} \geq -x_2 x_{5l} - x_5 x_{2l} + x_{2u} x_{5l}$$

$$x_1 x_{4u} + x_4 x_{1u} - x_{1u} x_{4u} + \sqrt{2}/2\ x_{10} \leq 40.$$

$$x_1 x_{4l} + x_4 x_{1l} - x_{1l} x_{4l} + \sqrt{2}/2\ x_{10} \leq 40.$$

$$-x_1 x_{4u} - x_4 x_{1l} + x_{4u} x_{1l} + \sqrt{2}/2\ x_{11} \leq -40.$$

$$-x_1 x_{4l} - x_4 x_{1u} + x_{1u} x_{4l} + \sqrt{2}/2\ x_{11} \leq -40.$$

$$x_{12} \geq x_2 x_{8u} + x_8 x_{2u} - x_{2u} x_{8u}$$

$$x_{12} \geq x_2 x_{8l} + x_8 x_{2l} - x_{2l} x_{8l}$$

$$x_{13} \geq -x_2 x_{8u} - x_8 x_{2l} + x_{8u} x_{2l}$$

$$x_{13} \geq -x_2 x_{8l} - x_8 x_{2u} + x_{2u} x_{8l}$$

$$\sqrt{2}/2\ x_{10} + x_3 x_{6u} + x_6 x_{3u} - x_{3u} x_{6u} \leq 0.$$

$$\sqrt{2}/2\ x_{10} + x_3 x_{6l} + x_6 x_{3l} - x_{3l} x_{6l} \leq 0.$$

$$\sqrt{2}/2\ x_{11} - x_3 x_{6u} - x_6 x_{3l} + x_{3l} x_{6u} \leq 0.$$

$$\sqrt{2}/2\ x_{11} - x_3 x_{6l} - x_6 x_{3u} + x_{3u} x_{6l} \leq 0.$$

$$x_7 x_{1u} + x_1 x_{7u} - x_{1u} x_{7u} + \sqrt{2}/2\ x_{12} \leq 0.$$

$$x_7 x_{1l} + x_1 x_{7l} - x_{1l} x_{7l} + \sqrt{2}/2\ x_{12} \leq 0.$$

$$-x_7 x_{1u} - x_1 x_{7l} + x_{1u} x_{7l} + \sqrt{2}/2\ x_{13} \leq 0.$$

$$-x_7 x_{1l} - x_1 x_{7u} + x_{7u} x_{1l} + \sqrt{2}/2\ x_{13} \leq 0.$$

$$\sqrt{2}/2\ x_{12} + x_9 x_{3u} + x_3 x_{9u} - x_{3u} x_{9u} \leq 20.$$

$$\sqrt{2}/2\ x_{12} + x_9 x_{3l} + x_3 x_{9l} - x_{3l} x_{9l} \leq 20.$$

$$\sqrt{2}/2\; x_{13} - x_9 x_{3u} - x_3 x_{9l} + x_{3u} x_{9l} \leq -20.$$

$$\sqrt{2}/2\; x_{13} - x_9 x_{3l} - x_3 x_{9u} + x_{3l} x_{9u} \leq -20.$$

$$- x_4 + x_5 - x_6 = 0$$

$$- x_7 + x_8 - x_9 = 0.$$

Each of these problems is defined for a set of x_{il} and x_{iu} for i=1,..9

Fig 6.1,6.2,6.3 represent combinatorial trees obtained when solving the three bar truss by the several B & B strategies described in this Chapter.

