OPTIMAL TOLERANCE ALLOCATION

OPTIMAL TOLERANCE ALLOCATION

by

WAHEED K. MICHAEL, B.Sc. (Eng.), M.Eng.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

July 1980

DOCTOR OF PHILOSOPHY (1980)
(Mechanical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: OPTIMAL TOLERANCE ALLOCATION

AUTHOR: Waheed K. Michael, B.Sc. (Cairo University)
M.Eng. (McMaster University)

SUPERVISOR: Professor J. N. Siddall

NUMBER OF PAGES: (xi), 155

ABSTRACT

This thesis addresses itself to one of the most general theoretical problems associated with the art of engineering design. Viewed in its entirety the proposed approach integrates the relation between the design and production engineers through the theory of nonlinear optimization. The conventional optimization problem is extended to include the optimal allocation of the upper and lower limits of the random variables of an engineering system. The approach is illustrated by an example using a sequence of increasingly generalized formulations, while the general mathematical theory is also provided. The method appears to offer a practical technique provided a satisfactory cost function can be defined.

The thesis presents an analytical approach to full acceptability design conditions as well as less than full acceptability or scrap design conditions. An important distinction between the design and the manufacturing scrap has been introduced and illustrated through examples.

The space regionalization technique is utilized to estimate the system design scrap. Optimization strategies are introduced to the mathematically defined upper and lower limits of the regionalization region. This region is then discretized into a number of cells depending upon the probabilistic characteristic of the system random variables.

The analytical approach exhibited does not rely explicitly on evaluation of partial derivatives of either the system cost objective or any of its constraints at any point. Moreover, the technique could be applied to engineering systems with either convex or nonconvex feasible regions. It could also be exercised irrespective of the shape of the probabilistic distributions that describe the random variables variation.

Industrially oriented design examples are furnished to justify the applicability of the theory in different engineering disciplines.

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude and appreciation to his supervisor Professor J. N. Siddall for his guidance and encouragement throughout the course of this study. He also thanks Dr. A. Smith, Dr. J. H. T. Wade and Professor W. R. Newcombe, members of his supervisory committee for their continuing interest.

The scholarship offered by the Natural Sciences and Engineering Research Council for two years, the Ontario Graduate Scholarship grant offered by the Ontario Government and the Department of Mechanical Engineering financial support through a Teaching Assistantship are gratefully acknowledged.

Thanks are due to Myrna Velasco my wife for her help and forebearance.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Tolerances are a recognition of the fact that perfection cannot be achieved. They can be defined generally as the limits imposed on the variability of some design variables or specifications. The purpose of this study is to examine how tolerances can be integrated into the overall optimization decision problem in an analytical way. Tolerances are commonly associated with machined dimensions of design components. They should be generalized, however, as bounds on any quantities. Examples are the yield point of a metal, the stiffness of a spring, and the horsepower of an engine. Although, modern machine tools are capable of machining to a high level of accuracy, their ability to duplicate a specific dimension, e.g., shaft size, on a repetitive basis is limited because of tool wear, deflections and vibrations of the machine and the workpiece, temperature changes, in addition to human errors. Similarly, steel cannot be made with an exact yield point, or an engine cannot be built with an exact maximum horsepower. The user, therefore, must accept some tolerance on the nominal values. Although, he would usually prefer tight tolerance, the tighter the tolerance the higher the cost.

In the design of a system the emphasis on its function, and the unawareness of the production difficulties, often lead to the application of tight tolerances that are difficult to attain economically, since their assignment has traditionally been done wholly by judgement. With few exceptions the usual methods of selecting tolerances do not optimize cost directly, since the designer's concern is to specify tolerances so that the system can first function and then hopefully be the least expensive.

In the conventional optimization design problem, the problem of interest is finding one single point in the feasible region which minimizes or maximizes the problem objective(s) [1,2]*. This optimum point is the vector of the problem deterministic design variables. Since many other points can also meet the required system specifications, the designer can assign tolerances on the system component dimension values so as to minimize the total production cost. The main concern in this research, therefore, is the problem of the best possible trade-off between tolerance and cost, which could be stated more generally as one of choosing the tolerances which maximize the overall value of the system [3].

1.2 Literature Review

The tolerance problem has attracted deep interest among designers in different disciplines. The tolerance studies,

* Number between brackets designates references at the end of the thesis.

however, were first introduced as applied on mechanical systems where a large amount of literature is available [4-35] in the areas of tolerance specifications and cost analysis. Muskets built by Eli Whitney in 1812 are among the earliest examples of mass-produced items with fully interchangeable components. The most appropriate level of interchangeability is not necessarily the highest. Under many circumstances, complete universal interchangeability is neither necessary nor economically justifiable. Standardization and interchangeability are closely related. The first standard tolerance limits system was established in 1902 by Newall. This was followed by a British Standard No. 164 in 1906 which was superseded by the present standards such as: ASA - American Standards Association, ASME - American Society of Mechanical Engineers, BS - British Standards, CSA - Canadian Standards Association, ISA - International Standards Association and ISO - International Organization for Standardization. Even though in practice one usually has to choose from a finite set of discrete standard tolerances which follows one of the previous systems. The continuous tolerance solution, however, yields an absolute minimum cost which is definitely of interest; since it can serve as the basis for selection of discrete tolerance values.

The full acceptance, FA design principle - also known as infallible interchangeability - was observed when assigning tolerances to components and when considering the effect of these tolerances upon the assembly and the functioning of the finished products. Where safety is of paramount importance, e.g., missiles, nuclear equipments, and elevators; it is understandable that designers should insist upon FA no matter what the cost. However, there are many cases where insistance upon FA is not justified, and an occasional failure to assemble or to function would not be serious, particularly when judged in the light of the overall economic gain in production. Statistical dimensioning analysis, therefore, was proposed to guide the selection of the tolerances where the probabilistic distributions of the associated processes are assumed known.

The dominant statistical approach suggested in the literature, [4-29], depends on the validity of some unrealistic assumptions. They are:

i) Each machined dimension in an assembled component should come from a process that follows a normal distribution.

ii) The size of any individual component is independent of the size of any other component.

iii) The total tolerance spread on each part is equal to a predetermined multiple of the standard deviation of the part normal distribution.

iv) The mean values of the randomly distributed dimensions coincide with the corresponding mean of the production process.

v) The percentage of assemblies permitted to include any deviation in any of their parts from the blue-print tolerance zone sometimes has to be guessed or specified.

vi) The performance function, F, which usually describes the relationship between different physical dimensions in a design component, has to be in a simple linear form. It is a function of the random dimension variables, $\underline{X}$. The variance of a sum or difference is proportional to the sum of the variances of individual items. Generally, it could be expressed mathematically as follows:

$$\sigma_F = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial F(\underline{X})}{\partial x_i}\right)^2 \sigma_i^2} \qquad (1.1)$$

where, σ_F is the standard deviation of the assembled dimension, N is the total number of the random variables x_i, and σ_i^2 is the variance of the random variable i which should follow a symmetrical normal distribution.

The main objective of the research published in the literature mentioned before could be categorized into two broad sections. In the first category, the object is to determine the tolerances of the individual dimensions in a chain based on a specified tolerance of the sum dimension. This has been done by adapting different assumptions such as: equal tolerances, tolerances which are proportional to their associated nominal dimensions, tolerances which are proportional with process deviation or with process relative cost. In the second category, the object is to compute the resulting tolerance for an assembly when the tolerances of the components are given. The problem has been tackled for both the full acceptance design and the stochastic conditions.

In spite of the large amount of literature available in the area of tolerance specification as applied to mechanical systems, many authors have considered the multiple constraint system as a single constraint system - e.g., the performance function constraint $F(\underline{X})$ - and have also dealt with nonlinear constraints as linear by utilizing a truncated Taylor series expansion.

The primary objective of an engineering system design and of the dimensional specifications for its component parts is to ensure that the system will give the service desired. This could be mathematically expressed as a set of governing inequality constraints. The secondary objective is to facilitate the manufacture or the purchase of the system component parts as cheaply as possible. This raises the necessity of expressing the system cost objective as a function of both the nominal variables and their tolerances.

The relationship between tolerances and their associated cost is strongly influenced by the manufacturing methods and the lot size. Not only is the precision which can be maintained in a given machine tool difficult to determine, but also, the relationship between precision and cost is difficult to fit to analytical cost models with a reasonable accuracy. Depending upon the required precision, different processes have to be selected. A chain of processes may be rough turning, finish turning, grinding, etc.

Tolerance cost models, therefore, have to cover the cost characteristics for both individual processes as well as a sequence of processes. Also, distinction between different cost components should be clearly made; e.g., machining cost, repairing cost, scraping cost, inspection cost, assembly cost, etc.

Some of the cost functions, U, mentioned in the literature [5, 7, 28, 30-37] are:

$$U = \sum_{i=1}^{N} (c_{1i}/t_i) \tag{1.2.a}$$

$$U = \sum_{i=1}^{N} (c_{1i}/t_i^2) \tag{1.2.b}$$

$$U = \sum_{i=1}^{N} (c_{1i} + c_{2i} (t_i^2)^{c_{3i}}) \tag{1.2.c}$$

$$U = \sum_{i=1}^{N} (c_{1i} + c_{2i} e^{c_{3i} t_i}) \tag{1.2.d}$$

$$U = \sum_{i=1}^{N} (x_i^o/t_i) \tag{1.2.e}$$

$$U = \sum_{i=1}^{N} c_{1i}/1-S \tag{1.2.f}$$

$$U = \sum_{i=1}^{N} t_i^{-1}/1-S \tag{1.2.g}$$

where N is the number of the random design variables, and x_i^o and t_i are the nominal and the associated absolute symmetrical tolerance values, respectively. c_{1i}, c_{2i} and c_{3i} are constants; S is the system scrap percentage. The design cost objectives (1.2.a-d) are inversely proportional to the x_i^o, and they are used generally to determine the optimum symmetrical tolerance values for FA design conditions or for a specified maximum allowable scrap. The desing objective (1.2.e), on the other hand, was aimed at optimizing both the nominals and the tolerances of a system. The design criterion expressed in Equation (1.2.f) minimizes the system scrap percentage while Equation (1.2.g) maximizes the tolerances as well. The optimum outcome will be a trade-off between the inflated system tolerances and the corresponding increase in the scrap percentage.

The validity of the simulation of a system is bounded by the accuracy of the mathematical cost model as well as the performance constraints formulation. Therefore, the closer the system cost model represents the manufactured system design, machining, inspection, assembly, testing and repairing conditions, the more accurate the optimum solution will be.

Even though mechanical systems have a longer history in tolerance specification, electrical systems are comparatively more advanced in tolerance design. Emphasis will be placed in this literature review on some of the more ingenious analytical methods for statistical circuit analysis.

Generally, a system scrap, S, could be mathematically expressed as

$$S = 1 - \mathrm{Prob}(\underline{X} \in R_c) = \iint_{R_c} \cdots \int f(\underline{X})\, d\underline{X} \qquad (1.3)$$

where $R_c = \{\underline{X} | \underline{\Phi}(\underline{X}) \geq 0\}$; $\underline{X} = [x_1, x_2 \ldots x_N]^T$

$\underline{X}$ is the system random design variables.

$\underline{\Phi}(\underline{X})$ is the system set of inequality constraints.

$f(\underline{X})$ is the joint probability density function of the random variables $\underline{X}$.

R_c is the system feasible closed region.

Since R_c is an implicit function of $\underline{X}$, S cannot be evaluated analytically and the usual method of evaluating it was by using Monte Carlo analysis, [4, 31]. It is expensive, however, to combine Monte Carlo techniques with optimization to accurately predict the optimum system scrap. This is because of the large number of system simulations per optimization iteration which may be required.

The simplical approach [38, 39] approximates the boundary of the feasible region of an N-dimensional design space with a polydrom of bounding (N-1) simplices. The feasible region contains all the design outcomes that satisfy the system performance constraints. A crude estimate of the system scrap percentage - the complement of the system level of acceptability or yield - is obtained by performing Monte Carlo analysis directly in the variables space outside the approximated feasible region which could be updated in the mean time using the Monte Carlo results.

If the system random variables are assumed to be statistically independent and symmetrically distributed, Karafin [40] proposed an analytical method which approximates the system scrap by computing its upper and lower bounds. This has been done by applying truncated Taylor series approximations on the system constraints which have to be normally distributed.

The space regionalization technique [41, 42] divides the tolerance region into finite number of non-overlapping cells, each covering a sector of the joint density space, and a weight is assigned to it accordingly. The center of the cells located outside the full acceptance region are checked against the system nonlinear constraints to determine whether the whole weight of the cell will contribute to the system scrap or not. The method can handle sets of dependent as well as independent variables.

Elias [43] developed a program which minimizes the system scrap when statistics for the random varialbes are given. The program adjusts the variables' specifications iteratively and the value of the system scrap is updated by repeating the Monte Carlo analysis.

To reduce the number of Monte Carlo simulations while keeping high confidence in the scrap estimate, the importance sampling [44] approach was adapted. It concentrates the distribution of sample points at some critical regions instead of spreading them evenly.

The methods described above do not explicately optimize either the system random variables' nominal values or their assigned tolerances. Pinal and Roberts [45] minimized a cost function and approximated the system constraints by truncated

Taylor series expansions. They considered the values of the nominal variables fixed, while Bandler et al. [37, 46], permitted the nominal point to move. An orthotope describing the tolerance region is to be inflated within the feasible region and beyond it. The center of the orthotope provides the nominal parameter values and the lengths of the orthotope edges are twice the absolute tolerances. Abdel-Malek [36] used multidimensional linear cuts of the tolerance orthotope to estimate the system scrap. While using arbitrary statistical distributions for the random variables, the tolerance orthotope is partitioned into a collection of orthocells and a weight is assigned to each. The accuracy of the system scrap estimated using this method depends on the validity of the one-dimensional convexity condition which should be preserved for all the system inequality constraints. Also, the calculation procedures rely heavily on the exact evaluation of the first derivative of both the system objective and constraints.

Throughout the work presented in this thesis, the engineering system random variables are described in terms of nominal values and tolerance distributions. The tolerance limits are considered design variables and can be defined in terms of the nominal value, x^o, of the base variable. The random variable, x, thus has a value in the region

$$x^o - t^- \leq x \leq x^o + t^+ \qquad (1.4)$$

where t^- and t^+ are the lower and upper tolerance deviations, respectively.

1.3 Original Contributions Claimed

Having briefly scanned the tolerance assignment state of the art, various limitations and approximations of each of the reported concepts have been pointed out. This has led to the necessity of searching for a more versatile methodology that relies on optimization techniques with as few assumptions as possible, and that describes practical engineering systems with minimum diversity.

To provide insight into the tolerance assignment problem, Chapter 2 presents a simple design problem of two fitted cylinders that is solved in a step-by-step fashion on seven stages. The mathematical definition of the toleranced design problem with full acceptance is given, and the level of a design acceptability is defined.

The problem with less than full acceptance is constructed in Chapter 3, where a distinction between the design and the manufacturing scrap is introduced. Chapter 3 also presents an analytical approach which not only provides a system positive and negative tolerances for each random design variable but also facilitates the evaluation of the optimum scrap percentages of the system. This is made possible by utilizing the space regionalization technique. The approach is general enough to be used in conjunction with any statistical distribution. An emphasis however is placed on uniform and beta distributions, and justifications for this emphasis are given.

Chapter 4 introduces an algorithm to allocate the upper bound of the regionalization region. Definitions and concepts as well as geometric interpretations are given. The procedure is simple and cheaply evaluated. It leads to considerable computational savings while defining the scrap optimum design.

The ideas presented in Chapter 4 are implemented in the strategy given in Chapter 5 to overcome the convexity assumption that had to be fulfilled earlier. The possible inaccuracy in the system scrap estimates is discussed, and a sensitivity analysis of the estimated errors is also done.

The last part of Chapter 5 is devoted to some practical implementations of the approach and the algorithms previously presented.

The mathematical proof for identifying the worst condition constraints is given in Appendix A.

Original contributions claimed for this thesis are:

1. A formulation of the design problem in optimization terms embodying nonsymmetrical tolerancing and system scrap.
2. Proposal of a more realistic cost function.
3. The distinction between manufacturing and design scrap.

4. Incorporation of the space regionalization technique with optimization algorithms.

5. Partitioning the regionalization domain according to the random variables' distributions.

6. Elimination from the regionalization region of all the cells adjacent to the active corners of the full acceptable region, in order to increase the procedure efficiency.

7. Algorithms to mathematically define the acceptable regionalization upper bound region.

8. A procedure to allocate the optimum acceptable lower bound region for a system with non-convex feasible region.

CHAPTER 2

TOLERANCE ASSIGNMENT WITH FULL ACCEPTANCE

2.1 Introduction

In general, from the manufacturing point of view, engineering problems can be categorized into two broad sections, depending upon the number of components produced and their functions. They are either mass produced components or job produced components. Acceptability can be defined as the fraction of components satisfying manufacturing specifications. A high acceptability level is essential in job production, but in mass production low production cost is the more typical criterion. Batch production will fall somewhere in between the two. In this chapter only production with 100 percent acceptability (sometimes known as worst case design) [47] will be considered. This might be applicable in either job or batch production, where there is less likelihood of scrapping components which do not meet design specifications than in mass production.

Designs in general are subjected to manufacturing tolerances on the physical dimensions or properties of the components; and also must meet performance requirements. In the presentation and discussion which will follow, the manufacturing limits will be the only variable tolerances to be determined optimally, even though a design might also be subjected to uncertainties in its other specifications, or in its in-service or off-service environmental conditions. At this stage all of the design specifications and parameters, except tolerances, will be assumed deterministically known with no deviation. The "full acceptability" optimum design with a deterministic treatment could be justified as an end in itself, or it might be considered as a preliminary exercise leading up to "scrap optimum design", where we have less than full manufacturing acceptability, and quantities other than tolerances may also be treated as random.

The conventional way of introducing a new idea or theory in the engineering field is to state the theory and its governing assumptions, verify it mathematically, and then elaborate its application using examples and practical problems. We will, however, tackle the situation with an opposite approach - solving a simple practical design problem in a step-by-step fashion by releasing some assumptions in each design stage until the problem approaches a real life practical case; then the optimum full acceptability tolerance design procedure will be generalized and the theory will be verified.

2.2 The Problem

The basic problem, [P], will be a cylinder subjected to internal pressure [48]. It is diagramatically sketched in Figure 2.1, and consists of two cylinders shrunk together. The strength requirement is therefore fulfilled by utilizing the trapped prestress. The following deterministic specifications are set for the design.

Applied internal pressure	$p_o = 50 \times 10^6$ Pa
Allowable yield strength	$S = 150 \times 10^6$ Pa
Nominal inner radius	$r_o = 0.1$ m
Maximum outer radius	$r_{max} = 0.2$ m
Maximum interference	$D_{max} = 10^{-4}$ m
Minimum cylinder thickness	$t_{min} = 0.01$ m
Modulus of elasticity	$E = 2 \times 10^{11}$ Pa

2.2.1 The Deterministic Problem $[P]_1$

The problem design variables, $[\underline{X}^o]_1$, are

$$[\underline{X}^o]_1 = \begin{bmatrix} r_1 \\ r_2 \\ p_f \end{bmatrix}_1 = \begin{bmatrix} \text{intermediate radius, m} \\ \text{outer radius, m} \\ \text{shrink fit pressure, Pa} \end{bmatrix} \tag{2.1}$$

The optimization criterion, U_1, is the minimization of the overall material cost (i.e., the cylinder volume, or simply the outer radius squared because the inner radius is fixed).

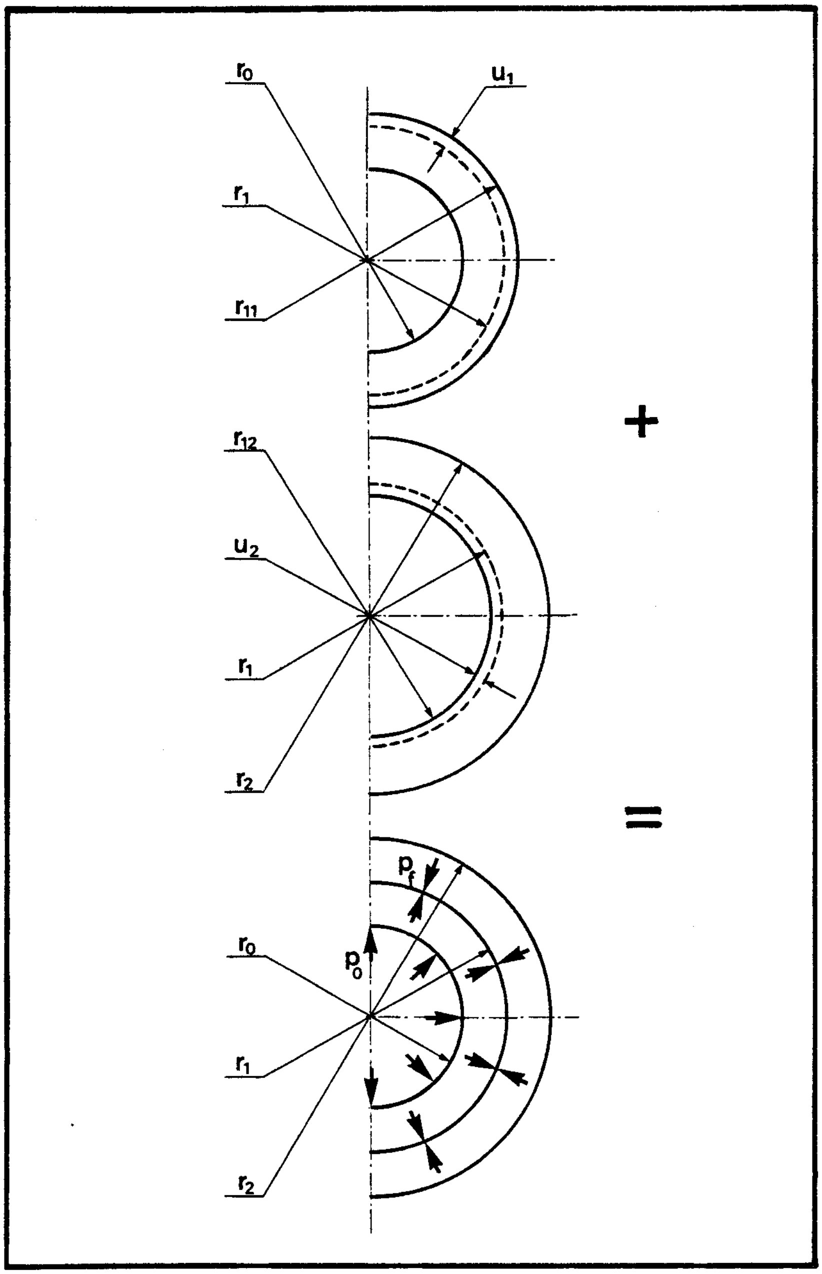


Figure 2.1 Shrink fitted cylinders subjected to internal pressure, p_o.

$$U_1 = r_2^2 = \text{minimum} \tag{2.2}$$

Constraints, $[\underline{\Phi}]_1$

$$\phi_{1,1} = r_1 - r_o - t_{min} \geq 0$$

$$\phi_{2,1} = r_2 - r_1 - t_{min} \geq 0$$

$$\phi_{3,1} = r_{max} - r_2 \geq 0 \tag{2.3}$$

$$\phi_{4,1} = p_f \geq 0$$

$$\phi_{5,1} = D_{max} - D_a[r_1,r_2] \geq 0$$

$$\phi_{6,1} = S - 2\ \tau_1[r_1,r_2] \geq 0$$

$$\phi_{7,1} = S - 2\ \tau_2[r_1,r_2] \geq 0$$

where;

$D_a[r_1,r_2]$ = actual interference between the inner and outer cylinders, m

$$= \frac{2\ p_f}{E} \frac{r_1^3(r_2^2-r_o^2)}{(r_2^2-r_1^2)(r_1^2-r_o^2)} \tag{2.4}$$

$\tau_1[r_1,r_2]$ = maximum shear stress in the inner cylinder, Pa

$$= p_o\left(\frac{r_2^2}{r_2^2-r_o^2}\right) - p_f\left(\frac{r_1^2}{r_1^2-r_o^2}\right) \tag{2.5}$$

$\tau_2[r_1,r_2]$ = maximum shear stress in the outer cylinder, Pa

$$= p_o\left(\frac{r_2^2(r_o^2/r_1^2)}{r_2^2-r_o^2}\right) + p_f\left(\frac{r_2^2}{r_2^2-r_1^2}\right) \qquad (2.6)$$

Optimum Solution

$$U_1^* = 0.0225 \qquad (2.7)$$

$$[\underline{X}^o]_1^* = [r_1,r_2,p_f]^T = [0.123,\ 0.150,\ 5.074\times10^6]^T$$

The radial (σ_r), tangential (σ_Θ) and shear (τ) stress distributions inside the optimum prestressed cylinders are shown in Figure 2.2. The objective function, U_1, contour lines and some constraints are plotted in the two dimensional domain r_1 and r_2 in Figure 2.3, where the third variable is kept constant at its optimum value p_{f1}^*. Point A designates the optimum vector $[\underline{X}^o]_1^*$. The constraints $\phi_{6,1}$ and $\phi_{7,1}$ are the only active constraints for this particular set of specifications.

2.2.2. Centering the Nominal Optimum Inside the Feasible Region $[P]_2$

The problem design variables are the same as $[P]_1$. In this stage, however, the optimization criterion differs. It is the minimization of the absolute difference between the maximum shear stresses in the two cylinders which in turn will guarantee maximum utilization of the space available

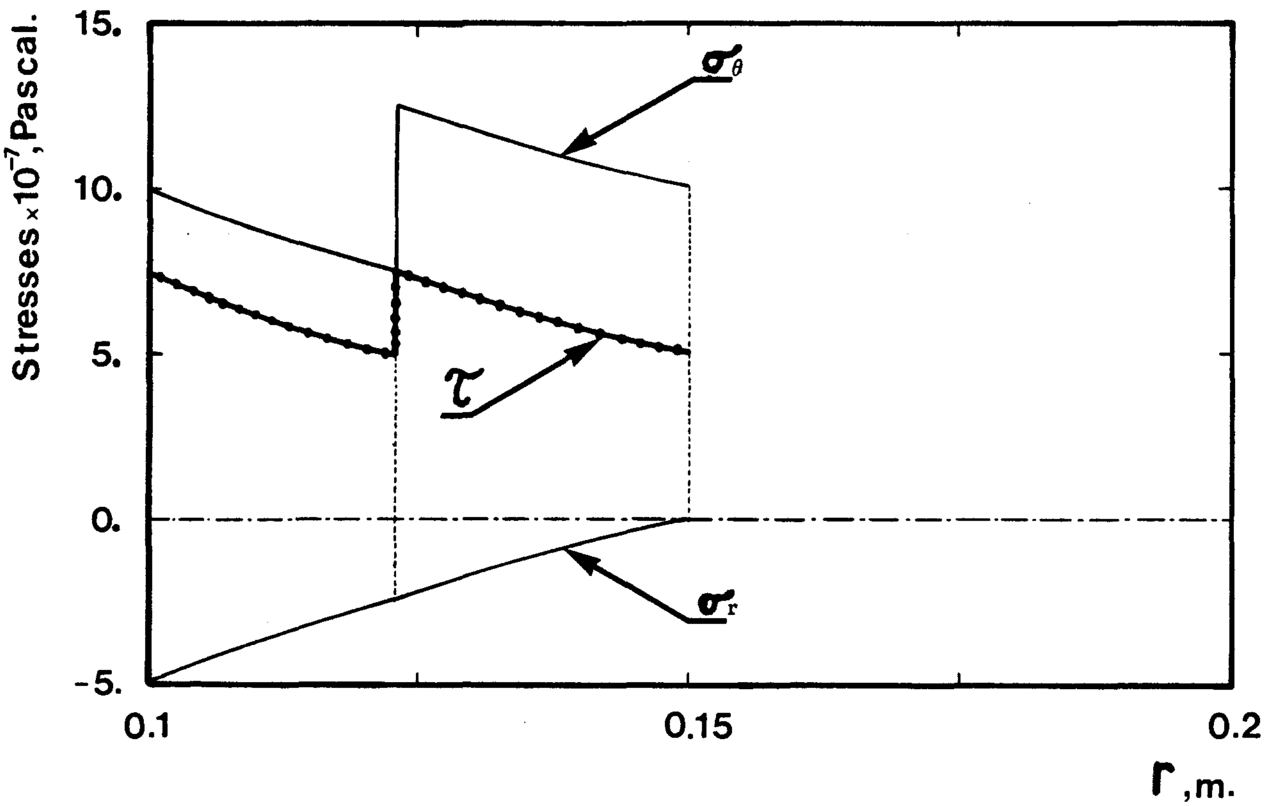


Figure 2.2 Radial (σ_r), tangential (σ_θ) and shear (τ) stresses of $[P]_1$.

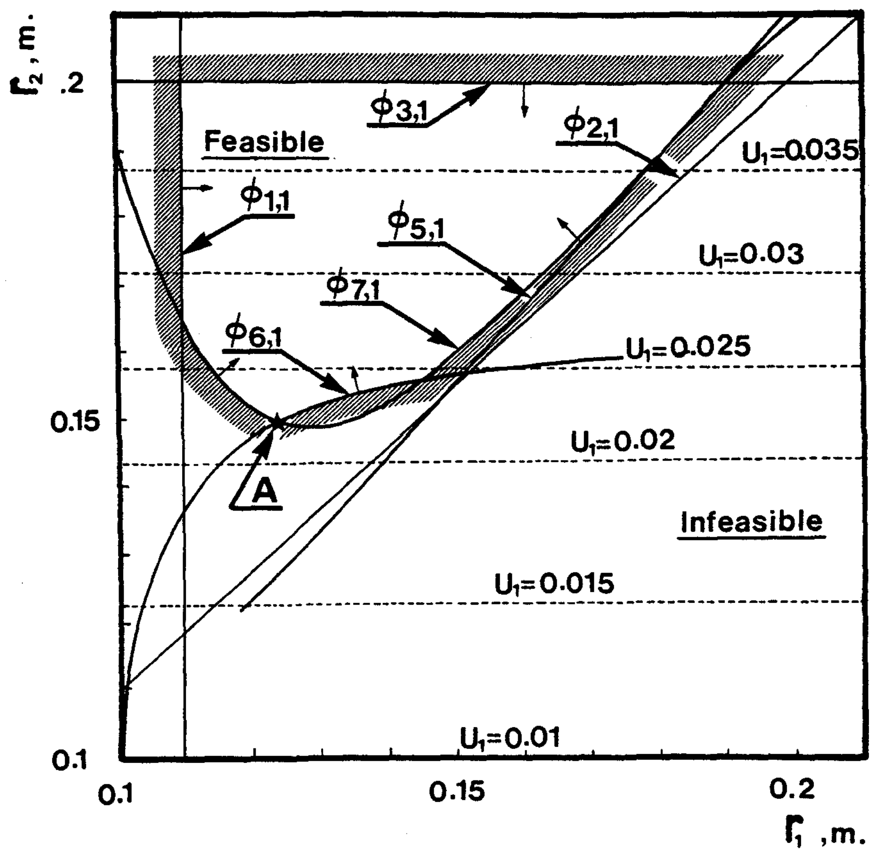


Figure 2.3 Feasible and unfeasible regions of $[P]_1$.

$(r_{max}-r_o)$ with the best stress distribution.

$$[\underline{X}^o]_2 = [\underline{X}^o]_1 = [r_1, r_2, p_f]^T$$

$$U_2 = |\tau_1 - \tau_2| = \text{minimum} \tag{2.8}$$

$$[\underline{\Phi}]_2 = [\underline{\Phi}]_1$$

Optimum Solution

$$U_2^* = 28.8 \tag{2.9}$$

$$[X^o]_2^* = [r_1, r_2, p_f]^T = [0.126,\ 0.185,\ 5.725\text{x}10^6]^T$$

The radial, tangential and shear stresses for the optimum design $[P]_2$ are shown in Figure 2.4. The objective function, U_2, contour lines and the optimum vector $[\underline{X}^o]_2^*$, designated as point B, are plotted in Figure 2.5, where p_f is taken as p_{f2}^*.

Apart from the difference in $[P]_1$ and $[P]_2$ design objectives, their optimum outcomes $[\underline{X}^o]_1^*$ and $[\underline{X}^o]_2^*$ are significantly different. The optimum design point B is centered inside the $[P]_2$ feasible region while point A is bounded by two constraints. This reduces the freedom of point A to deviate from its optimum value without violating the design specification. Point B, however, could possibly deviate and still be feasible even though it will increase the level of the design objective.

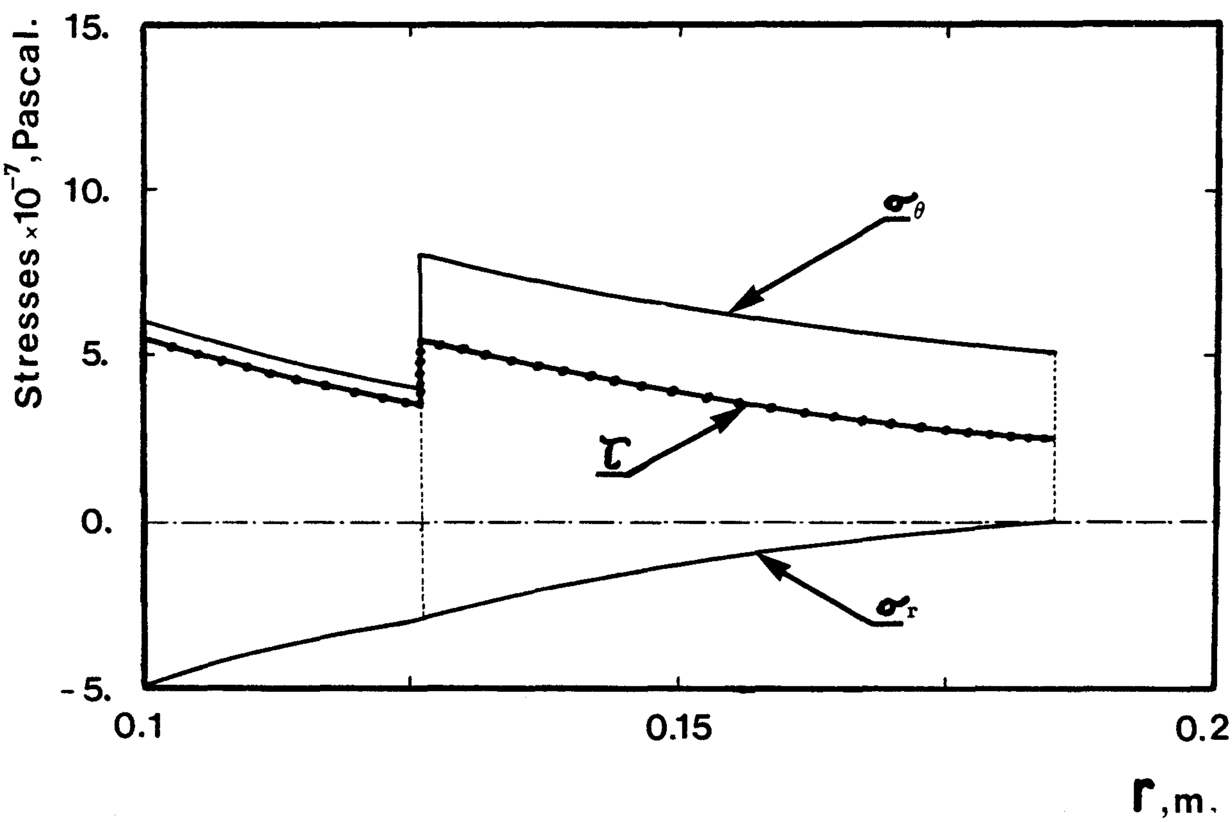


Figure 2.4 Radial (σ_r), tangential (σ_θ) and shear (τ) stresses of $[P]_2$.

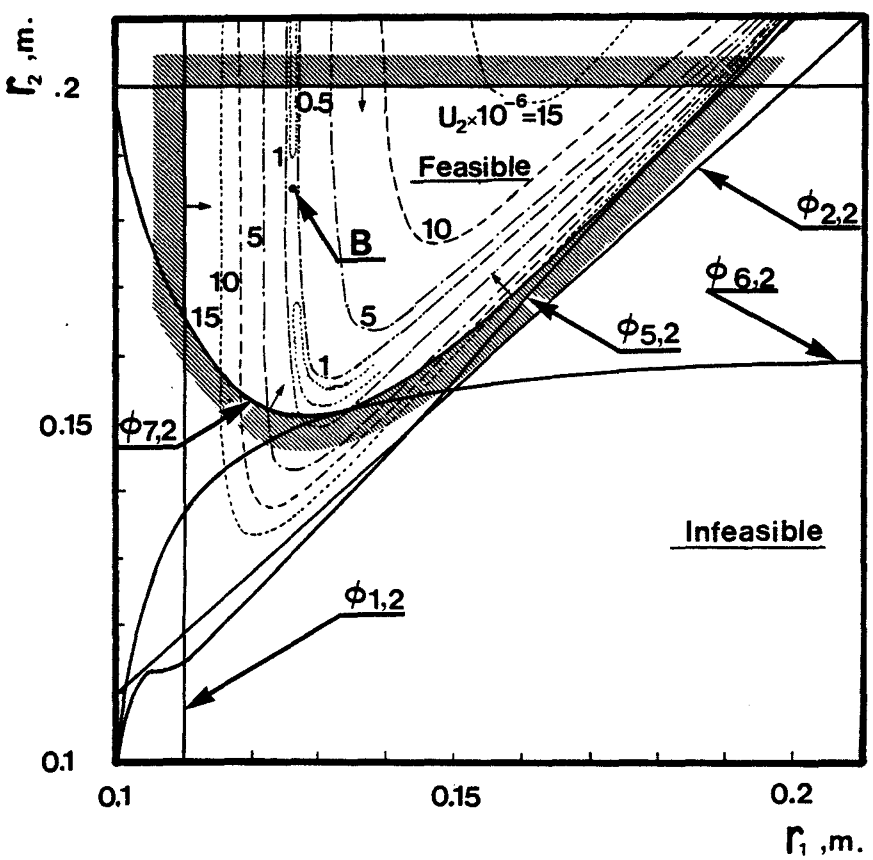


Figure 2.5 Feasible and infeasible regions of $[P]_2$.

2.2.3. Fixed Nominal-Optimum Symmetrical Tolerance $[P]_3$

At this stage we assume that the optimum nominal design variables, $[\underline{X}]_1^*$ or $[\underline{X}]_2^*$, are known and the question is what are the maximum feasible symmetrical tolerances which could be allocated to both r_1^* and r_2^* without violating any of the design constraints. That is to say the r_1, for instance, could feasibly take any value between $(r_1^*-t_1^*)$ and $(r_1^*+t_1^*)$, where t_1^* is the maximum associated tolerance to r_1^*. The problem objective, therefore, is to maximize t_1 and t_2 which could be done using one of several different optimization criterion, e.g., $(t_1^{-1}+t_2^{-1})$, $-(t_1+t_2)$, $(t_1 t_2)^{-1}$ and $(t_1+t_2)^{-1}$. However, all will lead to the same optimum hence they have the same objective of inflating a rectangle inside the feasible region.

$$[\underline{X}]_3 = [t_1, t_2]^T$$

$$U_3 = \frac{1}{t_1} + \frac{1}{t_2} = \text{minimum}$$

$$[\underline{\Phi}]_3 = \phi_{1,3} = (r_1^*-t_1) - r_o - t_{min} \geq 0$$

$$\phi_{2,3} = (r_2^*-t_2) - (r_1^*+t_1) - t_{min} \geq 0$$

$$\phi_{3,3} = r_{max} - (r_2^*+t_2) \geq 0$$

$$\phi_{4,3} = D_{max} - D_a[r_1^*-t_1),(r_2^*-t_2)] \geq 0 \qquad (2.10)$$

$$\phi_{5,3} = D_{max} - D_a[(r_1^*-t_1),(r_2^*+t_2)] \geq 0$$

$$\phi_{6,3} = D_{max} - D_a[(r_1^*+t_1),(r_2^*+t_2)] \geq 0$$

$$
\begin{aligned}
\phi_{7,3} &= D_{max} - D_a[(r_1^*+t_1),(r_2^*-t_2)] \geq 0 \\
\phi_{8,3} &= S - 2\tau_1[(r_1^*-t_1),(r_2^*-t_2)] \geq 0 \\
\phi_{9,3} &= S - 2\tau_1[(r_1^*-t_1),(r_2^*+t_2) \geq 0 \\
\phi_{10,3} &= S - 2\tau_1[(r_1^*+t_1),(r_2^*+t_2)] \geq 0 \\
\phi_{11,3} &= S - 2\tau_1[(r_1^*+t_1),(r_2^*-t_2)] \geq 0 \\
\phi_{12,3} &= S - 2\tau_2[(r_1^*-t_1),(r_2^*-t_2)] \geq 0 \\
\phi_{13,3} &= S - 2\tau_2[(r_1^*-t_1),(r_2^*+t_2)] \geq 0 \\
\phi_{14,3} &= S - 2\tau_2[(r_1^*+t_1),(r_2^*+t_2)] \geq 0 \\
\phi_{15,3} &= S - 2\tau_2[(r_1^*+t_1),(r_2^*-t_2)] \geq 0 \\
\phi_{16,3} &= t_1 \geq 0 \\
\phi_{17,3} &= t_2 \geq 0
\end{aligned}
\qquad (2.10)
$$

where $D_a[r_1,r_2]$, $\tau_1[r_1,r_2]$ and $\tau_2[r_1,r_2]$ are defined by Equations (2.4), (2.5) and (2.6), respectively.

The first three constraints, $\phi_{1,3}$, $\phi_{2,3}$ and $\phi_{3,3}$ are linear and are formulated on the basis of worst case design. $\phi_{1,3}$ actually represents two extreme cases

$$(r_1^*-t_1) - \text{Constant} \geq 0 \quad \ldots \text{(i)}$$

and

$$(r_1^*+t_1) - \text{Constant} \geq 0 \quad \ldots \text{(ii)}$$

The feasibility of (i) implies the satisfaction of (ii).

Therefore, (i) is the worst $\phi_{1,3}$. The same applies for the other two linear ϕ's. The sign of the worst tolerance is always the opposite of that of the associated nominal variable, [49], as will be proven in the Appendix. Because of the nonlinearity of $\phi_{5,1}$, $\phi_{6,1}$ and $\phi_{7,1}$, the four combinations of the two extreme values of r_1 and r_2 are now checked for each of these constraints, as described by $\phi_{4,3}$ through $\phi_{15,3}$. Even this simple but expensive method of checking a worst nonlinear constraint is only applicable if the constraint is convex for all the values of r_1 and r_2 between their extremes, which is the case for both $\phi_{5,1}$ and $\phi_{7,1}$. The problem of convexity will be discussed later; also mathematical verification and a suggestion for a scheme to reduce the number of ϕ's will be mentioned. This concept of checking corners may be understood by referring to Figure 2.6.