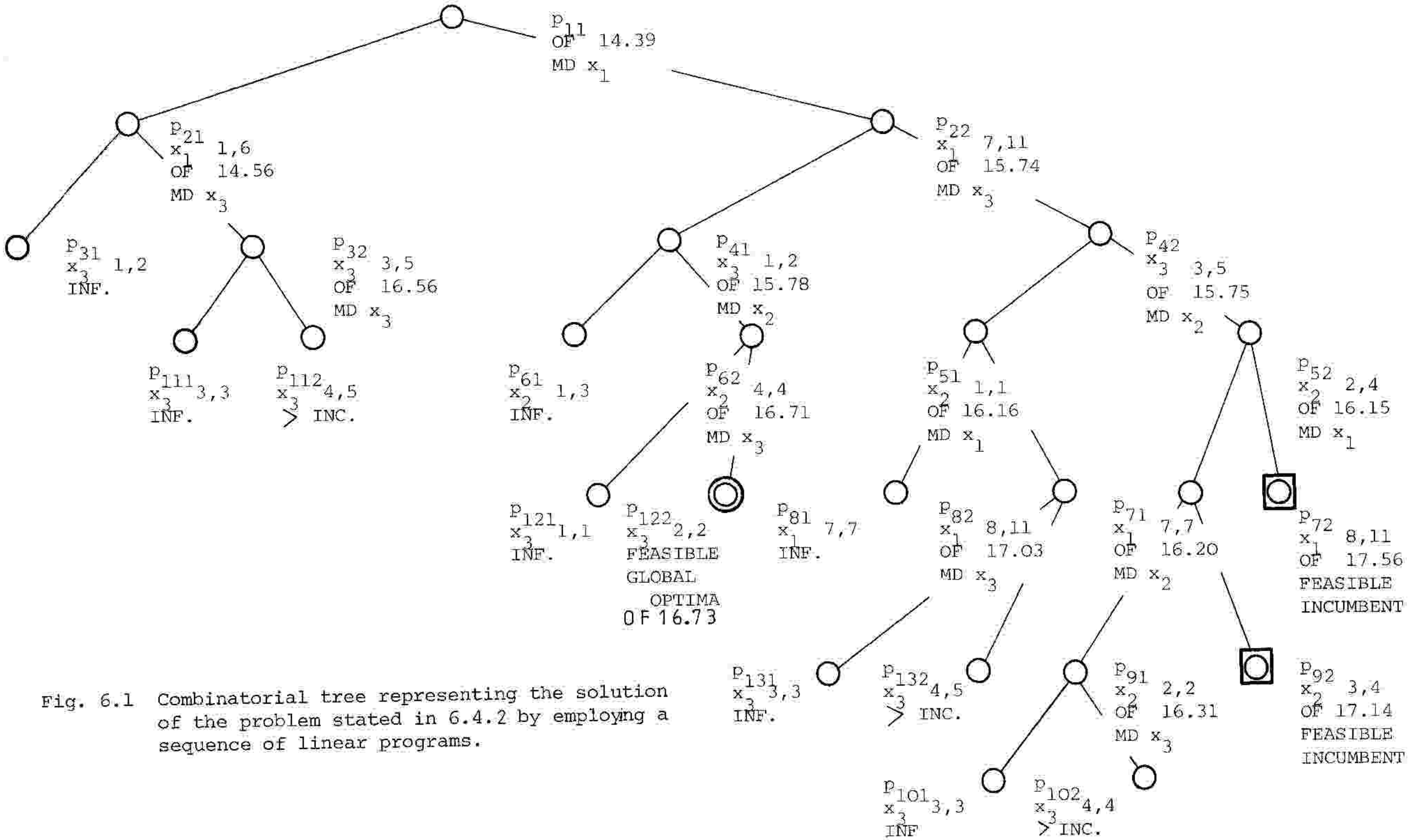


Fig. 6.1 Combinatorial tree representing the solution of the problem stated in 6.4.2 by employing a sequence of linear programs.

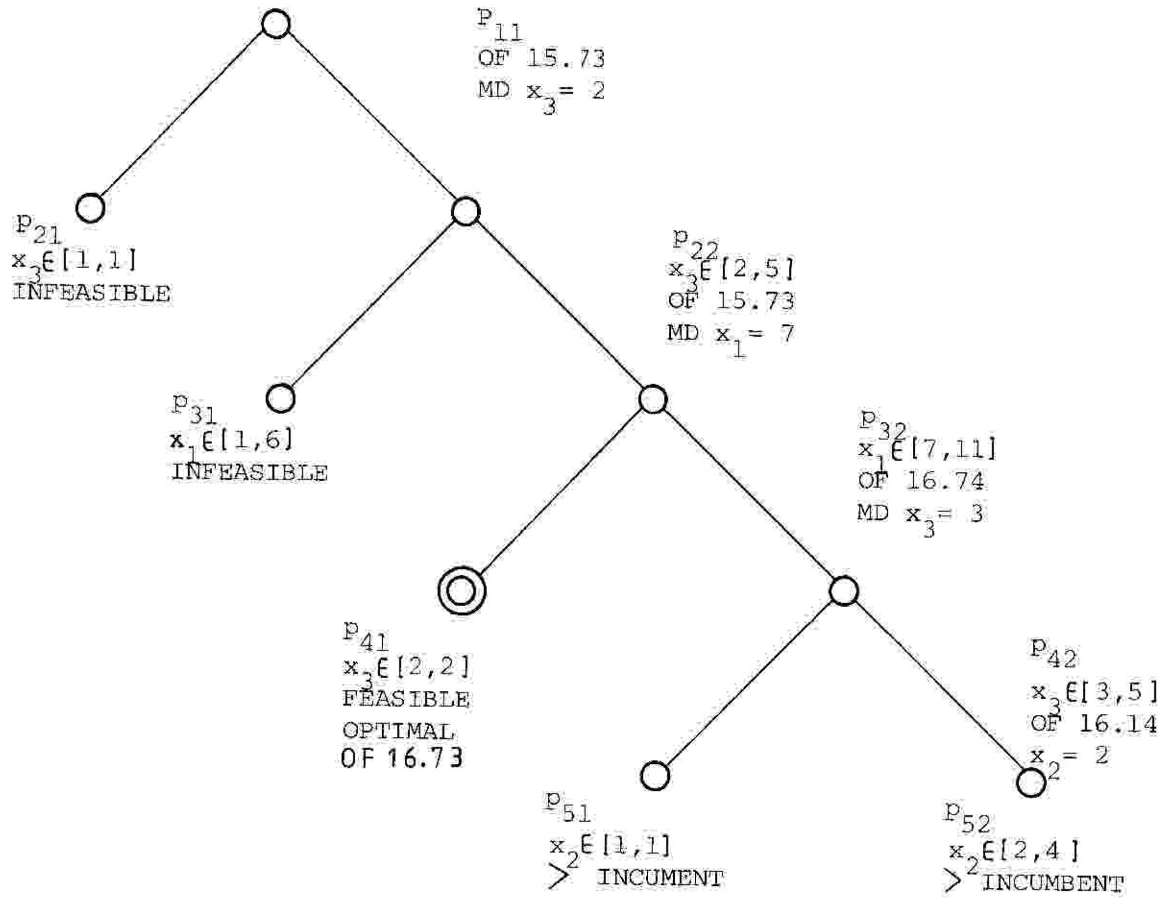


Fig. 6.2 Combinatorial tree representing the solution of the problem stated in 6.4.2 and employing a sequence of Linear Integer Programs strategy.