Optimum Solutions

1. When using $[\underline{X}]^*_1$ as the fixed optimum nominal design

$$U^*_{3,1} = 4\text{x}10^6$$

$$[\underline{X}]^*_{3,1} = [t_1, t_2]^T = [0.5\text{x}10^{-6},\ 0.4\text{x}10^{-6}]^T \qquad (2.11)$$

2. When using $[\underline{X}]^*_2$ as the fixed optimum nominal design

$$U^*_{3,2} = 1.3\text{x}10^2$$

$$[\underline{X}]^*_{3,2} = [t_1, t_2]^T = [0.016,\ 0.015]^T \qquad (2.12)$$

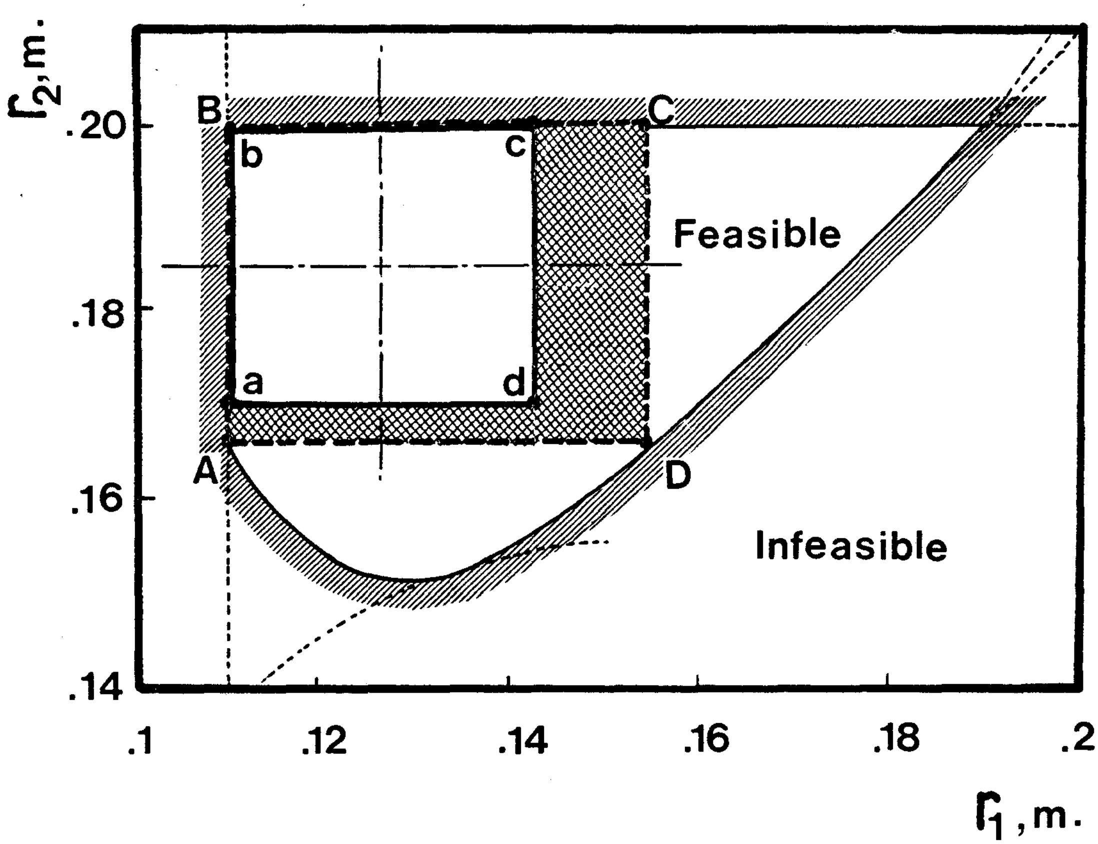


Figure 2.6 $[P]_3$ and $[P]_4$ optimum inflated toleranced solution.

The following remarks can be made about $[P]_3$ optimum solution.

(a) Due to truncation error, inefficiency of the optimization strategy and its stopping criterion, and the limited capacity of the computer word, $U^*_{3,1}$ did not reach to infinity and $[\underline{X}]^*_{3,1}$ did not become a true null vector, as it should have because of the location of $[\underline{X}]^*_1$. This also assumes that $[\underline{X}]^*_1$ is an exact optimum, (i.e., $\phi_{6,1}$ and $\phi_{7,1} = 0.0$), which is not the case because of the same mentioned reasons.

(b) The optimum solution of $[P]_3$ does not depend on the weighting factors in the objective function U_3. Thus if

$$U_3 = \sum_{i=1}^{N} \left(\frac{w_i}{t_i}\right),$$

U^*_3 and $[\underline{X}]^*_3$ will be the same irrespective of the values of w_i's.

(c) The rectangle a,b,c,d shown in Figure 2.6 describes the optimum solution of $[P]_3\big|_{[\underline{X}]^*_2}$ in which r_1 could take any value between 0.11 m and 0.142 m and r_2 could take any value between 0.17 m and 0.2 m. Figure 2.7 displays the shear stress distribution for the four extreme feasible design cases a,b,c and d. Any combination of r_1 and r_2, however, if chosen inside the feasible rectangle could be considered as an optimum design in a sense that it will withstand the applied internal pressure and follow the other specifications and is still manufactured with the maximum equally deviated tolerance.

(d) There are three main unrealistic approximations in $[P]_3$. They are:

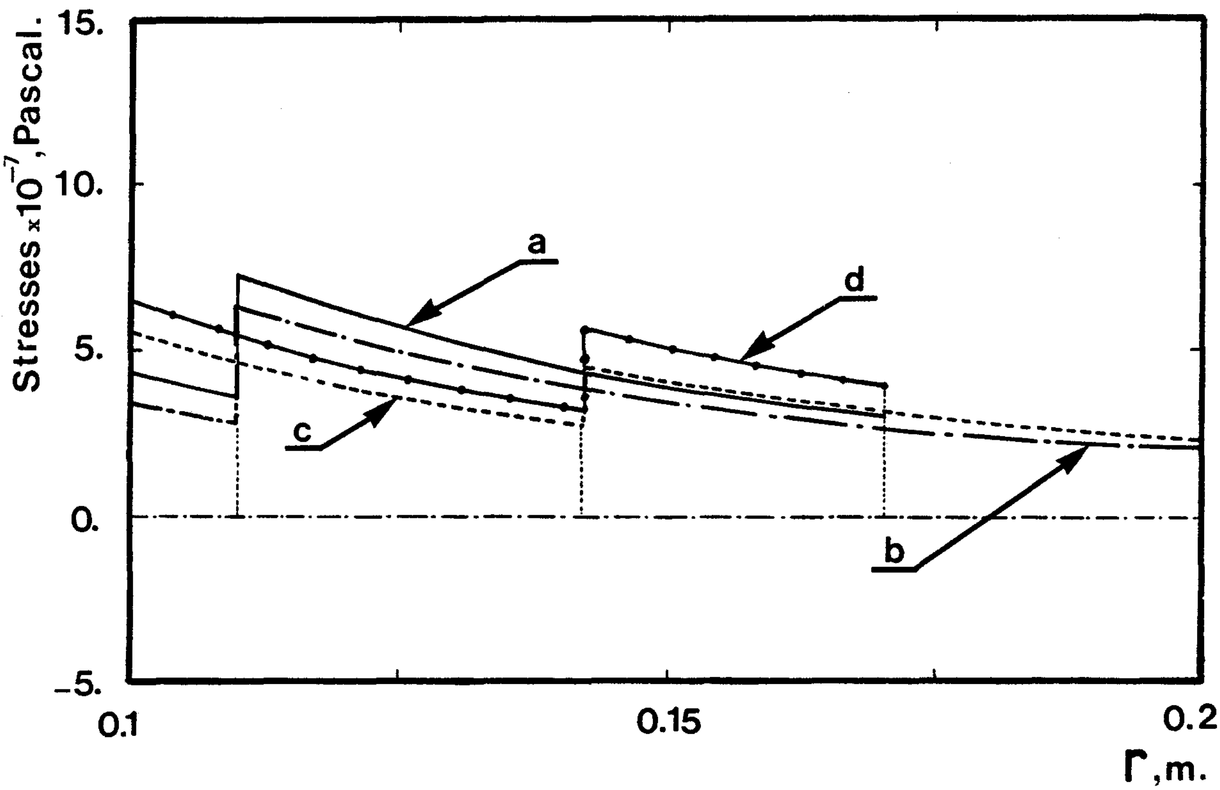


Figure 2.7 $[P]_3$ shear stresses distribution for the four worst designs a, b, c and d.

(i) the tolerances t_1 and t_2 are assumed to be equally deviated from the mean values of r_1 and r_2, respectively, which is not necessarily true.

(ii) the optimum mean values r_1^* and r_2^* are assumed fixed and have no influence on the objective function.

(iii) the objective function served its purpose of inflating the feasible tolerance rectangle, even though it did not include the increase in assembly cost which is due to the increase in the interference between the cylinders. This, in turn, is due to the increase in r_1 and r_2.

2.2.4 Fixed Nominal-Optimum Non-Symmetrical Tolerance , $[P]_4$

The optimum nominal design variables $[\underline{X}^o]_2^*$ are assumed known and the problem objective is to allocate the maximum feasible tolerance rectangle by maximizing both the negative and the positive tolerances for each of the radii r_1^* and r_2^*.

$$[\underline{X}]_4 = [t_1^+,\ t_1^-,\ t_2^+,\ t_2^-]^T$$

$$U_4 = \frac{1}{t_1^+ + t_1^-} + \frac{1}{t_2^+ + t_2^-} = \text{minimum} \qquad (2.13)$$

$$[\underline{\Phi}]_4 = \begin{bmatrix} [\underline{\Phi}]_{31} \\ \phi_{18,4} = t_1^- \geq 0 \\ \phi_{19,4} = t_2^- \geq 0 \end{bmatrix}$$

where, $[\underline{\Phi}(t_1^+, -t_1^-, t_2^+, -t_2^-)]_{31} = [\underline{\Phi}(t_1, -t_1, t_2, -t_2)]_3$

e.g. $\phi_{2,4} = \phi_{2,31} = (r_2^* - t_2^-) - (r_1^* + t_1^+) - t_{min} \geq 0$

Optimum Solution

$$U_4^* = 52.14 \tag{2.14}$$

$$[\underline{X}]_4^* = [t_1^+, t_1^-, t_2^+, t_2^-]^T = [0.028,\ 0.016,\ 0.015,\ 0.019]$$

The optimum tolerance rectangle is ABCD in Figure 2.6 and gives an increase of about 56% in the tolerance area. Figure 2.8 shows the shear stress distribution for the extreme design cases.

The optimization method used in this work was based on a random adaptive search strategy followed by accelerated pattern moves, ADRANS [50]. A constrained optimization problem, in general, consists of an objective function and a set of equality and inequality constraints. It can be transformed, using penalty or barrier functions, and expressed by an artificial unconstrained objective function. ADRANS falls into the category of direct search methods, which do not rely explicitly on evaluation or estimation of partial derivatives of the artificial objective function at any point. Therefore, it is not a prerequisite for the problem objective function and the subjected constraint equations to be continuous over the range of the design variables. Consequently, the number of inequality constraints in $[\underline{\Phi}]_4$ could be decreased from nineteen to ten by checking only the worst corner of the design rectangle. For example,

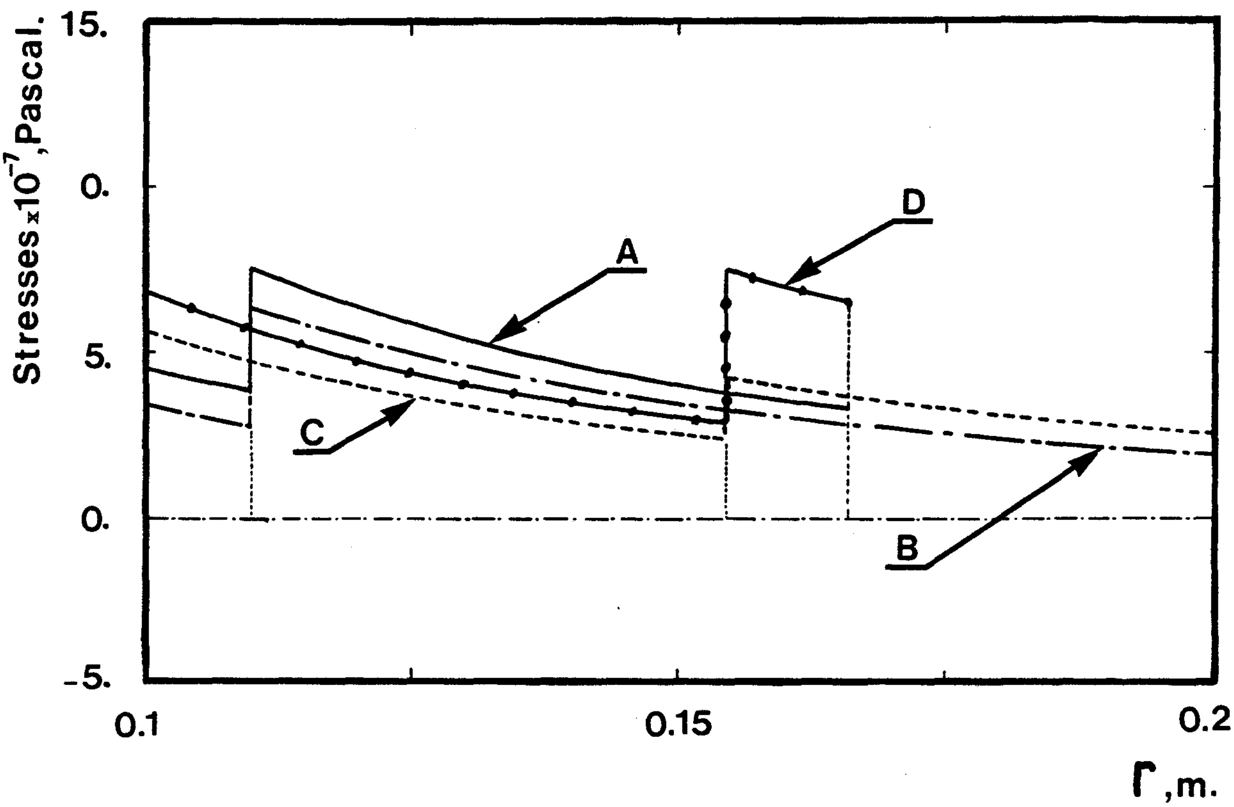


Figure 2.8 $[P]_4$ shear stresses distribution for the four worst designs A, B, C and D.

$$\phi_{3,4} = D_{max} - D_a[(r_1^* - t_1^-),(r_2^* - t_2^-)_A \geq 0$$

$$\phi_{4,4} = D_{max} - D_a[(r_1^* - t_1^-),(r_2^* + t_2^+)]_B \geq 0$$

$$\phi_{5,4} = D_{max} - D_a[(r_1^* + t_1^+),(r_2^* + t_2^+)]_C \geq 0$$

and $$\phi_{6,4} = D_{max} - D_a[(r_1^* + t_1^+),(r_2^* - t_2^-)]_D \geq 0$$

could be replaced with

$$\phi'_{3,4} = D_{max} - \mathrm{Max}[D_a(A,B,C\&D)] \geq 0 \qquad (2.15)$$

2.2.5. Optimum Nominal and Symmetrical Tolerance Allocation, $[P]_5$

Stage 5 design variables, $[\underline{X}]_5$, are the combination of $[\underline{X}^o]_1$ and $[\underline{X}]_3$ where the nominal design variables, $[r_1,r_2,p_f]^T$, are not fixed but are allowed to adjust their values to minimize the objective function, U_5, which consists of four parts to express the relationship between the relative costs of product material, machining and assembly.

$$[\underline{X}]_5 = [r_1^o,r_2^o,p_f,t_1,t_2]^T$$

$$U_5 = u_{15} + u_{25} + (u_{35} + u'_{35}) + u_{45} = \text{minimum}$$

$$[\underline{\Phi}]_5 = \phi_{1,5} = (r_1^o - t_1) - r_o - t_{min} \geq 0$$

$$\phi_{2,5} = (r_2^o - t_2) - (r_1^o + t_1) - t_{min} \geq 0 \qquad (2.16)$$

$$\phi_{3,5} = r_{max} - (r_2^o + t_2) \geq 0$$

$$\phi_{4,5} = D_{max} - \mathrm{Max}[D_a(r_1,r_2)_a, D_a(r_1,r_2)_b, D_a(r_1,r_2)_c, D_a(r_1,r_2)_d] \geq 0$$

$$\phi_{5,5} = S - 2\text{Max}[\tau_1(r_1,r_2)_a,\tau_1(r_1,r_2)_b,\tau_1(r_1,r_2)_c,\tau_1(r_1,r_2)_d] \geq 0$$

$$\phi_{6,5} = S - 2\text{Max}[\tau_2(r_1,r_2)_a,\tau_2(r_1,r_2)_b,\tau_2(r_1,r_2)_c,\tau_2(r_1,r_2)_d] \geq 0$$

$$\phi_{7,5} = p_f \geq 0 \qquad (2.16)$$

$$\phi_{8,5} = t_1 \geq 0$$

$$\phi_{9,5} = t_2 \geq 0$$

where,

$$(r_1,r_2)_a = [(r_1^o-t_1),(r_2^o-t_2)]$$

$$(r_1,r_2)_b = [(r_1^o-t_1),(r_2^o+t_2)]$$

$$(r_1,r_2)_c = [(r_1^o+t_1),(r_2^o+t_2)] \qquad (2.17)$$

$$(r_1,r_2)_d = [(r_1^o+t_1),(r_2^o-t_2)]$$

u_{15} = material cost

$$= c_{11}[(r_2^o+t_2)^2-r_o^2]$$

where $c_{11} = 32.0$

u_{25} = inner cylindrical surface machining cost

$$= c_{21}(2t_1)^{c_{22}}.\ \exp(c_{23}.2t_1) \qquad (2.18)$$

where $[c_{21},c_{22},c_{23}] = [12.8,-0.458,-0.0707]$

$$u_{35} = \text{outer cylindrical surface machining cost}$$
$$= c_{31}(2t_1)^{c_{32}} . \exp(c_{33}.2t_1)$$

where $[c_{31},c_{32},c_{33}] = [1.3,-0.63,-0.17\times10^{-8}]$

$$u'_{35} = c_{31}(2t_2)^{c_{32}} . \exp(c_{33}.2t_2) \quad (2.18)$$

$$u_{45} = \text{assembly cost}$$

$$= c_{41}.\{\text{Max}[D_a(r_1,r_2)]_j.c_{42}\}^{c_{43}} , \; j=i,ii,iii \text{ and } iv$$

where $[c_{41},c_{42},c_{43}] = [3.16,10^3,1.72]$

The c's are the best fitted cost model coefficients which were determined using a nonlinear least squares technique [51]. The tolerance cost ratio for the different machining conditions are estimated after Peat [52].

The cost represented by u_{25} and u_{35} models include the actual time taken to produce a completely acceptable element. Gauges, tools and fixture costs, overhead cost, inspection cost, etc. are also included in the cost models. Figure 2.9 shows u_{25} and u_{35} as a function of a product tolerance. Tolerance cost models must cover the cost characteristics for individual processes as well as for a sequence of processes. For all of the cost models discussed herein, (u_{25}, u_{35} and u_{45}), it was assumed that the same process must be used independently of the precision requirements. Also, 'scrapping cost', which

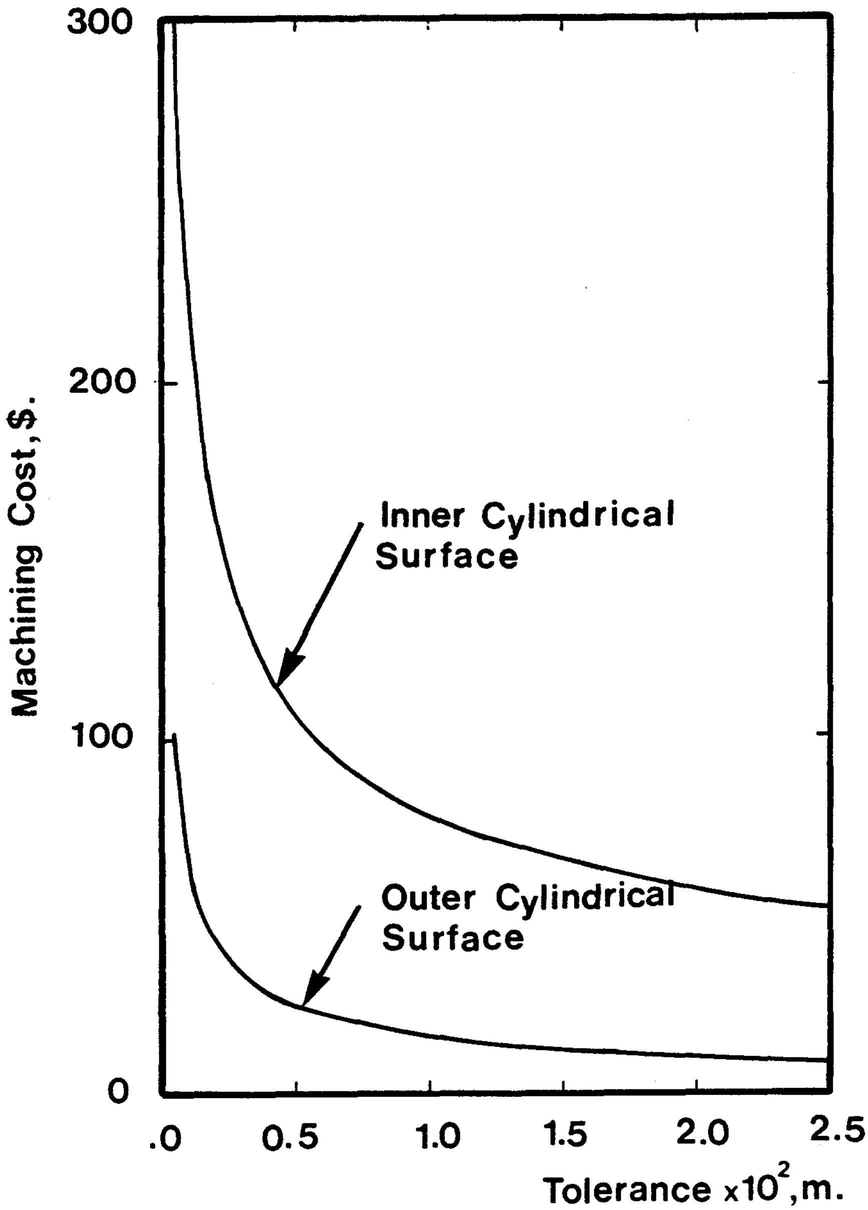


Figure 2.9 Inner and outer cylindrical surfaces machining cost models.

consists either of the cost of repairing or the cost of the whole part, if it has to be rejected, will not be taken into account. Hence we are interested only in full acceptability, and there is no part that will fall outside the specified feasible tolerance.