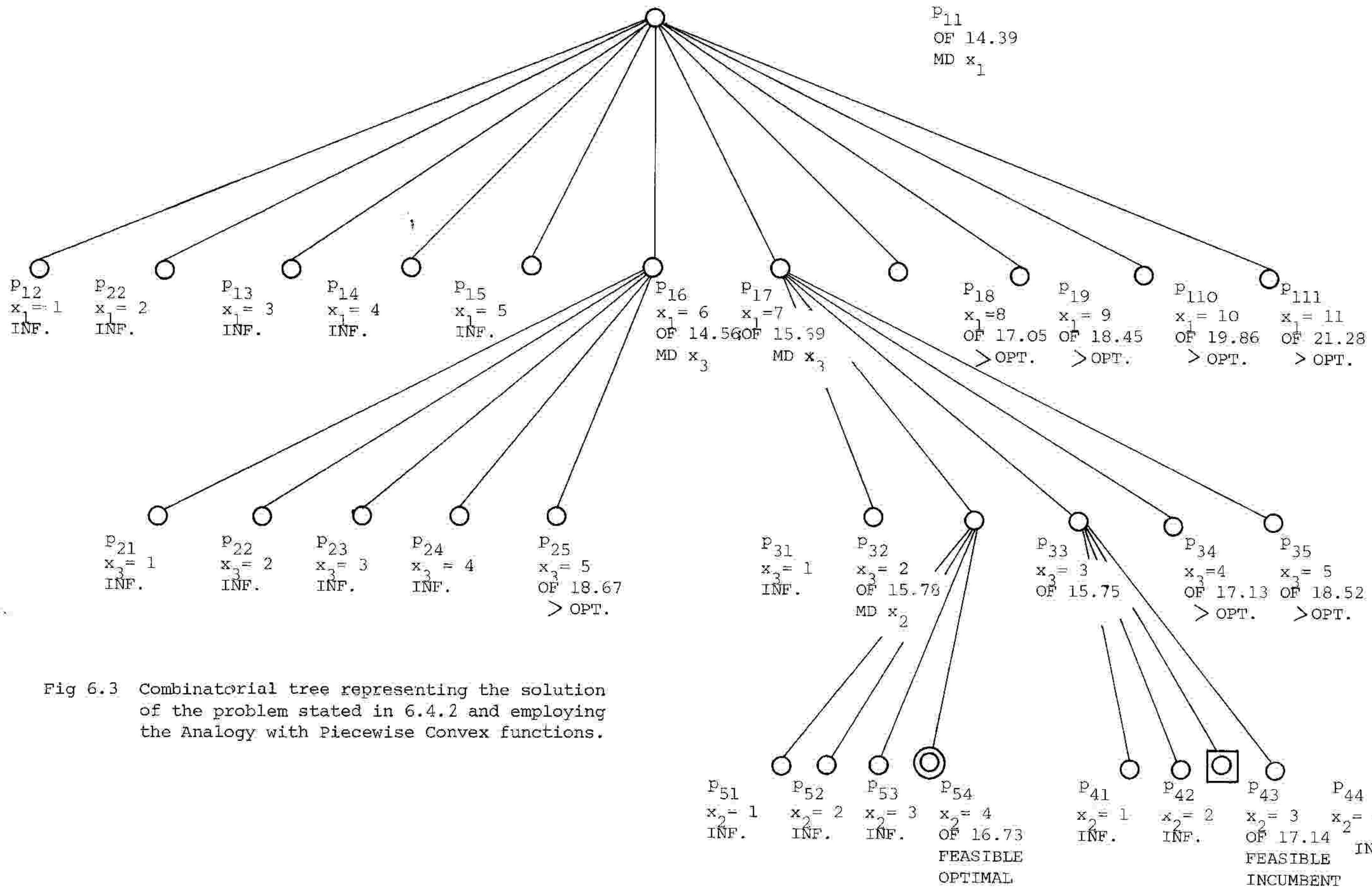


Fig 6.3 Combinatorial tree representing the solution of the problem stated in 6.4.2 and employing the Analogy with Piecewise Convex functions.

6.5. Computational Experience

The several strategies described in this Chapter were applied to find the solution of a three bar truss subjected to two alternative loading conditions with limits on the maximum allowable stresses and integer area variables. The global solution given by the formulation described as sequence of LP each in 13 variables and subject to 26 constraints was found after solving 24 subproblems and spending 6 ieu. By imposing the restriction that the design variables should always be integer the solution was found after solving 11 mixed linear integer problems and using 5 ieu. In this case the Lagrangian based Land and Powell subroutine has provided the solution to the subproblems defining each node. The strong branching rule described in the paragraph refering to the analogy with piecewise convex functions yield the global solution after only 5 ieu but having to solve 30 linear subproblems. The improvement in CPU time is due to the fact that a number of subproblems is infeasible and can easily be ruled out.

The (RIM) algorithm was proved to be more efficient than the B & B based methods. Two binary variables plus two real variables and five constraints were added to the relaxed master problem after each iteration. Land and Powell general purpose routine was used again and the global solution was found after 4 iterations and ellapsing 3 ieu. The discrete nature of the design variables leads to more efficient cut functions since only a limited number of areas make (RIM) feasible. The same statements holds true in the case of B & B strategy. A number of feasible points in the neighbourhood of the optimal solution are no longer feasible and can be ruled out.

In conclusion from the accumulated practical experience the introduction of discrete design variables renders the methods described more tractable. the R-D methods rely on a particular Ø-1 mixed LP and hope of progress is linked with the development of more efficient purpose built routine to solve this relaxed program.

CHAPTER SEVEN

CLOSURE

To bring this work to a closure some conclusions and indications for further research on the mathematical model presented in this Thesis are considered necessary.

Classical optimization principles constitute most of the setting for mathematical programming theory that has been previously used in structural engineering. They have introduced great simplicity into the design of perfectly plastic models which parallels that of the linear elastic analysis. The existence of multiple optima solutions even in small structural design problems is the most striking cause for the development of methods more appropriate for nonconvex programming.

When member sections are allowed to vary continuously,in special situations such as statical determinacy it is possible to show that the least volume design of trusses is after all "convex". Although the argument is no longer valid when considering redundant structures convex algorithms with a fast rate of convergence have been used in the past to solve this nonconvex problem. Solution values obtained by such an algorithm need not to be global but uniqueness can be ensured if a sufficient condition is satisfied. This local test was derived by considering a curved path linking the local optima to any other prospective candidate and it is easily carried out by simple operations such as matrix inversion.

If a fully nonconvex behaviour manifest itself then this problem is better treated combinatorially. Several strategies and approximations were tested and based on the results the following comments seem to be appropriate. Although Reeves' approach starting with a local minimum as its base point averts the growth in nearly feasible solutions that occurs in Soland's method the latter works better in eliminating bigger intervals where no solution exist. The use of an hybrid method combining Reeves' algorithm for local optima and Soland's rules for infeasible design points might constitute an interesting hint for further research.

If the members have to be chosen from a discrete set of commercially available gauges Bender's algorithm has converged faster to the global solution. As it has been already pointed out further investigation should be carried out on solution methods for the relaxed master program constrained by Wolsey's cuts. Although Lagragian relaxation based solution methods were the more efficient of all we have tried by neglecting the results of previous iterations it does not fully explore all the potentialities of this route. A conventional Bender's decomposition applied to each relaxed master problem yields a Ø-1 linear integer programming for wich more powerful tools are available.

These methods are also an invaluable framework for the use in other bilinear problems arising from structural engineering models such as reliability studies.

REFERENCES

ARGYRYS,J.H. (1960) Energy Theorems and Structural Analysis Butterworths

ARORA,J.S. AND HAUG,E.J. (1976) "Efficient Optimal Design of Structures by Generalized Steepest Descent Programming"
Int. J. Num. Meth. Engng. 10,744-766

ARORA,J.S. (1980) "Analysis of Optimality Criteria and Gradient Projection Methods for Optimal Structural Design"
Comp. Meth. Appl. Mech. Engng. 23,185-213

ARORA,J.S. , HAUG,E.J. AND RAJAN,S.D. (1981) "Efficient Treatment of Constraints in Large Scale Structural Optimization"
Engineering Optimization 5,105-120

ATREK,E. (1981) "On Stress Constraints in Structural Optimization"
Mechanics Research Comunications 8,61-66

BALAS,E. (1969) "Duality in Quadratic Programming : The Quadratic Case"
Management Science 16,1

BALAS,E. (1975) "Disjunctive Programming : Cutting Planes from Logical Conditions"
in "Nonlinear Programming 2" by MANGASARIAN,O.L. , Academic Press

BANICHUCK,N.V. (1980) "Optimization of the Shapes of Elastic Bodies"
(in Russian) , Nauka Moskow

BARNETT,R.L. (1961) "Minimum Weight Design of Beams for Deflections"
J. Mech. Eng. Div.,Proc. ASCE 87,75-109

BARTHOLOMEW,P. (1979) "Dual Bound used for Monitoring Structural Optimization Problems"
Engineering Optimization 4,45-50

BAZARAA,M.S. (1973) "Geometry and Resolution of Duality Gaps"
Naval Research Logistics Quarterly 20,357-366

BAYER,M. (1978) Bilinear Formulations in Structural Optimization ,Ph. D. Thesis
UCL,University of London

BERKE,L. (1972) "Convergence Behaviour of Iterative Resizing Procedures based on OC"
USAF,AFFDL Tech. Memo 1-FRB

BERKE,L. AND KHOT,N.S. (1974) "Use of Optimality Criteria Methods for Large Scale Systems"
AGARD,Lect. Ser. 70

BROWN,D.M. AND ANG,A.H.S. (1966) "Structural Optimization by NLP"
J. Struct. Div.,Proc. ASCE 92,319-340

CANTU,E. AND CINQUINI,C. (1979) "Iterative Solutions for Problems of Optimal Design"
Comp. Meth. Appl. Mech. Engng. 20,257-266