Optimum Solution

$$U_5^* = 80.5 \qquad (2.19)$$

$$[\underline{X}]_5^* = [r_1, r_2, p_f, t_1, t_2]^T = [0.1376, 0.1876, 2.534\text{x}10^6, 0.0276, 0.01237]^T$$

$$0.110 \leq r_1 \leq 0.165$$

$$0.175 \leq r_2 \leq 0.200$$

Figure 2.10 shows the shear stress distribution for the corners of the optimum tolerance square. The $[P]_5$ tolerance area $(4t_1t_2)$ is less than that of $[P]_3$ or $[P]_4$. However, if the $[P]_5$ optimization criterion, U_5, is used as a base of comparison, there will be an increase in its value of 7.51 and 19.5%, when using $[\underline{X}]_3^*$ and $[\underline{X}]_4^*$, respectively, as the design solutions.

2.2.6. Optimum Nominal and Non-Symmetrical Tolerance Allocation, $[P]_6$

$[P]_6$ design variables are a combination of $[\underline{X}]_1$ and $[\underline{X}]_4$; and its objective function, U_6, consists of the same

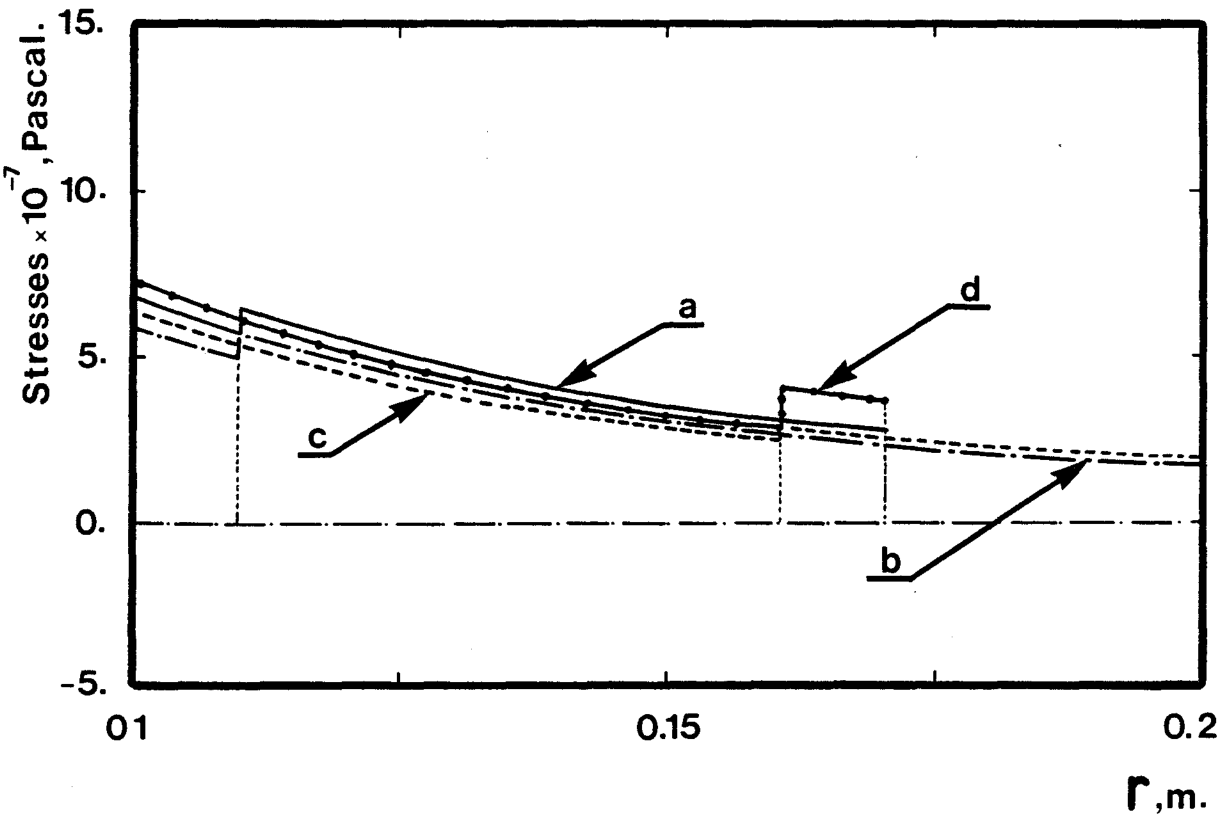


Figure 2.10 $[P]_5$ shear stresses distribution for the four worst designs a, b, c and d.

four parts as that of U_5, but taking the non-symmetrical tolerance distribution into consideration. The problem, therefore, could be restated as:

$$
\begin{aligned}
[\underline{X}]_6 &= [r_1^o, r_2^o, p_f, t_1^+, t_1^-, t_2^+, t_2^-]^T \\
U_6 &= u_{16} + u_{26} + (u_{36} + u_{36}') + u_{46} = \text{minimum} \\
u_{16} &= c_{11}[(r_2^o + t_2^+)^2 - r_o^2] \\
u_{26} &= c_{21} \cdot (t_1^+ + t_1^-)^{c_{22}} \cdot \exp[(t_1^+ + t_1^-) \cdot c_{23}] \\
u_{36} &= c_{31} \cdot (t_1^+ + t_1^-)^{c_{32}} \cdot \exp[(t_1^+ + t_1^-) \cdot c_{33}] \\
u_{36}' &= c_{31} \cdot (t_2^+ + t_2^-)^{c_{32}} \cdot \exp[(t_2^+ + t_2^-) \cdot c_{33}] \\
u_{46} &= c_{41} \cdot [\text{Max } D_a(r_1, r_2)_j \cdot c_{42}]^{c_{43}}
\end{aligned}
\qquad (2.20)
$$

$$
[\underline{\Phi}]_6 = \begin{bmatrix} [\underline{\Phi}]_5' \\ \\ \phi_{10,5} = t_1^- \geq 0 \\ \phi_{11,5} = t_2^- > 0 \end{bmatrix}
$$

where $[\underline{\phi}(t_1^+, -t_1^-, t_2^+, t_2^-)]_5' = [\underline{\phi}(t_1, -t_1, t_2, -t_2)]_5$

Optimum Solution

$$U_6^* = 80.7$$

$$[\underline{X}]_6^* = [0.135, 0.182, 1.3\text{x}10^6, 0.026, 0.025, 0.018, 0.012]^T$$

$$0.110 \le r_1 \le 0.161 \qquad (2.21)$$

$$0.170 \le r_2 \le 0.200$$

There is an increase of 12% in the tolerance area of $[P]_6$ over that of $[P]_5$. The following data gives the four worst cases and illustrates the range in values that are possible.

State	r_1 (m)	r_2 (m)	Δ (10^{-4}m)	u_1 (10^{-4}m)	u_2 (10^{-4}m)
I	0.11	0.17	0.0918	0.0725	0.0193
II	0.11	0.20	0.0878	0.0725	0.0153
III	0.16	0.20	0.0716	0.0204	0.0512
IV	0.16	0.17	0.1868	0.0204	0.1764

where r_{11} = the manufactured outer radius of the inner cylinder = $r_1 + u_1$ (2.22a)

r_{12} = the manufactured inner radius of the outer cylinder = $r_1 - u_2$ (2.22b)

u_1 = the inner cylinder displacement due to p_f

$$= r_1 \left[\frac{r_1^2+r_0^2}{r_1^2-r_0^2} - \nu\right] p_f/E \qquad (2.22c)$$

u_2 = the outer cylinder displacement due to p_f

$$= r_1 \left[\frac{r_2^2 + r_1^2}{r_2^2 - r_1^2} + \nu \right] p_f/E \tag{2.22d}$$

Δ = total interference between the inner and outer cylinders $= u_1 + u_2$ (2.22e)

The four distinct design states shown in the previous table are the corners of the inflated optimum rectangle, and their shear stress distributions are plotted in Figure 2.11.

Any assembled design, having dimensions between the upper and lower limits, is an optimum feasible design. This is only true, however, for selectively assembled cylinders, where for each feasible inner cylinder a matched feasible outer cylinder must be chosen having just the right amount of interference fit. Even though the selectively assembled cylinders may be considered as a feasible solution it is not practical since an additional overhead cost has to be added to the finished cost, which was neglected in the previous optimization. It will therefore be necessary to define a separate r_1 for the two cylinders, and the interference pressure, p_f, becomes a state variable.

2.2.7. Micro Design Optimum Nominal and Non-symmetrical Tolerance Allocation, $[P]_7$

The basic design variables in the previous six stages of analysis were r_1, r_2 and p_f. In the ensuing analysis,

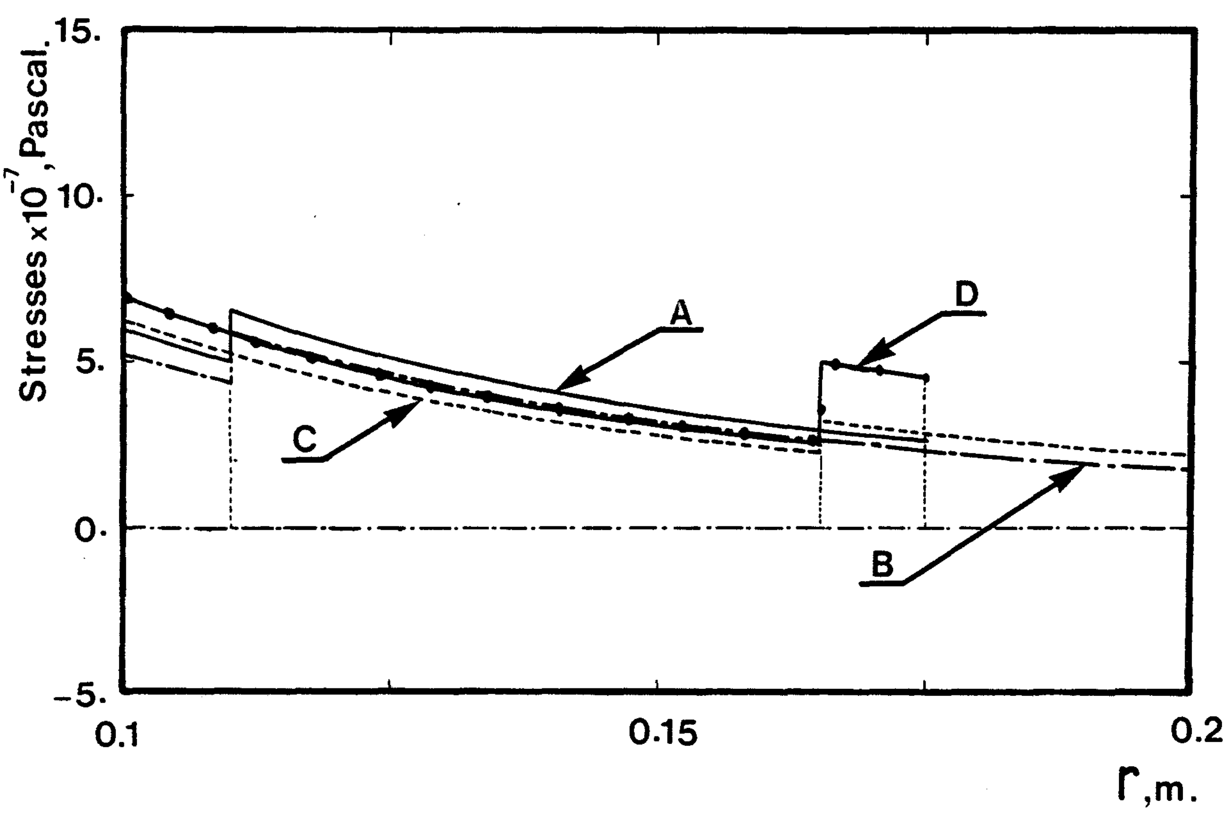


Figure 2.11 $[P]_6$ shear stresses distribution for the four worst designs A, B, C and D.

however, r_{11}, r_{12} and r_2, shown in Figure 2.12 and defined in Equations (2.22), will be taken as the problem basic design variables. In all of the preceding analysis we were concerned about determining the overall assembled structure (system) design variables and their tolerances, and this analysis could be classified as "Macro Design". On the other hand, in this stage, $[P]_7$, we are interested in the unassembled components optimum design variables and their associated positive and negative tolerances. Therefore, it could be classified as "Micro Design". Selective assembly will not now be required.

The $[P]_7$ design variables vector could be expressed as

$$[\underline{X}']_7 = [\mathring{r}_{11}, \mathring{r}_{12}, \mathring{r}_2, t^+_{11}, t^+_{12}, t^-_{12}, t^+_2, t^-_2]^T \qquad (2.23)$$

To satisfy the stress constraints the maximum shear stresses in the inner, τ_1, and outer, τ_2, cylinders must be computed. τ_1 and τ_2, however, are functions of r_1, among other variables, and not functions of either r_{11} or r_{12}. We can obtain the following expression by solving Equations (2.22a,b,c and d).

$$r_1 = f(r_1) = r_{11} - \left[\frac{(1-\nu)r_1^2+(1+\nu)r_o^2}{(r_1^2-r_o^2)}\right] \cdot \left[\frac{(r_1-r_{12})(r_2^2-r_1^2)}{(1+\nu)r_2^2+(1-\nu)r_1^2}\right] \qquad (2.24)$$

and a numerical analysis technique can be used, e.g., Newton Raphson, to determine the five roots of r_1. The iterative procedure, however, is costly, and conversion is not guaranteed, since it depends upon the iteration starting value of r_1. To avoid this mathematical inadequacy another

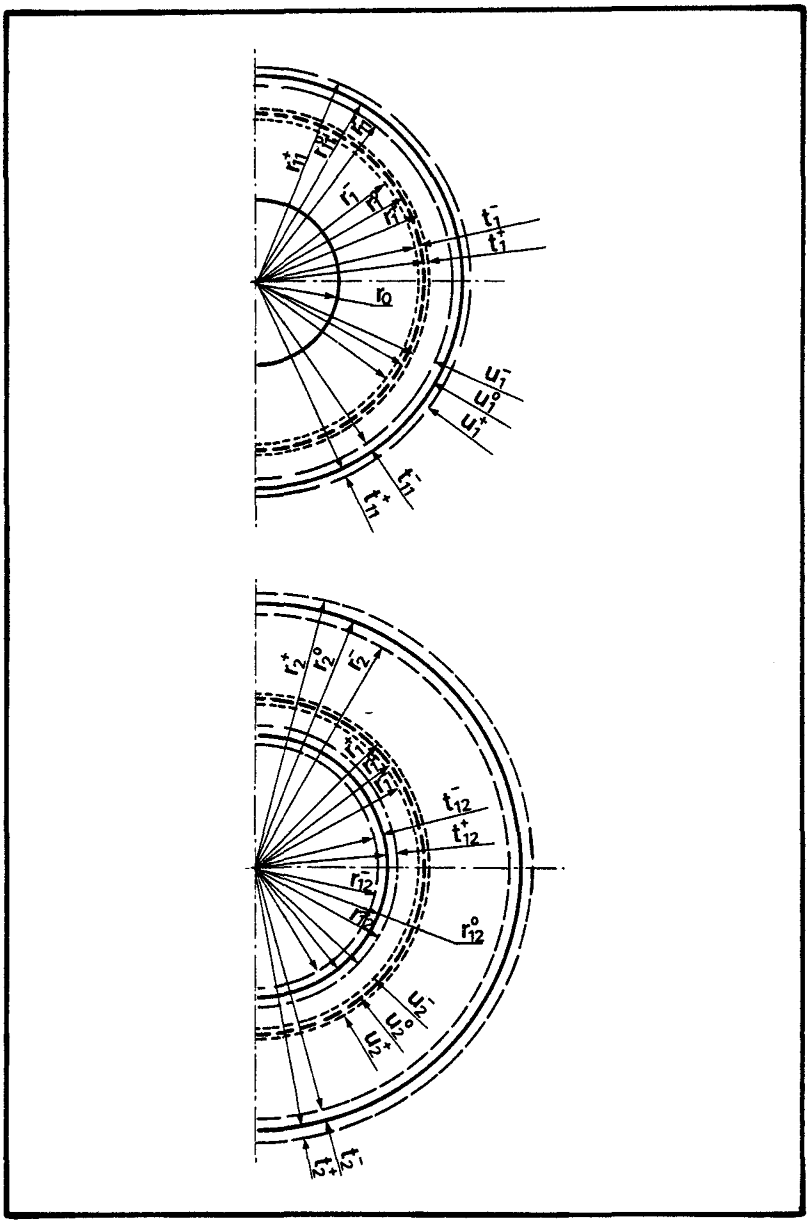


Figure 2.12 The inner and outer cylinders and their dimensional tolerance.

set of design variables is chosen, $[\underline{X}]_7$, in which r_1^o and r_2^o are mean radii and u_1^o is the mean interference.

$$[\underline{X}]_7 = [r_1^o, r_2^o, u_1^o, t_1^+, t_1^-, t_{11}^+, t_{11}^-, t_2^+, t_2^-]^T$$

$$\begin{aligned}
U_7 &= u_{17}+u_{27}+(u_{37}+u'_{37})+u_{47} = \text{minimum} \\
u_{17} &= c_{11}[(r_2^o+t_2^+)^2 - r_o^2] \\
u_{27} &= c_{21}(t_{12}^+ + t_{12}^-)^{C_{22}}. \exp[c_{23}(t_{12}^+ + t_{12}^-)]/100. \\
u_{37} &= c_{31}(t_{11}^+ + t_{11}^-)^{C_{32}}. \exp[c_{33}(t_{11}^+ + t_{11}^-)/100. \\
u'_{37} &= c_{31}(t_{12}^+ + t_{12}^-)^{C_{32}}.\exp[c_{33}(t_{12}^+ + t_{12}^-)]/100. \\
u_{47} &= c_{41}[\{(r_{11}^o + t_{11}^+) - (r_{12}^o - t_{12}^-)\}.c_{42}]^{C_{43}}
\end{aligned} \tag{2.25}$$

where $c_{11}, c_{21}, \ldots c_{43}$ are defined in Equation (2.18).

$$\begin{aligned}
[\underline{\phi}]_7 = \phi_{17} &= (r_{11}^o - t_{11}^-) - r_o - t_{min} \geq 0 \\
\phi_{27} &= (r_2^o - t_2^-) - (r_{12}^o + t_{12}^+) - t_{min} \geq 0 \\
\phi_{37} &= r_{max} - (r_2^o + t_2^+) \geq 0 \\
\phi_{47} &= (r_{11}^o - t_{11}^-) - (r_{12}^o + t_{12}^+) \geq 0 \\
\phi_{57} &= D_{max} - [(r_{11} + t_{11}^+) - (r_{12} - t_{12}^-)] \geq 0 \\
\phi_{67} &= S - 2\,\text{Max}[\tau_1(r_1,r_2)_i, \tau_1(r_1,r_2)_{ii}, \\
&\quad \tau_1(r_1,r_2)_{iii}, \tau_1(r_1,r_2)_{iv}] \geq 0 \\
\phi_{77} &= S - 2\,\text{Max}[\tau_2(r_1,r_2)_i, \tau_2(r_1,r_2)_{ii}, \\
&\quad \tau_2(r_1,r_2)_{iii}, \tau_2(r_1,r_2)_{iv}] \geq 0
\end{aligned} \tag{2.26}$$

$$\phi_{87} = t_{12}^+ \geq 0, \quad \phi_{97} = t_{12}^- \geq 0, \quad \phi_{107} = t_{u2}^+ \geq 0, \quad \phi_{117} = t_{u2}^- \geq 0$$

where, $r_1^- = r_1^o - t_1^-$, $r_1^+ = r_1^o + t_1^+$

$r_2^- = r_2^o - t_2^-$, $r_2^+ = r_2^o + t_2^+$

$r_{11}^o = r_1^o + u_1^o$

$r_{11}^- = r_{11}^o - t_{11}^-$, $r_{11}^+ = r_{11}^o + t_{11}^+$

$u_1^- = r_{11}^- - r_1^+$, $u_1^+ = r_{11}^+ - r_1^-$ (see Figure 2.12)

$u_2^o = u_1^o . A(r_1^o, r_2^o)$

$u_2^- = u_1^- . \text{Min}[A(r_1, r_2)_{i,ii,iii,iv}]$ (2.27)

$u_2^+ = u_1^+ . \text{Max}[A(r_1, r_2)_{i,ii,iii,iv}]$

$r_{12}^o = r_1^o - u_2^o$

$r_{12}^- = r_1^+ - u_2^+$, $r_{12}^+ = r_1^- - u_2^-$

$t_{12}^- = r_{12}^o - r_{12}^-$, $t_{12}^+ = r_{12}^+ - r_{12}^o$

$$A(r_1 , r_2) = \left[\frac{(1+\nu) r_2^2 + (1-\nu) r_1^2}{(1+\nu) r_o^2 + (1-\nu) r_1^2} \cdot \frac{r_1^2 - r_o^2}{r_2^2 - r_1^2}\right]$$

$\tau_1(r_1, r_2)$ and $\tau_2(r_1, r_2)$ are defined in Equations (2.5) and (2.6), respectively, where,

$$p_{f1} = u_1^- \; E/[r_1 . \{\frac{r_1^2 + r_o^2}{r_1^2 - r_o^2} - \nu\}]$$

= p_f used in calculating the worst case of $\tau_1(r_1, r_2)$

$$p_{f2} = u_1^+ E/[r_1\{\frac{r_1^2+r_o^2}{r_1^2-r_o^2} - \nu\}] \quad (2.28)$$

$= p_f$ used in calculating the worst case of $\tau_2(r_1,r_2)$.