CARPENTER,W.C. AND SMITH,E.A. (1977) "Computational Efficiency of Nonlinear Programming Methods on a Class of Structural Problems"
Int. J. Num. Meth. Engng. 11,1203-1223

CELLA,A. (1971) "Automated Optimum Design from Discrete Components"
J. Struct. Div.,Proc ASCE 97,175-189

CHERN,J.M. AND PRAGER,W. (1971) "Minimum Weight Design of SD Trusses subject to Multiple Constraints"
Int. J. Solids Structures 7,931-940

COTTLE,R.N.,GIANESSI,F. AND LIONS,J.L. (1980) Variational Inequalities and Complementarity Problems
Wiley

CRISTOFIDES,N. , MINGOZZI,A. , TOTH,P. AND SANDI,C. (1979) Combinatorial Optimization
Wiley

DANTZIG,G.B. AND WOLFE,P. (1960) "The Decomposition Algorithm for Linear Programming"
Op. Research 8,101,111

DANTZIG,G.B. (1963) Linear Programmming and Extensions
Princeton University Press

FALK,J.E. AND SOLAND,R.M. (1969) "An Algorithm for Separable Nonconvex Programming Problems" Management Science 15,551-568

FALK,J.E. (1976) "Exact Solutions of Inexact LP" Op. Research 24,783-787

FLEURY,C. (1979) "Unified Approach to Structural Weight Minimization" Comp. Meth. Appl. Mech. Engng 20,17-38

FLEURY,C. (1980) "Efficient Optimality Criteria Approach to the Minimum Weight Design of Elastic Structures" Computers and Structures11,163-173

FUCHS,M. (1980) "Linearized Homogeneous Constraints in Structural Design" Int. J. Mech. Sci. 22,33-40

GALILEO,G.L. (1638) "Discursi et Dimonstrazione Matematiche Intorno a Due Nuove Science Attenti alla Mechanica et i Movimenti Locale" Leida

GARFINKEL,R.S. AND NEMHAUSER,G.C. (1972) Integer Programming Wiley

GELLATLY,R.A. (1966) "Development of Procedures for Large Scale Automated Minimum Weight Structural Design" AFFDL,Tech. Rep. 180

GEOFFRION,A.M. (1970) "Primal Resource-Directive Approaches for Optimizing Nonlinear Decomposable Systems" Operations Research 17,375-403

GEOFFRION,A.M. (1972) "Generalized Bender's Decomposition" J. Opt. Theory Appl. 10,237-260

GEOFFRION,A.M. AND GRAVES,G.W. (1974) "Multicommodity Distribution System designed by Bender's Decomposition" Management Science 5, 822-844

GEOFFRION,A.M. (1977) "How can Specialized Discrete and Convex Optimization Methods be Married" Annals of Discrete Mathematics 1,205-220

GIANESSI,F. AND NICCOLUCCI,F. (1976) "Connections between NLP and IP problems" in Symposia Mathematica vol. XIX , Academic Press

GILLETT,B.E. (1976) Introduction to O.R. Methods McGraw-Hill

GILL,P.E. AND MURRAY,W. (1977) "The Computation of Lagrange Multiplier Estimates for Constrained Minimization" NPL NAC Report n. 77

GILL,P.E. AND MURRAY,W. (1978) "Numerically Stable Methods for QP" Math. Programming 14,349-372

HALMOS,E. AND RAPSACK,T. (1978) "Minimum Weight Design of Statically Indeterminate Trusses"
Math. Programming Study 9,109-119

HAUG,E.J. AND ARORA,J.S. (1979) Applied Optimal Design
Wiley

IMAI,K. AND SHOJI,M. (1981) "Minimum Cost Design of Framed Structures by the Mini-Max Dual Method"
Int. J. Num. Meth. Engng. 17,213-229

IUTAM (1975) Syposium on Optimization in Structural Design - Warsaw 1973 , SAWCZUCK,A. AND MROZ,Z. (Ed.)
Springer-Verlag Berlin

JOHNSON,D. (1982) "Selection of Variables for Linearized Truss Optimization"
Engng. Optimization 5,209-213

KANEKO,I. (1982) "On some Recent Engineering Applications of C omplementarity Problems"
Math. Programming Study 17,111-125