Optimum Solution

In the following table three distinct optimum designs are presented, which illustrate the sensitivity of the solution to changes in the internal pressure, p_o, and the inner cylindrical surface machining cost, u_{27}.

	Case 1	Case 2	Case 3
$p_o x10^{-6}$ Pa	5	5	6
u_{27}' =	u_{27}	$10xu_{27}$	u_{27}
U_7 \$	43	144	47
r_{11} (mm)	$137.747^{+0.014}_{-0.012}$	$164.632^{+0.0}_{-0.012}$	$147.945^{+0.019}_{-0.006}$
r_{12} (mm)	$137.716^{+0.019}_{-0.023}$	$164.559^{+0.054}_{-0.014}$	$147.917^{+0.009}_{-0.031}$
r_2 (mm)	$193.192^{+0.023}_{-0.022}$	$189.676^{+0.015}_{-0.018}$	$199.348^{+0.039}_{-0.015}$
t_{11} (mm) $(t_{11}^+ + t_{11}^-)$	0.026	0.012	0.025
t_{12} (mm) $(t_{12}^+ + t_{12}^-)$	0.042	0.068	0.040
t_2 (mm) $(t_2^+ + t_2^-)$	0.045	0.033	0.054

The shear stress distributions of the optimum nominal designs, $[\underline{X}_2^{o*}]_7$, for each of the three cases are plotted in Figure 2.13.

2.3 Mathematical Generalization

2.3.1 The Nominal Optimization Problem, $[\underline{X}^{o*}]$

The nominal optimization problem could be stated as follows: Minimize a scalar objective function, U, of n continuous independent design variables, $\underline{X}^o$, subject to a set of m inequality constraints, $\underline{\Phi}\ (\underline{X}^o) \geq \underline{0}$ where $\underline{0}$ is a zero vector. Therefore, the problem can be compactly written as

$$\text{minimize } U(\underline{X}^o) \tag{2.29}$$

subject to $\underline{\Phi}\ (\underline{X}^o) \geq \underline{0}$

where
$$\underline{X}^o = [x_1^o, x_2^o, \ldots x_n^o]^T$$
$$\underline{\Phi}\ (\underline{X}^o) = [\phi_1(\underline{X}^o),\ \phi_2(\underline{X}^o), \ldots \phi_m(\underline{X}^o)]^T$$
$$\underline{0} = [0,0,\ldots 0]^T$$

The feasible region, R_c, consists of a set of feasible points $\underline{X}^o$ which satisfies the constraint vector, $\underline{\Phi}$. It may be succinctly defined as

$$R_c = \{\underline{X}^o \mid \underline{\Phi}\ (\underline{X}^o) \geq \underline{0}\}$$

R_c is assumed to be a closed region, i.e., it contains all

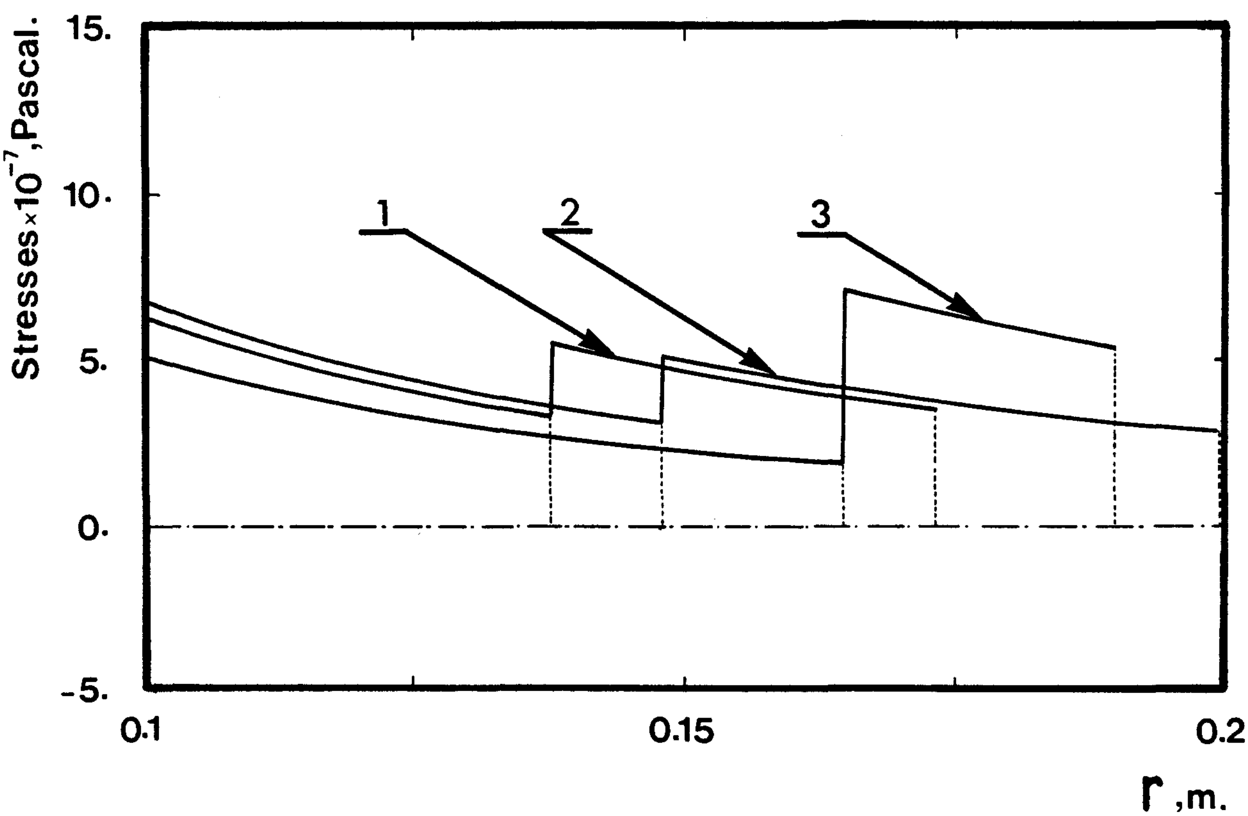


Figure 2.13 $[P]_7$ shear stresses distribution for case 1, 2 and 3 designs.

its boundary points. Therefore, R_c could further be defined as

$$R_c = \{\underline{X}^o \mid \phi_i(\underline{X}^o) \geq 0, \ i=1,2,\ldots m_a\} \qquad (2.30)$$

where m_a is the number of active constraints defining the closed region.

Problem (2.25) could be solved by employing a multi-dimensional optimization strategy.

2.3.2 The Toleranced Design Optimization Problem [$\underline{X}^*$]

The optimum toleranced problem is the combination of the nominal optimization design and the tolerance assignment problems. The nominal base design point [$\underline{X}^o$] is allowed to be allocated optimally as well as its associated tolerances so that the cost objective function, $U(\underline{X})$, will be at its minimum. The problem sometimes is known as "Optimal Design Centering" to distinguish it from the tolerance assignment problem, [$\underline{T}^*$], in which the nominal base design point is considered fixed at an arbitrary point. It is usually the problem nominal optimum, [$\underline{X}^{o*}$].

The toleranced design optimization problem could be mathematically represented as follows

$$\text{minimize } U(\underline{X})$$

$$\text{subject to} \quad \underline{\Phi}(\underline{X}) \geq \underline{0} \qquad (2.31)$$

where, $\underline{X} = \underline{X}^o + \underline{T}^+ \underline{\alpha}^+ + \underline{T}^- \underline{\alpha}^-$

$$\underline{T}^{+} = \begin{bmatrix} t_1^{+} & & & 0 \\ & t_2^{+} & & \\ & & \ddots & \\ 0 & & & t_n^{+} \end{bmatrix} , \quad \underline{T}^{-} = \begin{bmatrix} t_1^{-} & & & 0 \\ & t_2^{-} & & \\ & & \ddots & \\ 0 & & & t_n^{-} \end{bmatrix} \tag{2.32}$$

$$\underline{\alpha}^{+} = [\alpha_1^{+}, \alpha_2^{+}, \ldots, \alpha_n^{+}]^T \quad , \quad \underline{\alpha}^{-} = [\alpha_1^{-}, \alpha_2^{-}, \ldots \alpha_n^{-}]^T$$

In this expression t_i^{+} and t_i^{-} are scalar variables which represent the absolute or relative nominal positive and negative tolerances associated with t_i^{o}, respectively. The quantities α_i^{+} and α_i^{-} are the normalized random positive and negative tolerances and their probability density functions are assumed to be truncated and vary between zero and one and between negative one and zero, respectively. Positive tolerance distribution might vary considerably from that of the negative tolerance for the same nominal variable, for example, internal cylindrical turning, piercing and deep drawing operations.

The level of a design acceptability, A, is defined by

$$A = \frac{\text{number of designs which met specifications(constraints)}}{\text{total number of designs}} \tag{2.33}$$

In this analysis we are only concerned with one hundred percent acceptability design, in which the tolerance region, R_t, must be completely within the constraint region, R_c.

$$R_t \subset R_c \tag{2.34}$$

where,

$$R_t=\{\underline{X}|\underline{X}=\underline{X}^o+\underline{T}^+\underline{\alpha}^++\underline{T}^-\underline{\alpha}^-,\ 0\leq\alpha_i^+\leq 1, -1\leq\alpha_i^-\leq 0,\ i=1,2..n\}$$

and

$$R_c=\{\underline{X}|\underline{\Phi}(\underline{X})\geq \underline{0}\}$$

We assume the R_c is convex, where all the points created by a linear interpolation between any two points in the region lie inside the region as shown in Figure (2.14). It is only sufficient, therefore, for the tolerance region, R_t, to satisfy

$$R_{tc}\subset R_c \tag{2.35}$$

where,

$$R_{tc}=\{\underline{X}|\underline{X}=\underline{X}^o+\underline{T}^+\underline{\alpha}^++\underline{T}^-\underline{\alpha}^-, \alpha_i^+\varepsilon\{0,1\}\ \&\ \alpha_i^-\varepsilon\{-1,0\}\ , i=1,2..n\}$$

The 2^n distinct worst designs of the tolerance region, R_{tc}, are taken as a unique representative of the full-acceptability design domain because it is impractical and very expensive to consider explicitly the infinite number of designs contained in the tolerance region, R_t, even if a Monte Carlo technique is employed with a reasonable confidence.

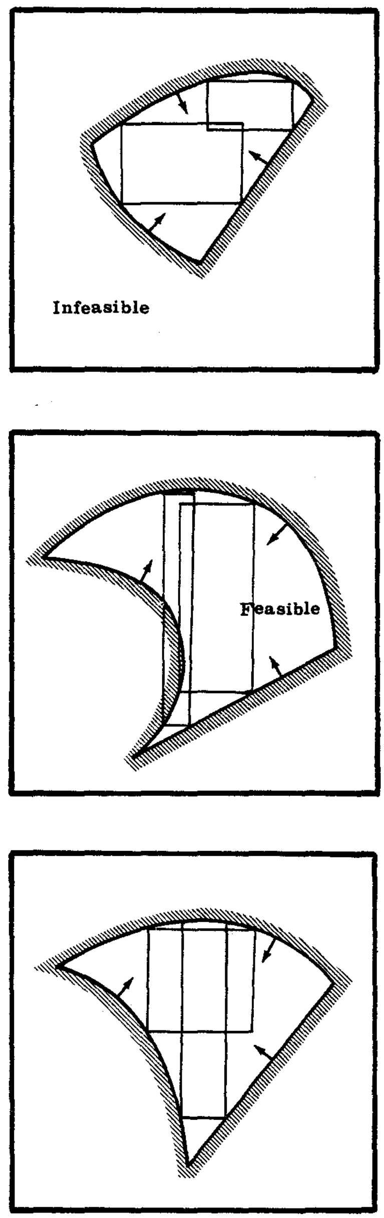
It is a sufficient but not a necessary condition for R_c to be a strictly convex region so that Equation (2.35) can be validly applied. It is evident that the feasible region of the design problem $[P]_1$ is not strictly convex, even though the R_{tc} condition is still valid. It is only necessary and sufficient, therefore, for R_c to be a parallel convex region for R_{tc} to be a subset R_c domain. A parallel convex region is a closed region where all the points created by a linear interpolation between any two points in the region, which form a line parallel to the domain variables axis, lie inside the region as shown in Figure 2.14.

2.4 Conclusions

Unlike a conventional optimization problem, $[\underline{X}^o]$, where a single point is of interest, the optimization scheme introduced in this chapter creates a region of interest, $[\underline{X}]^*$. If the constraints region satisfies the parallel convexity assumption, then all points inside and on the boundaries of the tolerable optimization region, R_{tc}, will represent an optimum feasible design.

The optimum tolerance range, R_{tc}, increases significantly when the design variables associated tolerances are allowed to be unsymmetrically allocated, i.e., $t_i^+ \neq t_i^-$, as can be seen from Figure 6.2 any by comparing $[P]_5$ and $[P]_6$.

It is also evident from comparing $[P]_3$ and $[P]_5$ or $[P]_4$ and $[P]_6$ that by allowing the nominal base point, r_1^o and r_2^o, to move, a set of larger tolerances is not always obtained. This is because the optimum tolerance area, A_t, in the toleranced design optimization problem, $[\underline{X}]^*$, mainly depends on the optimization criterion. On the other hand,



a)

b)

c)

Figure 2.14 a) Convex region b) Non-convex region
c) Parallel convex region.

the optimization criterion has no effect on A_t if $\underline{X}^o$ is fixed. For example, if U_5 was used as the optimization criterion for $[P]_3$ the same optimum tolerance would have been found.

$[P]_2$ was mainly introduced to demonstrate a possible practical situation where there is no meaningful objective function, and the designer is concerned with centering the design so as to be as remote as possible from all feasibility bounds. Thus none of the problem constraints were active, $(\phi_i(\underline{X}^o)^* \neq 0,\ \ i=1,2,..m)$. Unlike $[P]_1$, where two constraints approach zero, $[P]_2$, by having a centered fixed optimum design, $[\underline{X}^o]_2^*$, helps in demonstrating the generation of the associated tolerance domain, R_{tc}, as explained in $[P]_3$ and $[P]_4$. A centered optimum solution could in a like manner be achieved by minimizing the errors created by the problem constraints where the optimization criterion might be expressed as

$$\text{minimize } U = \sum_{i=1}^{m} |\phi_i(\underline{X}^o)|^P; \quad p \geq 1 \tag{2.36}$$

where the problem constraints, $\underline{\Phi}(\underline{X}^o)$, must be normalized or weighted. In general, however, the optimal centering process is not essential to reach an optimal tolerable design similar to those explained in $[P]_5$, $[P]_6$ and $[P]_7$, once the optimization criterion is established.

The importance of having an accurate cost model is demonstrated in $[P]_7$ by altering the cost level of one of the objective components. Depending upon the cost model, however, a modest number of statistically designed experiments could lead to a good estimate of the chosen model coefficients and consequently a reasonably accurate toleranced design.

In all of the above analysis the input specifications, $\underline{S}$, e.g., the cylinders' material strength and the internal applied pressure, were assumed deterministic quantities. Their randomness could be expressed by

$$\underline{S}=\underline{S}^{o}+\underline{\Delta}\ \underline{\gamma} \tag{2.37}$$

where,

$$\Delta = \begin{bmatrix} \delta_1 & & \\ & \delta_2 & 0 \\ & 0 & \ddots & \delta_k \end{bmatrix}$$

$$-1\leq\gamma_i\leq1 \quad ; \quad i=1,2,..k$$

k is the number of the problem statistically varied input specifications.

δ_i is a scalar variable which resembles the variation of the specification s_i around its mean s_i^o.

γ_i is a random variable which follows the same symmetrical distribution as that of s_i.

The problem of full acceptance toleranced optimum design could be further extended to include the uncertainties in $\underline{S}$

and mathematically stated as

$$\text{minimize } U(\underline{X},\underline{S}) \qquad (2.38)$$

$$\underline{\Phi}(\underline{X},\underline{S}) \underline{>} 0$$

The non-convexity of the constraint region could be overcome by checking the linear boundaries of R_{tc} using several equidistanced stations, depending upon the level of confidence required in the solution. This should be carried out after a primary solution is reached. If any of the checking stations for a specific linear boundary fails to satisfy any of the $\underline{\Phi}(\underline{X})$ constraints, R_{tc} will be reduced by displacing the check line a small distance ε towards the nominal optimum $\underline{X}_o^*$. The checking procedure is only repeated for those lines which were found to intersect with the non-convex region boundaries and it is successively repeated until feasability is guaranteed for the R_{tc} region. This point also will be discussed in further depth in Chapter 5.

The true optimum solution for either $[P]_1$ or $[P]_2$ will not differ if their design variables are taken to be $[r_1, r_2, p_f]$ or $[r_1, r_2, u_1]$. However, the choice of the design variables in the toleranced design problem is more critical since it affects the final solution and the whole problem objective and implementation.

CHAPTER 3

TOLERANCE ASSIGNMENT WITH LESS THAN FULL ACCEPTANCE

3.1 Introduction

The determination of the upper and lower limits of design variables of an engineering system, as discussed in the previous chapter, were based on the full acceptance of the design outcomes. Any design that happens to lie between these limits would fulfill all the system performance and geometrical constraints.

The limits on the random design variables, $\underline{X}$, however, could be increased if we allowed a portion of the selected design outcomes to violate the system constraints. The percentage of this portion of the violated designs to the total accepted designs defines the design scrap percentage, S_d.

Manufacturing scrap percentage, S_m, on the other hand, may differ from the design scrap depending upon the nature of the manufacturing process and the system function and design. The manufacturing scrap percentage is the portion of the rejected manufactured components, which do not meet any of the system design variables' upper and lower limits. These limits, however, may include some designs which violate the system constraints. That is to say, the design scrap percentage may be greater than zero. The distinction between the design and the manufacturing scrap will be further elaborated in the ensuing sections.

The optimum design scrap percentage, S_d^*, of a system may be estimated by utilizing any suitable unconstrained optimization technique. And consequently the system optimum tolerance limits would be determined as well. The design and the manufacturing scrap percentages should be estimated statistically for every optimization iteration. The regionalization technique, developed by Gopal[42] is modified and adapted in conjunction with an optimization strategy to save some computational time in estimating the system design scrap.

The design example of the two fitted cylinders, previously illustrated, will be used to aid in elaborating the various design objectives as well as various probabilistic distributions is studied, using a different hypothetical manufacturing-to-design scrap relationship.

3.2 Manufacturing and Design Scrap

The probabilistic distribution, $f(x_i)$, of a certain dimension, x_i, in a manufactured component at the time of assembly, is not always equal to that given by the manufacturing process because after production, the parts usually have to be inspected, and a part will be rejected if any of its specifications, say dimension x_i, happens to be outside the design tolerance region

$$x_i^o - t_{si}^- \leq x_i \leq x_i^o + t_{si}^+ \qquad (3.1)$$

where, t_{si}^+ and t_{si}^- are the positive and negative tolerance values as specified by the design engineer for the nominal dimension x_i^o. However, depending upon the manufacturing process the outcome machined dimension x_i may vary between

$$\overline{x}_i - t_{min} \leq x_i \leq \overline{x}_i + t_{max} \qquad (3.2)$$

where, $\overline{x}_i$ is the process statistical mean which does not have to coincide with x_i^o, and t_{min} and t_{max} are the process lower and upper deviations. Using simple linear transformation, x-plane variables could be mapped to the tolerance t-plane domain or to the unit z-plane domain, as shown in Figure 3.1. This mapping will make the statistical recognition of the problem much easier.

The distribution of the component dimension, z_i, going to assembly, will be the distribution given by the process, $f(z_i)$, truncated at the optimum positive and negative design tolerance values, z_{si}^+ and z_{si}^-, as shown in Figure 3.2. Manufacturing scrap percentage, S_m, for a single variable i, is the percentage of the shaded areas to the total area under the $f(z_i)$ distribution, and it could be mathematically defined as follows

$$S_m = \left(\int_0^{z_{si}^-} f(z_i)\, dz_i + \int_{z_{si}^+}^{1} f(z_i)\, dz_i \right) \times 100\% \qquad (3.3)$$

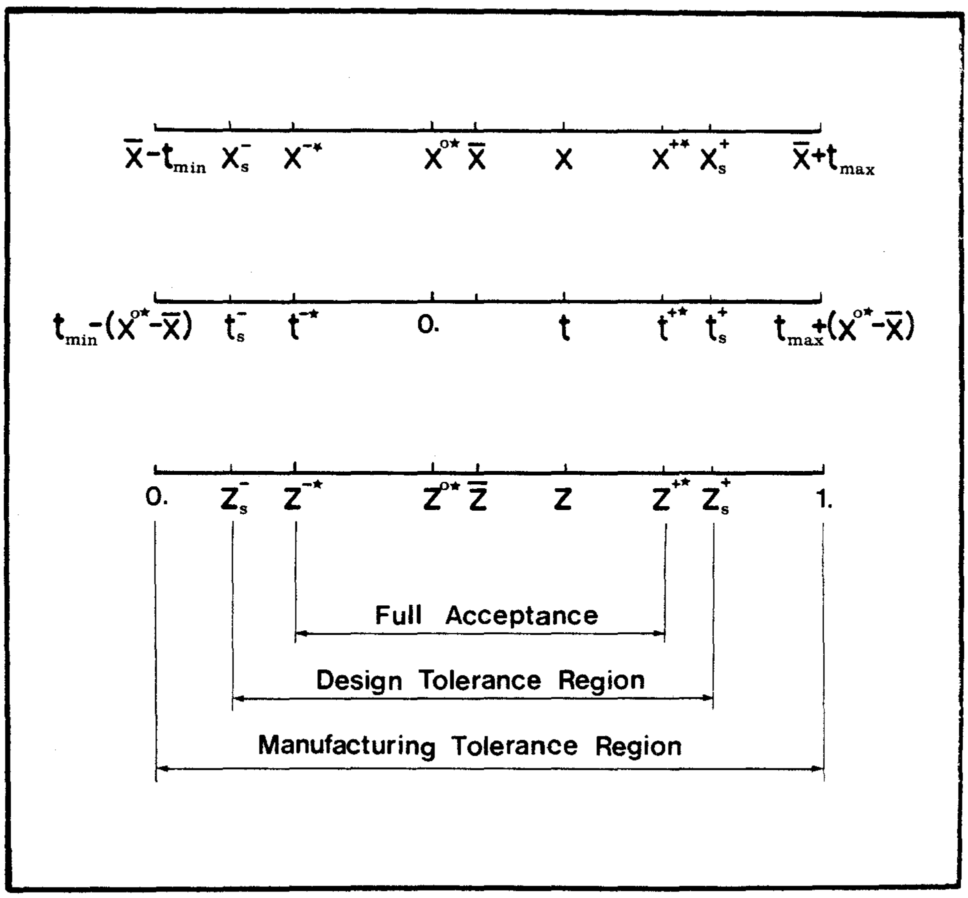


Figure 3.1 Mapping of the random variable x to the random tolerance t and the unit domain z.

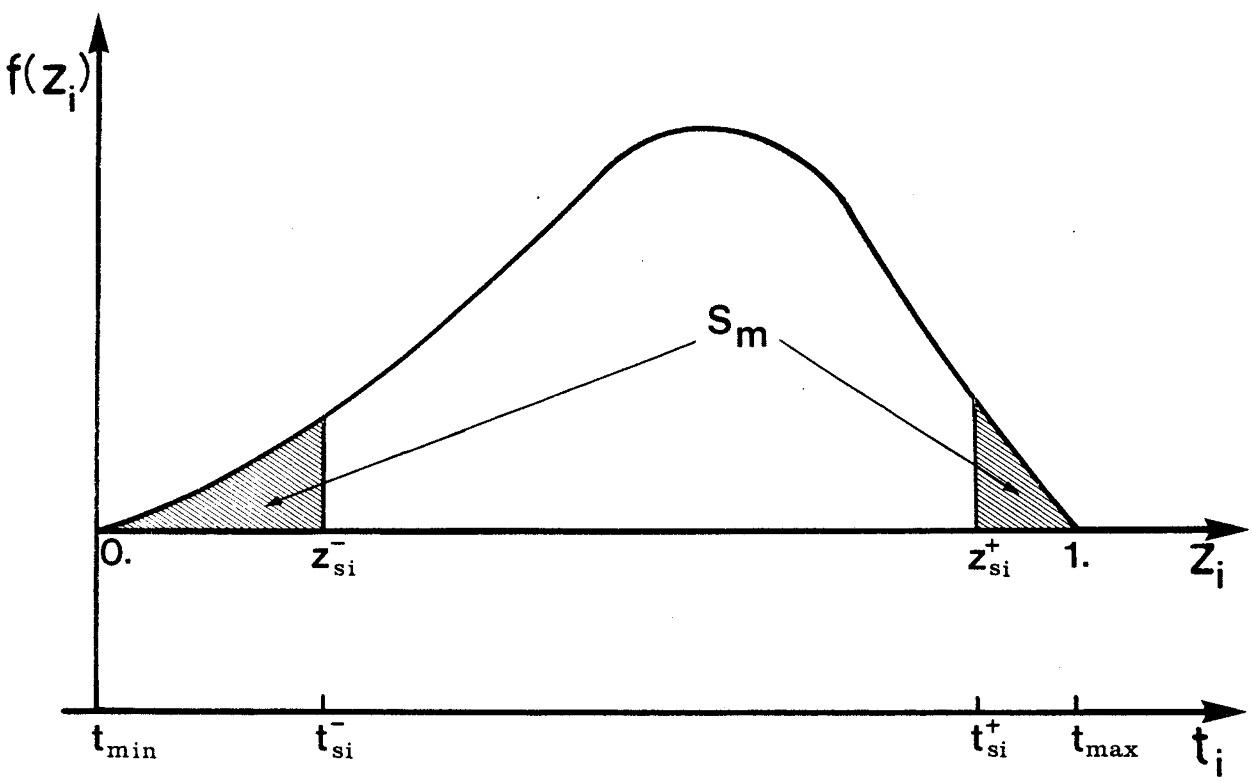


Figure 3.2 Single variable manufacturing scrap, S_m.

The manufacturing scrap might equal zero, however, if the process capability, $t_{min} + t_{max}$, is less than or equal to the design tolerance range, $t_s^- + t_s^+$, as shown in Figure 3.3.

Design scrap, S_d, on the other hand, is the percentage of the assembled design components which fails to fulfill the system constraints. The assembled fitted two cylinders, for example, might pass the inspections of their manufactured dimensions before and after the assembly, but still not satisfy one or more of the design constraints. The design scrap percentage, for a single variable i, is the percentage of the hatched areas shown in Figure 3.4 to the total area under the truncated $f(z_i)$ distribution, and it could be mathematically expressed as

$$S_d = \left(\int_{z_{si}^-}^{z_i^-} f(z_i)\, dz_i + \int_{z_i^+}^{z_{si}^+} f(z_i)\, dz_i \right) \Bigg/ \int_{z_{si}^-}^{z_{si}^+} f(z_i)\, dz_i \tag{3.4}$$

The design scrap could be constrained to zero if the design tolerance range, $t_s^- + t_s^+$, will be within the full acceptance tolerance range, $t^{-*} + t^{+*}$.

3.3 Probability Distributions

To determine a system scrap percentage the probabilistic distributions of the system random design variables, x_i's, have to be known. The distribution $f(x_i)$ depends on

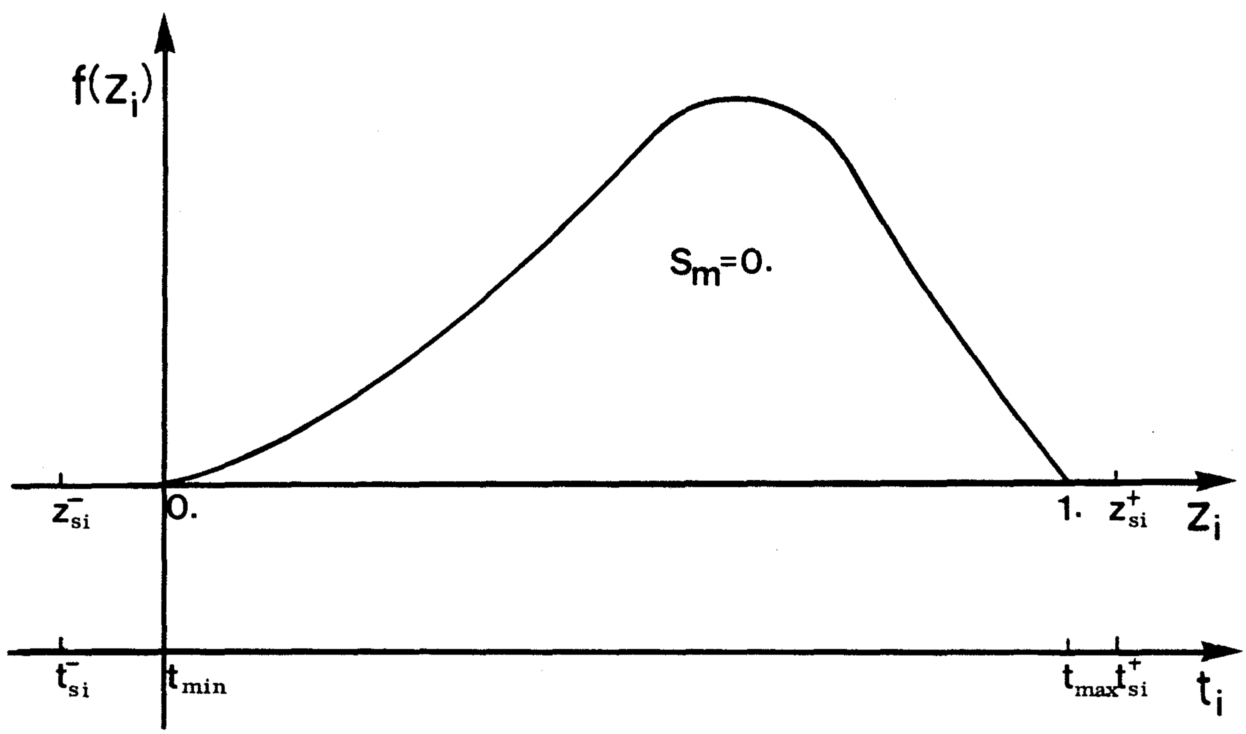


Figure 3.3 Zero manufacturing scrap percentage.

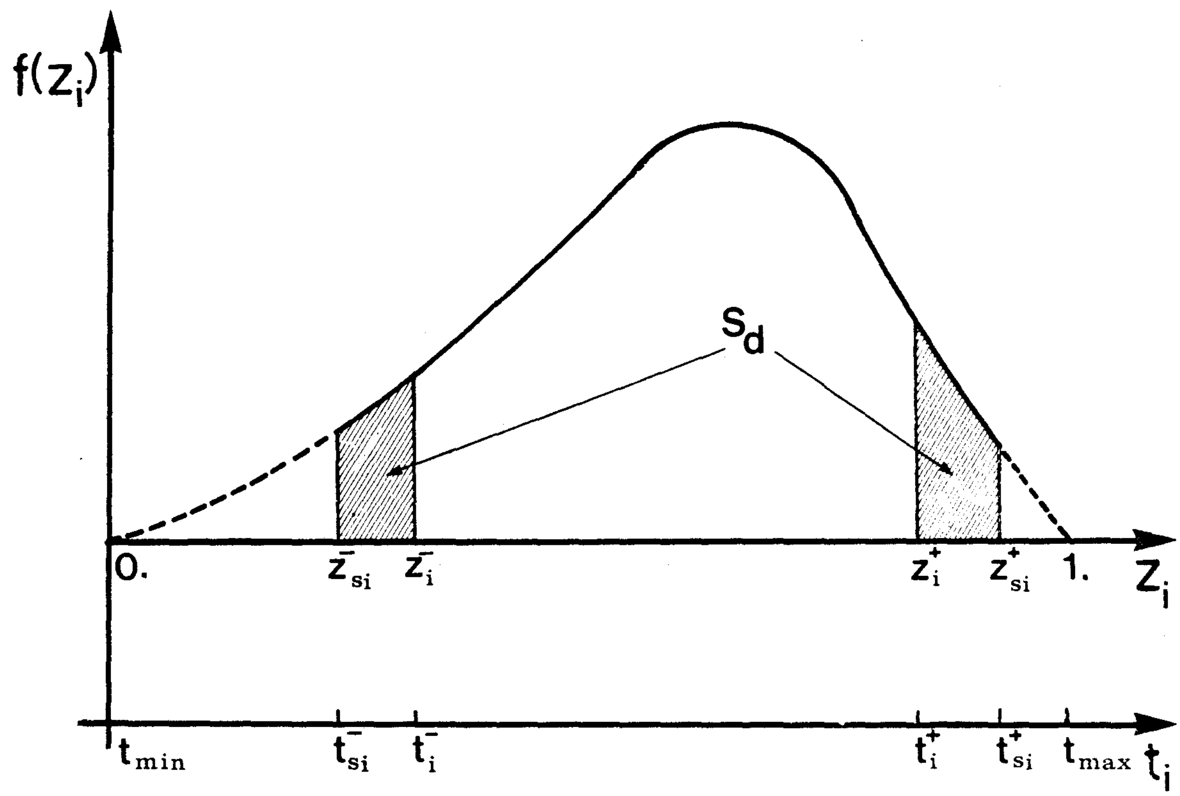


Figure 3.4 Single variable design scrap, S_d.

the process used to manufacture the design component which contains the random variable, x_i. Although the probabilistic distributions of manufacturing process outcomes may be identified by either conducting a set of experiments or utilizing some historical production control data, these distributions are not always known. Most processes, however, give distributions varying between a normal to rectangular distributions. A rectangular distribution is the worst estimate of the probabilistic distribution given by a process, and this makes scrap calculations possible even without process distributions knowledge.

The central limit theory implies that the distribution of the sum dimension is asymptotically normal, independent of the distributions of the individual dimensions, as long as the number of the assembled dimensions is large enough or if the individual dimensions distributions are normal. Neither condition, however, applied to the engineering systems. First, variations of system random variables have finite range and they never vary between positive and negative infinity. That is to say, all the assumed normal distributions should be truncated. Secondly, the number of the assembled random dimensions is not likely to be large enough to allow the validity of the central limit theory. Moreover, the confidence in the estimation of a system scrap is particularly influenced by the shapes of the tails of the random variables' distributions.

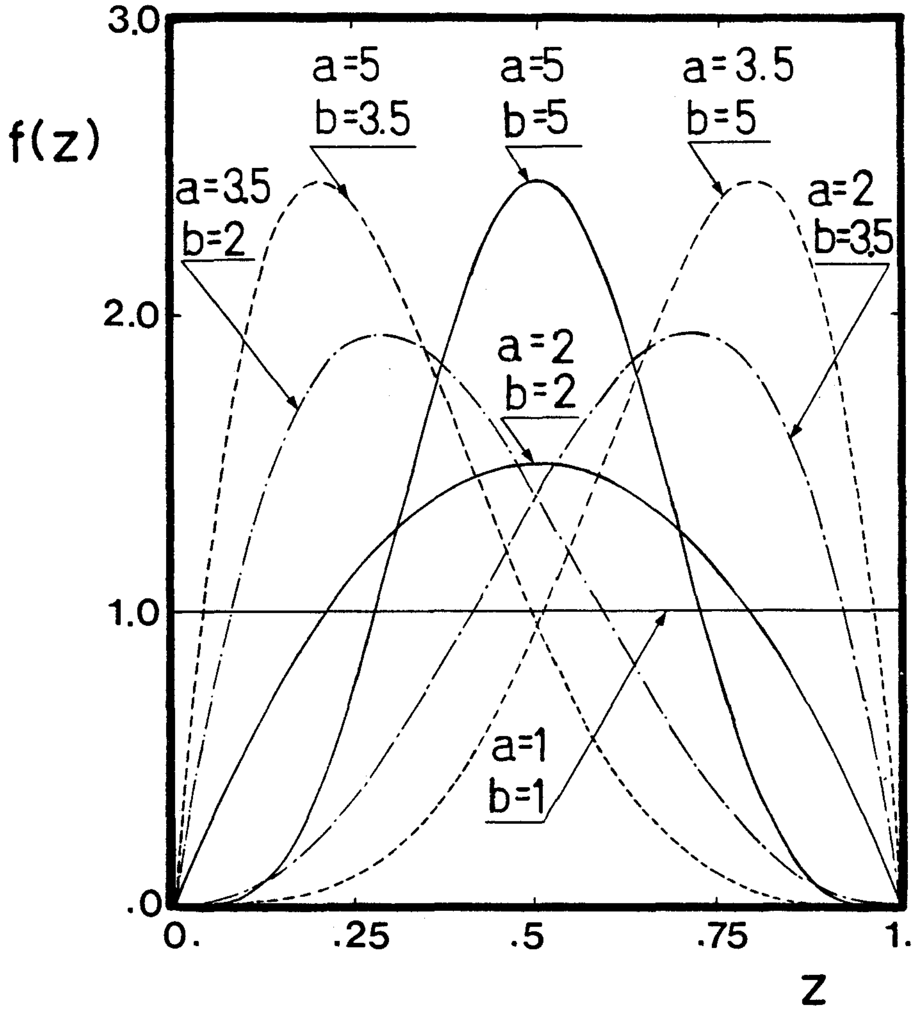


Figure 3.5 Unit beta distribution.

The beta distribution, therefore is recommended and will be taken as the manufacturing processes statistical model. This distribution, besides being easy to use, has a finite range from zero to one. It also covers asymmetrical as well as symmetrical cases, and its shape could vary from rectangular to normal, as shown in Figure 3.5. The probability density function of the unit beta distribution in the interval [0,1] is

$$f(z) = z^{a-1}\,(1-z)^{b-1} \Big/ \int_0^1 t^{a-1}\,(1-t)^{b-1}\,dt \qquad (3.5)$$

where, a and b are the distribution parameters and they could be expressed as a function of the distribution mean $\overline{z}$ and the standard deviation σ as follows

$$a = \overline{z}[\{\overline{z}(1-\overline{z})/\sigma^2\}-1]$$

and (3.6)

$$b = a(1-\overline{z})/\overline{z}$$

The statistical parameters a and b could, therefore, be estimated for a particular manufacturing process if its mean and deviation are known, or by mapping a set of an observed outcomes x into z and best fitting the z observations into f(z).

3.4 Space Regionalization

To estimate a system design scrap percentage the system random design variables have to be checked for feasibility in every point on the function surface. By utilizing a regionalization technique, however, computational savings might be accomplished, since the technique approximates the multi-dimensional joint probability function, $f(\underline{X})$, by a finite number of discrete points. Each of these points is centered inside a distinct cell and represents all the points occurring within the cell boundaries, as illustrated in Figure 3.6. While using the regionalization technique the system random variables, e.g., r_1 and r_2, are discretized into R_i regions over the variables' tolerance interval $-t_i^-$ to t_i^+. Therefore

$$N_{cell} = \prod_{i=1}^{N_t} R_i \tag{3.7}$$

where, N_{cell} is the total number of cells and, N_t is the total number of the random design variables of the system. The joint probability p_I associated with different cells I can be evaluated as follows

$$p_I = \prod_{j=1}^{N_t} p_{ij} \tag{3.8}$$

where $I \in [1 , N_{cell}]$

$i \in [1 , R_i]$

and $p_{i1}, p_{i2}, \ldots, p_{iN_t}$ are the independent probabilities of occurrence for each of the random design variables between the boundaries of the I cell, as described in Figure 3.7 for a system of two variables. The design scrap percentage, therefore, could be estimated by

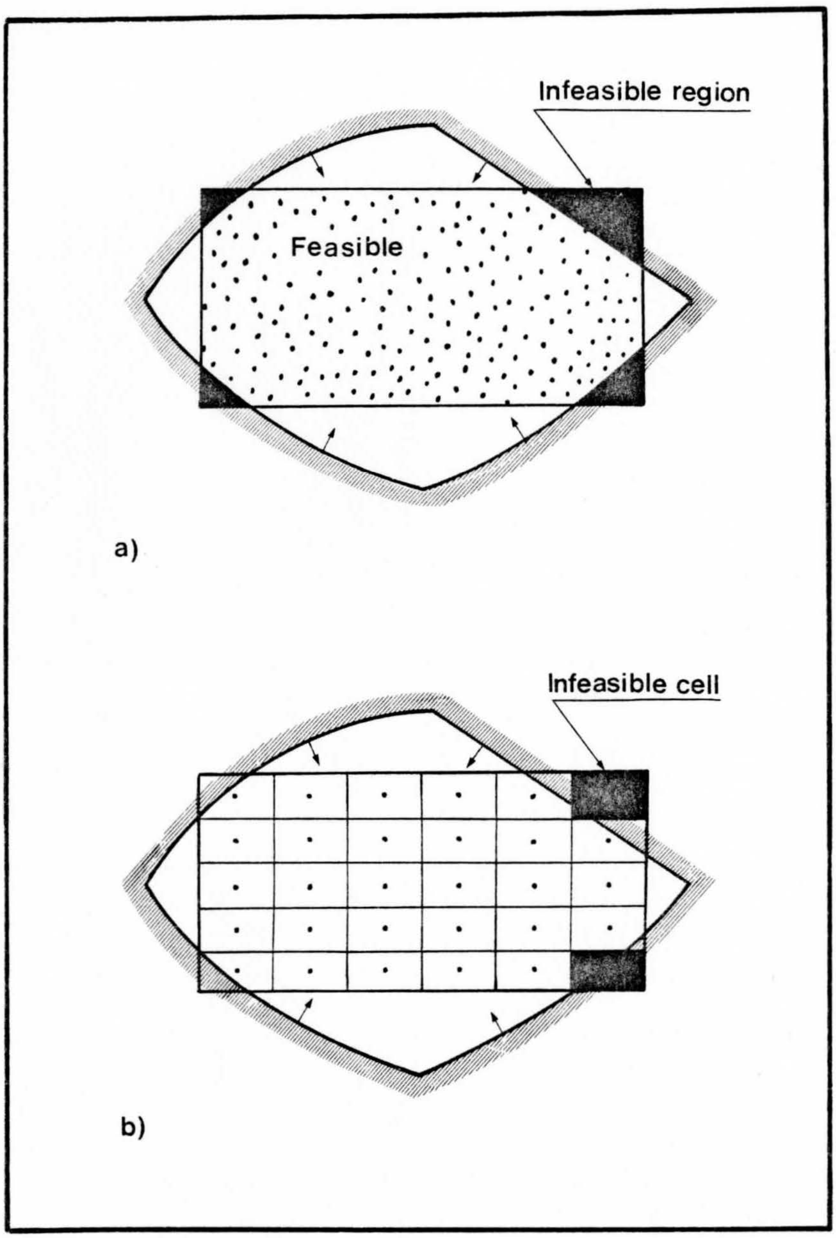


Figure 3.6 Random sample generation versus regionalization.