KHOT,N.S. (1981) "Algorithms Based on OC to Design Minimum Weight Structures"
Eng. Optimization 5,73-90

KIRSCH,U. AND MOSES,F. (1979) "Decomposition in Optimum Structural Design"
J. Struct. Div. Proc. ASCE 105,85-100

KIRSCH,U. (1982) "Optimal Design Based on Approximate Scaling" , J. Struct. Div. Proc ASCE 108,888-909

KOHNO,H. (1976) "Cutting Plane Algorithm for Solving Bilinear Programs"
Math. Programming, 11,14-27

KOVACS,L.B. (1980) Combinatorial Methods of Discrete Programming
Akademiai Kiado , Budapest

KRISHNAMOORTHY,C.S. AND MOSI,D.R. (1979) "A Survey on Optimal Design of Civil Engineering Structural Systems"
Eng. Optimization 4,73-88

LAND,A.H. AND POWELL,S. (1973) FORTRAN Codes for Mathematical Programming
Wiley

LASDON,L.S. (1970) Optimization Theory for Large Systems
McMillan

LAZIMY,R. (1982) "Mixed Integer Quadratic Programming"
Math. Programming, 22,332-349

LAWLER,E.L. AND WOOD,D.E. (1966) "Branch and Bound Methods: a Survey"
Operations Research 14,699-719

LUENBERGER,D.G. (1969) Optimization by Vector Space Methods
Wiley

MANCINI,L.J. AND McCORMICK,G.P. (1976) "Bounding Global Minima"
Math. Op. Research 1,50-53

MARCAL,P.V. AND GELLATLY,R.A. (1968) "Applications of the Created Response Surface Technique to Structural Optimization",USAF,2nd Conference on Matrix Methods in Structural Mechanics AFFDL,Tech. Rep. 150

MELLO,E.L. (1980 Some Applications of Generalized Inverse Theory to Structural Problems , Ph. D. Thesis IC , University of London

MICHELL,A.G.M. (1904) "The Limits of Economy of Materials in Framed Structures" Philosophical Magazine , series 6 , Vol. 8, 589-597

MORRIS,A.J. (1976) "Unified Approach to Fully Stressed Design" , Eng. Optimization 2,3-15

MOSES,F. (1964) "Optimum Structural Design using LP" J. Struct. Div.,Proc. ASCE, 90,89-104

McCORMICK,G.P. (1972) "Attempts to Calculate the Global Solution of Problems that may have Local Maxima" in Numerical Methods for Nonlinear Optimization , LOOTSMA,F.A. (Ed.) , Academic Press

McCORMICK,G.P. (1976) "Computability of Global Solutions to Factorable Nonconvex Programs: Part I Convex Underestimating Problems" Math. Programming 10,147-175

McCORMICK,G.P. (1979) "Future Directions in MP" in NATO-ASI on Engineering Plasticity by MP - University of Waterloo 1977 , COHN,M.Z. AND MAIER,G. (Ed.) , Pergamon Press

McCORMICK,G.P. (1980) "Locating an Isolated Global Minimizer of a Constrained Nonconvex Program" Math. Op. Research 5,435-443

NATO-ASI (1981) on Optimization of Distributed Parameter Structures Vol. 1 - Iowa City 1980 , HAUG,E.J. AND CEA,J. (Ed.)
Sijthoff and Noordhoff

NIORDSON,F.I. AND PEDERSEN,P. (1975) "A Review of Optimal Structural Design"
in IUTAM on Optimization in Structural Design - Warsaw 1973 , SAWCZUK,A. AND MROZ,Z. (Ed.) , Springer-Verlag Berlin

NOBLE,B. (1968) Applied Linear Algebra
Prentice-Hall

POWELL,M.J.D. (1978) "Algorithms for Nonlinear Constraints that use Lagrangian Functions"
Math. Programming 14,224-248

PRAGER,W. (1974) "Introduction to Structural Optimization. Courses and lectures : International Centre for Mechanical Science , n. 212
Vdine , Springer-Verlag Vienna

PRZEMIENIECKI,J.S. (1968) Theory of Matrix Structural Analysis
McGraw-Hill

RAJARAMAN,A. AND SCHMIT,L.A. (1981) "Basis Reduction Concepts in Large Scale Structural Optimization"
Eng. Optimization 5,91-104