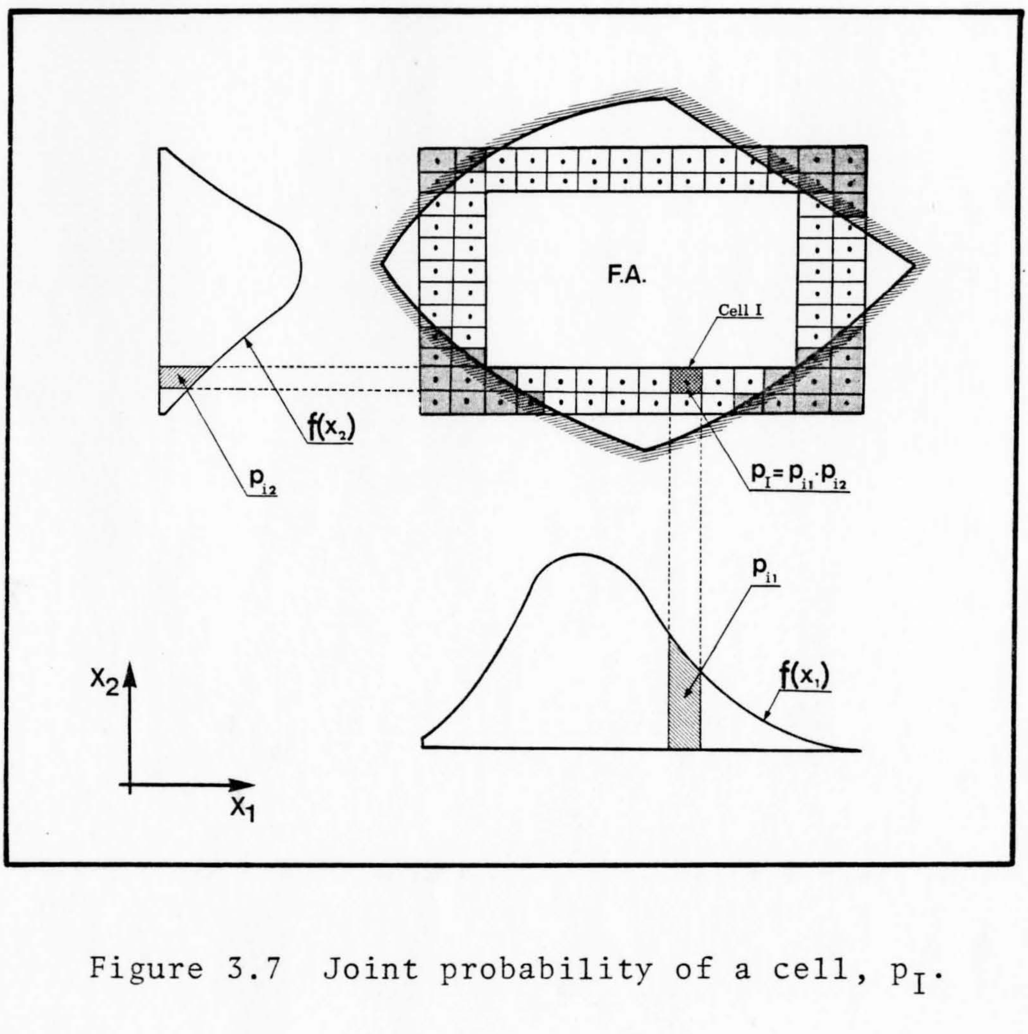


Figure 3.7 Joint probability of a cell, p_I.

$$S_d = \frac{\sum_{I=1}^{N_{cell}} p_I \; \alpha_I}{\sum_{I=1}^{N_{cell}} p_I} \times 100\% \qquad (3.9)$$

where $\alpha_I = 0$ if the representative central point of the I cell proved to satisfy all the system constraints, $\underline{\Phi}$.

$= 1$ otherwise; infeasible cell - shown shaded in Figure 3.6.

Considerable computational savings, however, can be achieved in estimating S_d if the system full acceptance, F.A., region is known. Since all the cells in the FA region are feasible, their α_I's equal zero and there is no need to check them. Figure 3.8 illustrates the reduction in the number of cells. The system design scrap percentage, therefore, could be computed as follows

$$S_d = \sum_{I=1}^{N_{cell}} p_I \; \alpha_I \bigg/ \left(\sum_{I=1}^{N_{cell}} p_I + p_{FA} \right) \times 100\% \qquad (3.10)$$

where p_{FA} is the probability of the full acceptance region.

To determine the total number of cells, N_{cell}, needed to estimate S_d with a reasonable accuracy, the number of cells R_i for each of the random variables must be adjusted first depending upon the range of each variable and its probability distribution. The number of cells per variable

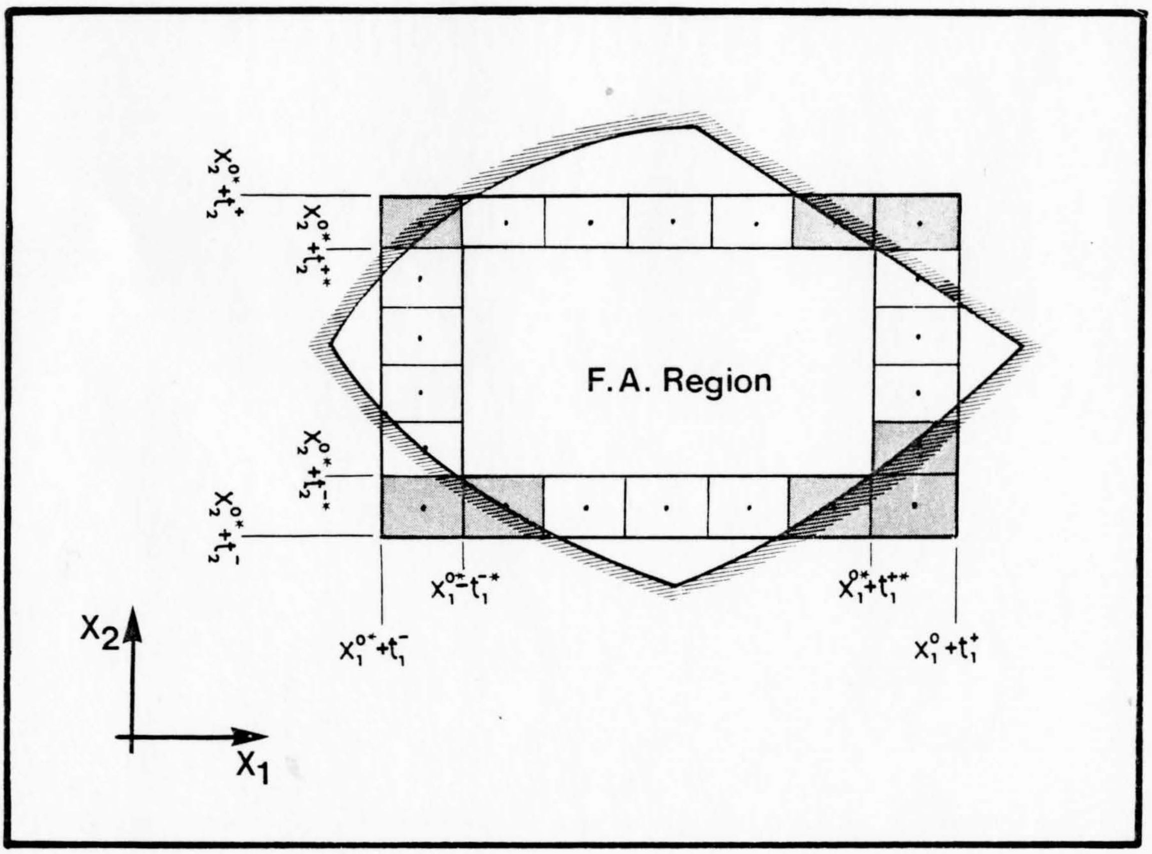


Figure 3.8 Estimating the system scrap utilizing the known full acceptance solution.

may be divided into three sectors as shown in Figure 3.9 and expressed mathematically as follows

$$R_i = R_{ti}^- + R_{mi} + R_{ti}^+ \tag{3.11}$$

where R_{ti}^+ is the number of cells between t_i^+ and t_i^{+*};

$$= \text{Int}\ [\lambda_{i1}\ (t_i^+ - t_i^{+*})/(t_i^{+*} + t_i^{-*})] + 1$$

R_{mi} is the number of cells between t_i^{-*} and t_i^{+*};

$$= \text{Int}\ [\lambda_{i2}] + 1$$

R_{ti}^- is the number of cells between t_i^{-*} and t_i^-;

$$= \text{Int}\ [\lambda_{i3}\ (t_i^- - t_i^{-*})/(t_i^{+*} + t_i^{-*})] + 1$$

Int [f] is the truncated integer value of the real function f.

λ_{i1}, λ_{i2} and λ_{i3} are scaling factors.

The values of λ_{i1}, λ_{i2} or λ_{i3} for different random variables should be proportional to the corresponding probability of occurrence for each sector per variable. For example,

$$\frac{\lambda_{11}}{\lambda_{21}} = \frac{\int_{z_1^{+*}}^{z_1^+} f(z_1)\ dz_1}{\int_{z_2^{+*}}^{z_2^+} f(z_2)\ dz_2} \tag{3.12a}$$

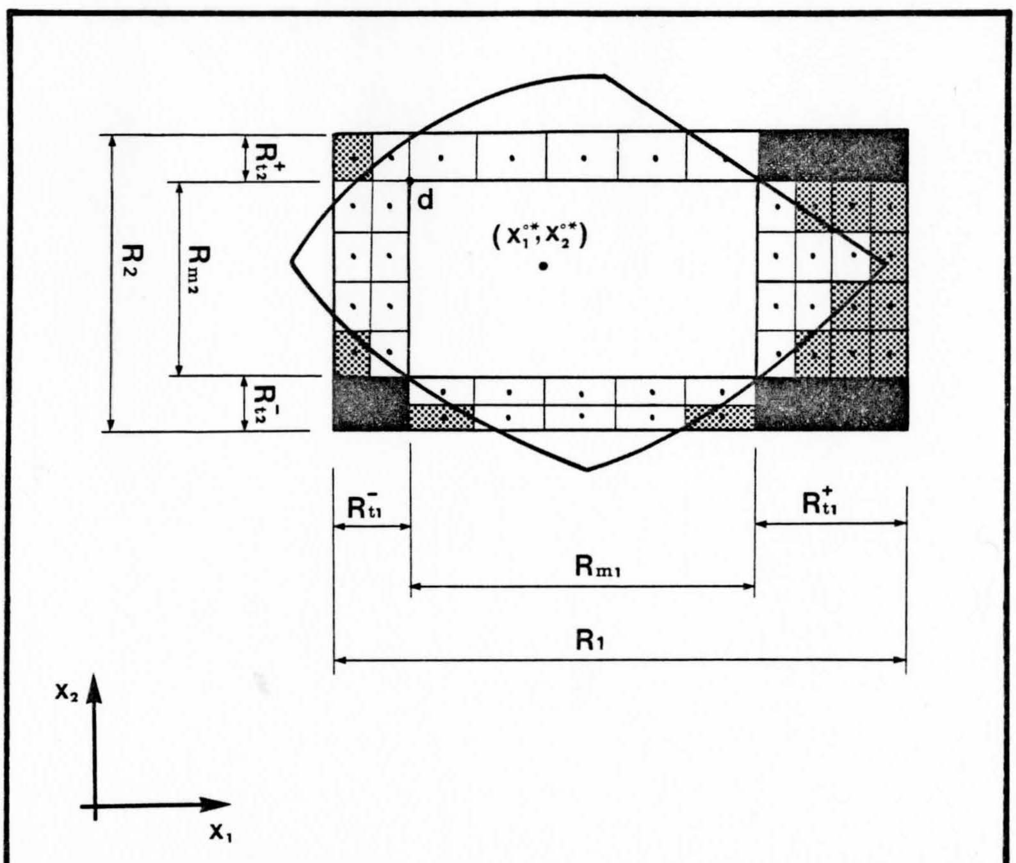


Figure 3.9 Segmentation of the system tolerance region.

and also for the same design variable and different sectors

$$\frac{\lambda_{11}}{\lambda_{13}} = \frac{\displaystyle\int_{z_1^{+*}}^{z_1^{+}} f(z_1)\, dz_1}{\displaystyle\int_{z_1^{-}}^{z_1^{-*}} f(z_1)\, dz_1} \qquad (3.12b)$$

The reference value of λ_{11} say and its relation to λ_{12} depends upon the confidence required in estimating S_d as well as the complexity of the feasible region bound.

Additional computational savings could be possible if the cells adjacent to the active corners of the full acceptance region are not checked against the system feasibility, since they will always be infeasible. A corner of the FA region is considered active if one or more of the system design constraints equals zero, as shown in Figure 3.9, where corners a, b and c are active and the cancelled cells identified by heavy shades. This leads to a saving of about 23% of the checking computational time for this particular problem setting.

Moreover, if the remaining cells are ranked in a descending order depending upon their probabilities, p_I, as shown in Figure 3.10, then, according to the desired accuracy in estimating S_d, those cells having negligible probability can be ignored.

The system design scrap percentage could finally be

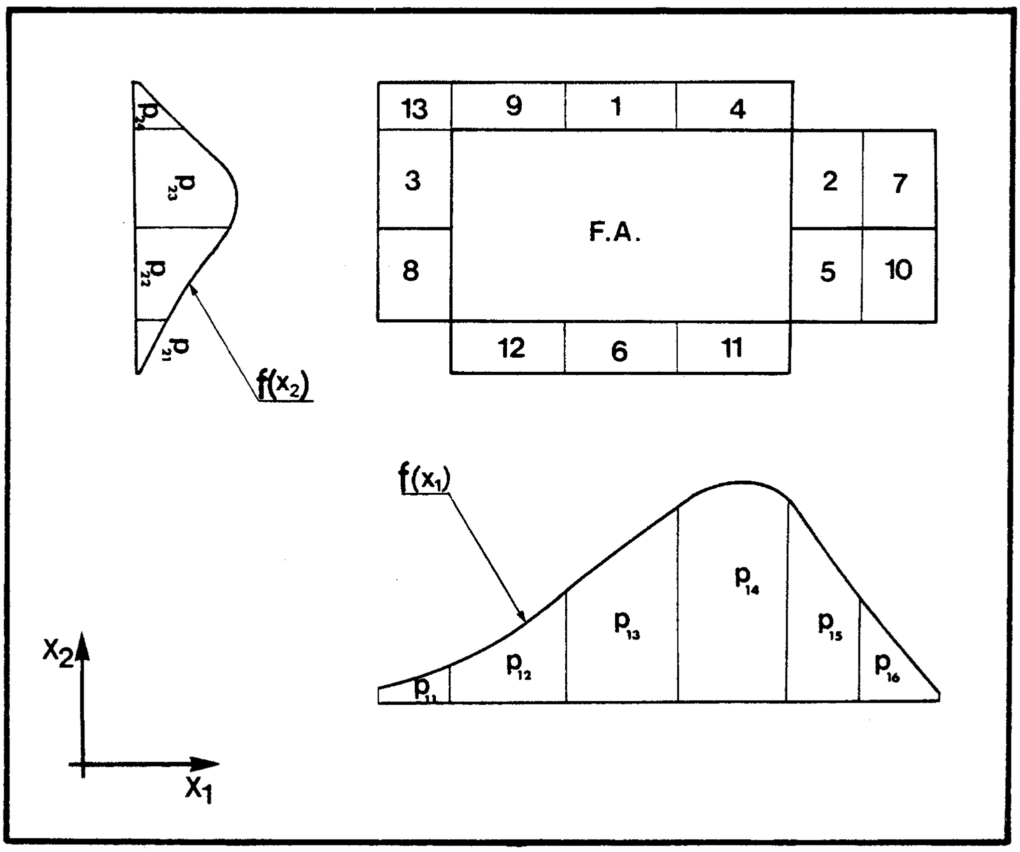


Figure 3.10 The ranked cells.

expressed by

$$S_d = \left[\sum_{I=1}^{N^*_{cell}} p_I \, \alpha_I + p_{ac}\right] \Bigg/ \left[\sum_{I=1}^{N^*_{cell}} p_I + p_{FA} + p_{ac}\right] \times 100\% \tag{3.13}$$

where, p_{ac} is the sum of the probabilities of the regions adjacent to the active corners.

N^*_{cell} is the number of the ranked cells to be checked.

Consequently, the system manufacturing scrap percentage is expressed as follows

$$S_m = 1 - \left[\sum_{I=1}^{N^*_{cell}} p_I + p_{FA} + p_{ac}\right] \times 100\%$$

$$= 1 - \left[\prod_{i=1}^{N_t} \int_{z^-_{si}}^{z^+_{si}} f(z_i)\, dz_i\right] \times 100\% \tag{3.14}$$

The complexity of the system under consideration depends upon the number of toleranced variables N_t, the random variables probability density function $f(x_i)$, the total number of cells N_{cell}, and also upon the adopted optimization strategy, the objective function surface and the feasible region boundaries.

3.5 Examples

The design example of the two fitted cylinders introduced in the previous chapter is used to elaborate the various concepts of less than full acceptance designs. The regionalization technique is utilized to estimate both S_d and S_m by discretizing the random variables domain into N_{cell} cells. The techniques suggested to gain more confidence in the computed system scrap, with minimum computational effort, are also used. Two unique cases derived from Equations (3.13) and (3.14) and matched with practical production situations, will also be discussed. They are:

i) When the production volume is small, or when it is required irrespective of economical considerations that all the manufactured components have to meet the system design constraints. The design scrap should equal zero; there should be full acceptance design; and manufacturing scrap is to have a minimum value.

$$S_d = 0\%; \quad S_m \geq 0\% \tag{3.15a}$$

ii) When the production volume is large, but the manufacturing processes capabilities are more tight than that imposed by the design.

$$S_d \geq 0\%; \quad S_m = 0\% \tag{3.15b}$$

i.e. $$\sum_{i=1}^{N_{cell}} p_I = 1$$

3.5.1 Predetermined Design Scrap, Zero Manufacture Scrap, $[P]_8$.

The design optimum nominal vector $[\underline{X}^o]_2^*$ is assumed known from $[P]_2$ as well as the fully accepted non-symmetrical tolerance boundaries $[\underline{T}^+]_4^*$ and $[\underline{T}^-]_4^*$ are also known from $[P]_4$. Since the manufacture scrap equals zero, the system scrap equals the design scrap. The problem objective is to maximize the design tolerances allowing a maximum of a specified scrap percentage, S_{sp}%.

3.5.1.a Uniform distribution

Both r_1 and r_2 are assumed to follow uniform distribution with a range of $r_1^{o*} - t_{1s}^- \leq r_1 \leq r_1^{o*} + t_{12}^+$ for r_1 and similarly for r_2.

$$[\underline{X}]_8 = [t_{1s}^+, t_{1s}^-, t_{2s}^+, t_{2s}^-]^T \tag{3.16}$$

$$U_8 = \frac{1}{t_{1s}^+ + t_{1s}^-} + \frac{1}{t_{2s}^+ + t_{2s}^-} = \text{minimum}$$

$$[\underline{\Phi}]_8 = \phi_{1,8} = S_{sp} - S_d$$

where $[r_1^{o*}, r_2^{o*}] = [0.126, 0.185]$

The following table and Figure 3.11 summarize the optimum results obtained for different S_{sp}%. And the percentage tolerance area gained over that of the full acceptance solution is also tabulated.

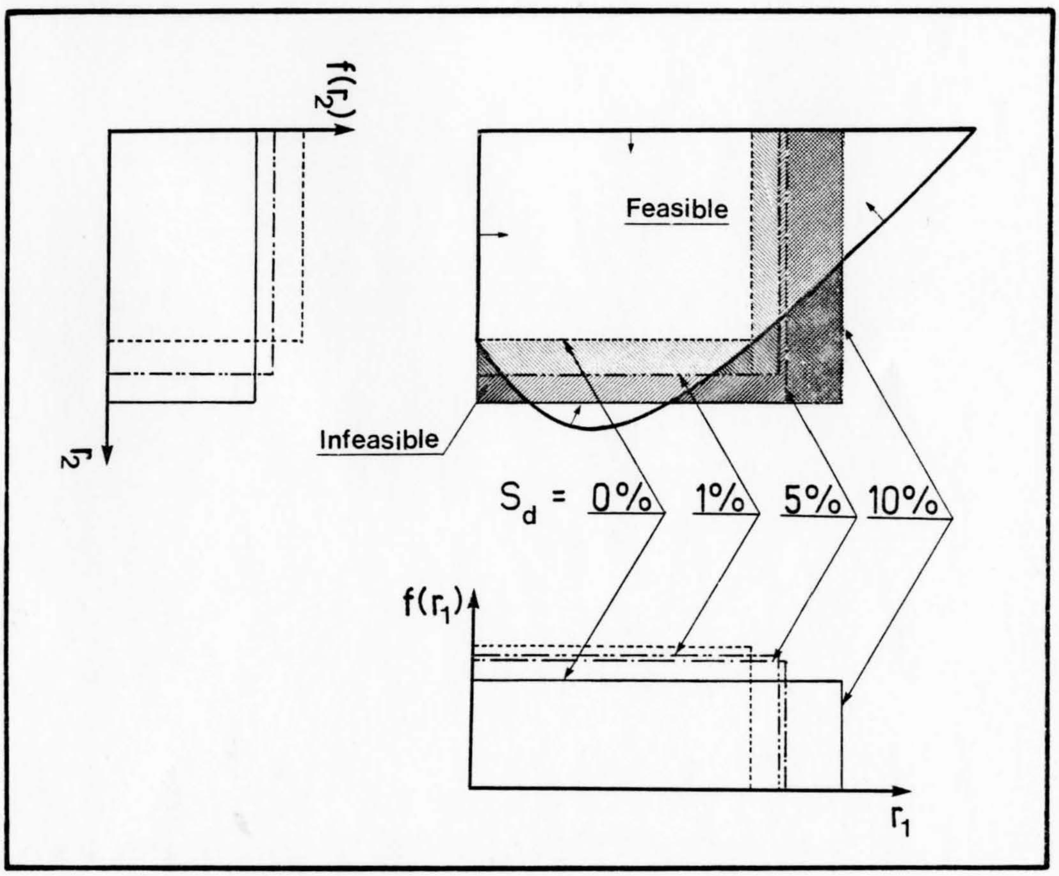


Figure 3.11 Predetermined design scrap - uniform distribution.