RAO,J.S. (1978) Optimization Theory and Applications
Wiley Eastern

RAZANI,R. (1965) "The Behaviour of Fully Stressed Design of Structures and its Relationship to Minimum Weight Design"
AIAA,3,2262-2268

REEVES,G.R. (1973) Nonconvex All Quadratic Programming , Sc. D. Thesis , Mathematics
Washington University

ROCKAFELLAR,R.T. (1976) "Solving a NLP by means of a Dual Problem"
in Symposia Mathematica vol. XIX , Academic Press

REINSCHMIDT,K.F. , CORNELL,A.C. AND BROTCHIE,J.F. (1968) "Iterative Design and Structural Optimization"
J. Struct. Div.,Proc. ASCE, 92,281-318

REINSCHMIDT,K.F. (1971) "Discrete Structural Optimization"
J. Struct. Div. Proc. ASCE 97,133-156

SAGLAM,M.R. (1981) "Computer Oriented Algorithms for Solving Problems with Discrete Programming Techniques"
in NATO-ASI on Optimization of Distributed Parameter Structures - Iowa City 1980 , HAUG,E.J. AND CEA,J. (Ed.) , Sijthoff and Noordhoff

SCHMIT,L.A. AND MALLETT,R.A. (1963) "Structural Synthesis and Design Parameter Hierarchy"
J. Struc. Div.,Proc. ASCE, 89,269-299

SCHMIT,L.A. AND MIURA,H. (1976) "Approximation Concepts for Efficient Structural Synthesis"
NASA,Contractor Report 2552

SECHLER,E.E. AND DUNN,C.G. (1963) Airplane Structural Analysis and Design
Dover

SHAPIRO,J. (1976) "A new Constructive Duality Theory for Mixed Integer Programming"
in Symposia Mathematica vol XIX , Academic Press

SHEU,C.Y. AND PRAGER,W. (1968) "Recent Developments in Structural Design"
Applied Mechanics Reviews, 21,985-992

SINGARY,N.M. AND RAO,J.K.S. (1975) "Optimization in Trusses using Optimal Control Theory"
J. Struct. Div.,Proc. ASCE, 101,1037-1052

SOLAND,R.M. (1971) "Algorithm for Separable Nonconvex Programming Problem : Nonconvex Constraints"
Management Science 17,759-773

SOLAND,R.M. (1973) "Algorithm for Piecewise Convex Programming Problems"
Naval Research Logistics Quarterly 20,325-339

SOYSTER,A.L. (1973) "Convex Programming with Set-Inclusive Constraints and Applications to Inexact LP"
Op. Research 21,1154-1157

SVED,G. AND GINOS,Z. (1968) "Structural Optimization under Multiple Loading"
Int. J. Mech. Science 10,803-805

TEMPLEMAN,A.B. (1976) "A Dual Approach to Optimum Truss Design" , J. Struct. Mechanics, 4,235-255

THESEN,A. (1978) "Branch and Bound Methodology"
in Computer Methods in Operations Research , Academic Press

THUENTE,D.J. (1980) "Duality Theory for GLP with Computational Methods"
Op. Research 28,1005-1011

TOACKLEY,A.R. (1968) "Optimum Design Using Available Sections" , J. Struct. Div. Proc. ASCE 94,1219-1241

VENKAYYA,V.B. (1978) "Structural Optimization : A Review and some Recomendations"
Int. J. Num. Meth. Engng. , 13,203-228

WASIUTYNSKY,Z. AND BRANDT,A. (1963) "The Present State of Knowledge in the Field of Optimum Design Structures"
Applied Mechanics Reviews , 16,341-350

WOLSEY,L.A. (1980) "Heuristic Analysis,Linear Programmming and Branch and Bound"
Math. Programmming Study 13,121-134

WOLSEY,L.A. (1981) "Resource-Decomposition Algorithm for General Mathematical Programs"
Math. Programmming Study 14,244-257

YATES,D.F.,TEMPLEMAN,A.B. AND BOFFEY,T.B. (1982) "The Complexity of procedures for Determining the Minimum Weight of Trusses with Discrete Members"
Int. J. Solids Structures 18,487-495