	Uniform Distribution			Beta Dist.
S_{sp}	10%	5%	1%	1%
r_{1s}^{+}	0.168	0.159	0.158	0.195
r_{1s}^{-}	0.220	0.110	0.110	0.110
r_{2s}^{+}	0.200	0.200	0.200	0.202
r_{2s}^{-}	0.156	0.156	0.161	0.144
% Area Gained	71	43	24	230

3.5.1b Beta Distribution

The same problem was tackled assuming beta distribution for both the system random variables r_1 and r_2. The distribution means are taken to be the fixed optimum nominals r_1^{o*} and r_2^{o*}, respectively. While their standard deviations are considered to be one eighth of each distribution range. The optimum result found is tabulated above and shown in Figure 3.12 for a specified design scrap of one percent.

The effect of varying the scrap levels as well as the random variable distributions on the allocation of the optimum tolerance domain was demonstrated above. It is evident that the tolerance domain is apt to cover the maximum possible area in the feasible domain while maintaining the same scrap level. It is also evident that the uniform distribution gives the worst tolerance area compared with

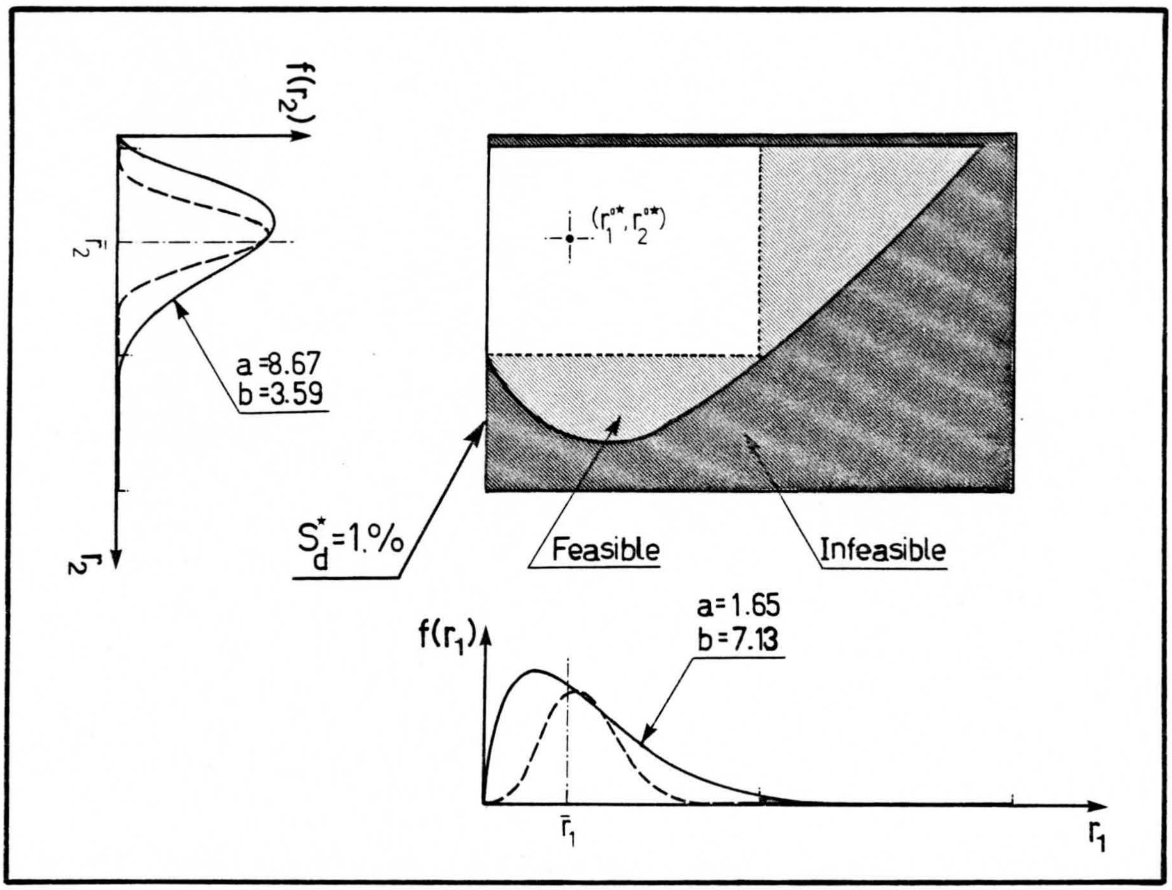


Figure 3.12 Predetermined design scrap - Beta distribution.

any other distribution.

3.5.2 Optimum Design Scrap, Zero Manufacture Scrap, $[P]_9$.

Similar to $[P]_8$, the manufacture scrap is assumed to equal zero which implies a hypothetical situation where a particular manufactureing process could be found to match the final optimum machining capabilities and to follow a known probabilistic distribution. Also, the system nominals r_1^{o*} and r_2^{o*} are assumed to be fixed and the random variables z_1 and z_2 follow a uniform distribution. The only change is that the system scrap percentage S_d will not have a predetermined value and it will be optimally determined. This can be done by allowing S_d to affect directly the problem objective, since S_d is a function of the tolerance design variables $[\underline{T}]_s$, which are greater or equal to the full acceptance tolerances $[\underline{T}]^*$. The production cost decreases and the system design scrap increases, if the tolerances $[\underline{T}]_s$ increase. However, the system cost increases if the design scrap increases. Therefore, a break even optimum value for both S_d and $[\underline{T}]_s$ could be found by selecting a suitable design objective which minimizes the total system cost. The unconstrained problem can be stated as follows

$$[\underline{X}]_9 = [\underline{T}]_s = [t_{1s}^+, t_{1s}^-, t_{2s}^+, t_{2s}^-]^T$$

$$U_9 = u_9 \ (1 + S_d/100.) = \text{minimum} \qquad (3.17)$$

where $u_9 = U_6 \Big|_{\substack{r_1^o = r_1^{o*} \\ r_2^o = r_2^{o*}}}$

Optimum Solution

$$0.110 \leq r_1 \leq 0.154$$

$$0.161 \leq r_2 \leq 0.200 \qquad S_d^* = 0.001\% \tag{3.18}$$

Area Gained = 14.7%

The solution is illustrated in Figure 3.13.

3.5.3 Optimum Design and Manufacture Scrap, $[P]_{10}$

The design nominals r_1^{o*} and r_2^{o*} are assumed known and fixed. Manufacturing processes to be used to produce the design are chosen and their probabilistic distributions are determined, which do not depend on any of the optimization variables. The objective is to determine the optimum tolerance values of r_1 and r_2 that give the best combination of S_d and S_m. The problem could be mathematically formulated as follows

$$[\underline{X}]_{10} = [\underline{T}]_s = [t_{1s}^+, t_{1s}^-, t_{2s}^+, t_{2s}^-]^T$$

$$U_{10} = u_9 \ (1 + S_d/100.) + u_{10} \cdot S_m/100. \tag{3.19}$$

where $u_{10} = (u_9 - u_{46})$ = machining cost.

u_{46} = assembly cost, Equation (2.20).

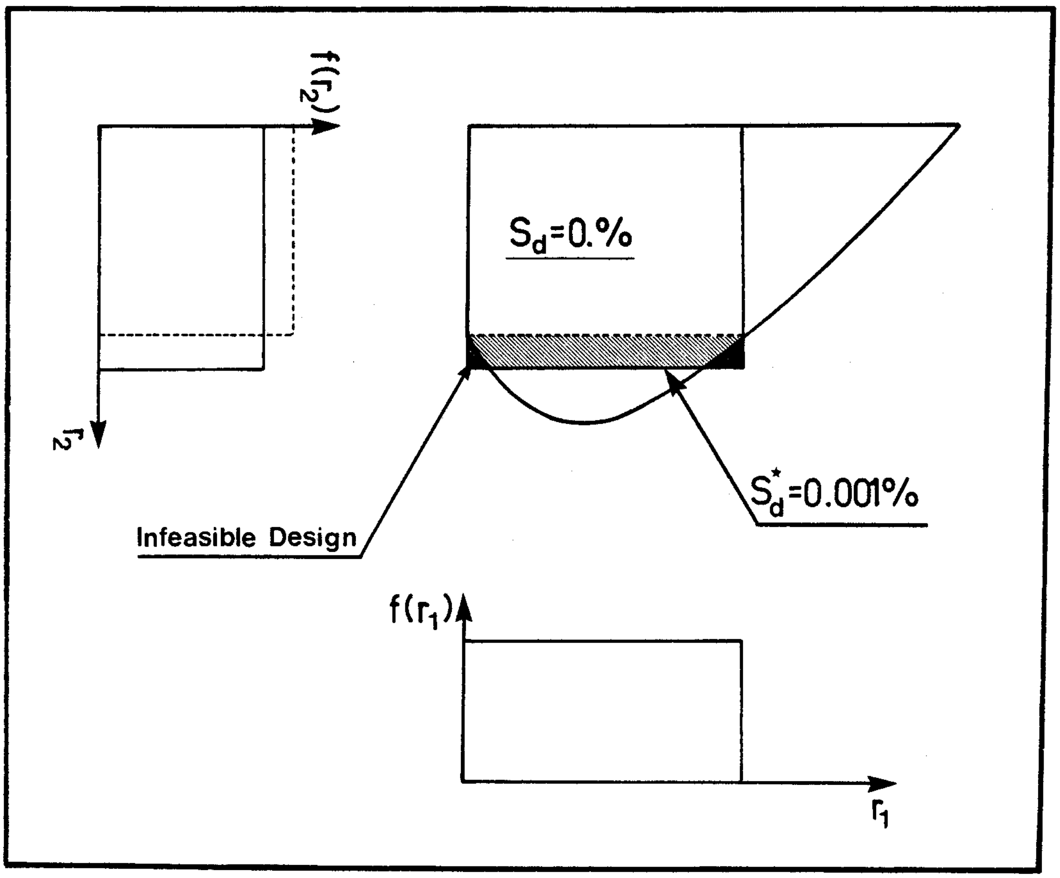


Figure 3.13 Optimum design scrap.

Both r_1 and r_2 are assumed to follow a uniform distribution between the limits

$$0.108 \leq r_1 \leq 0.160$$

$$0.150 \leq r_2 \leq 0.205$$

Optimum Solution

$$0.110 \leq r_1 \leq 0.156$$

$$0.161 \leq r_2 \leq 0.200$$

$$\text{Area Gained} = 20\% \qquad (3.20)$$

$$S_m^* = 38\%$$

$$S_d = 0.323\%$$

The solution is illustrated in Figure 3.14.

It is evident from the previous optimum result that the design variables $[\underset{\sim}{T}]_s$ are allocated to give the best trade-off between the manufacture scrap and the design scrap. Since for an incremental increase in the values of the system tolerances, the estimated value of the system design scrap increases while the computed value of the system manufacture scrap decreases. The problem solution, however, depends mainly upon both the shape of the cost objective and the distributions of the random variables. The assumed uniform distributions represent the worst condition, and their limits do not coincide with reality, however, the problem served its aim by exaggerating the effect of the manufacture scrap on the design scrap value.

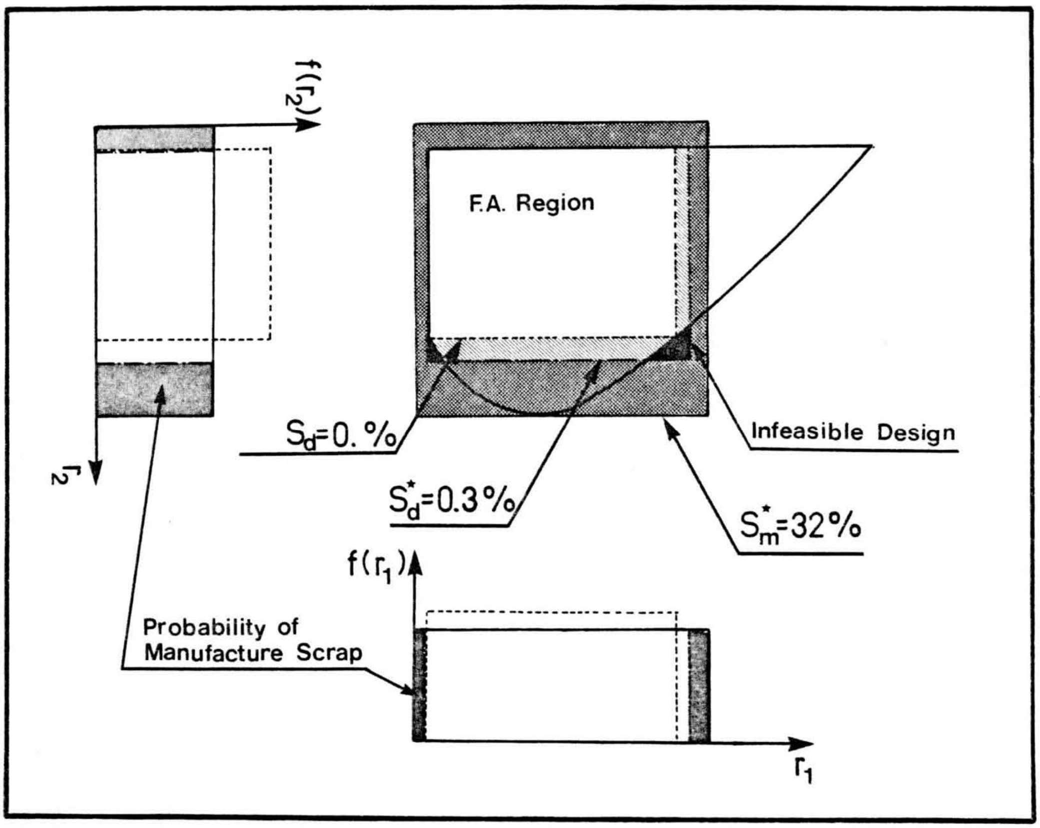


Figure 3.14 Optimum design and manufacture scrap.

3.6 Conclusions

Through the distinction between the design and the manufacturing scrap a link was constructed between designing a system to meet a set of specifications and manufacturing it. An engineering system in general could be decomposed into a set of smaller sub-systems, [53, 54]. In this event the tolerance assignment previously introduced facilitates the choice of standard or purchased sub-systems to match with the overall system performance, since the tolerances associated with the sub-system's specifications are widened as much as possible and as economically as possible.

The misuse of a normal distribution as applied to engineering phenomena is also pointed out, and the advantages of employing probabilistic distributions with a finite range - e.g., beta and uniform - are clearly elaborated.

The space regionalization technique originally suggested by Scott et al. [41], and further improved by Gopal [42] was used by the author in conjunction with optimization techniques to determine the system design and manufacturing optimum scrap percentages. Additional computational savings were also achieved by detecting the active corners and not analyzing their adjacent cells. Also a set of mathematical expressions were introduced to compute the relative number of cells needed for partitioning the regionalization space for the system random variables.

Four different cases were studied for the same example to draw the attention to various conclusions.

They are:

a) Comparing the tolerance percentage gained by allowing various amounts of system scrap.

b) Displaying the effect of the random variable distributions in allocating the system optimum tolerances and consequently its scrap.

c) Elaborating the fact that a uniform distribution leads to the worst estimates for the system optimum variables.

d) Introducing a simple but practical way of combining the system scrap with the objective function.

e) Differentiating the effect of the system design scrap versus the manufacturing scrap on the optimization outcomes.

CHAPTER 4

THE UPPER REGIONALIZATION BOUND

4.1 Introduction

The space regionalization technique described in Chapter 3 was introduced to save some computational effort in estimating the system design scrap. The full acceptance solution FA, has been elaborated and discussed in Chapter 2. It served as a lower bound within which the regionalization technique does not have to be utilized. In this chapter, on the other hand, the upper bound region, R^U, will be mathematically defined. An algorithm will be introduced and illustrated by geometric interpretations, to define the region in the system domain. The upper bound region contains the tolerance limits beyond which there will be no need to utilize the regionalization technique, thus reducing the computational effort required. This upper bound region just encloses the feasible region R_c, as illustrated in Figure 4.1.

4.2 Definitions

a) The system feasible region, R_c: It contains all the design outcomes that satisfy the system nonlinear inequality constraints. R_c is illustrated in Figure 4.1, and it could be defined as

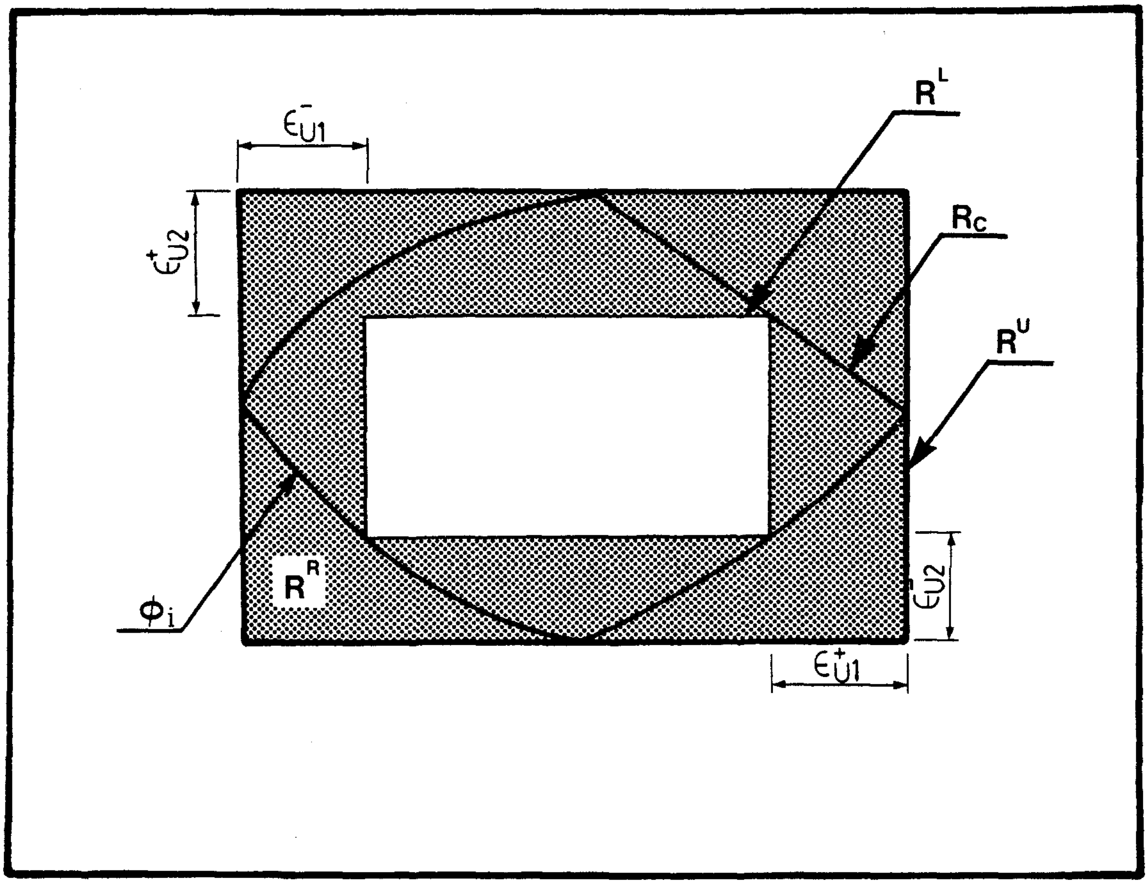


Figure 4.1 The lower and upper feasible bounds, R^L_a & R^U, versus the feasible region, R_c.

$$R_c = \{\underline{X} \mid \underline{\Phi} (\underline{X}) \geq \underline{0}\} \qquad (4.1)$$

b) The regionalization lower bound, R^L: It consists of 2n orthogonal surfaces that form a polytope, as illustrated in Figures 4.1 and 4.2. The lower bound lies fully inside the feasible region R_c, and it depends upon the location of the system nominals $\underline{X}^o$ and also on the shape of the closed feasible region. The lower bound has 2^n corners C, 2n norms S_n, and at least n of its corners have to be on the feasible region bound, where $\phi_i = 0$. The lower bound has to contain the full acceptance region R_{FA}, however, the reverse is not true, because R_{FA} depends on the system objective. The lower bound region could mathematically be expressed as follows:

$$R^L = \{\underline{X} \mid \underline{\Phi} (\underline{X}) \geq \underline{0};\ [\underline{X}^o + \underline{t}_L^+] \leq \underline{X} \leq [\underline{X}^o - \underline{t}_L^-]\}$$

$$R^L \subset R_c \qquad (4.2)$$

$$R_{FA} \subset R^L$$

where $\underline{t}_L^+$ and $\underline{t}_L^-$ are the maximum feasible values of the positive and the negative tolerances associated with the system nominals, respectively.

c) The regionalization upper bound, R^U: It contains both the lower bound and the feasible regions, and it also consists of 2n orthogonal surfaces. The upper bound is shown in Figure 4.1, and it could mathematically be expressed as follows:

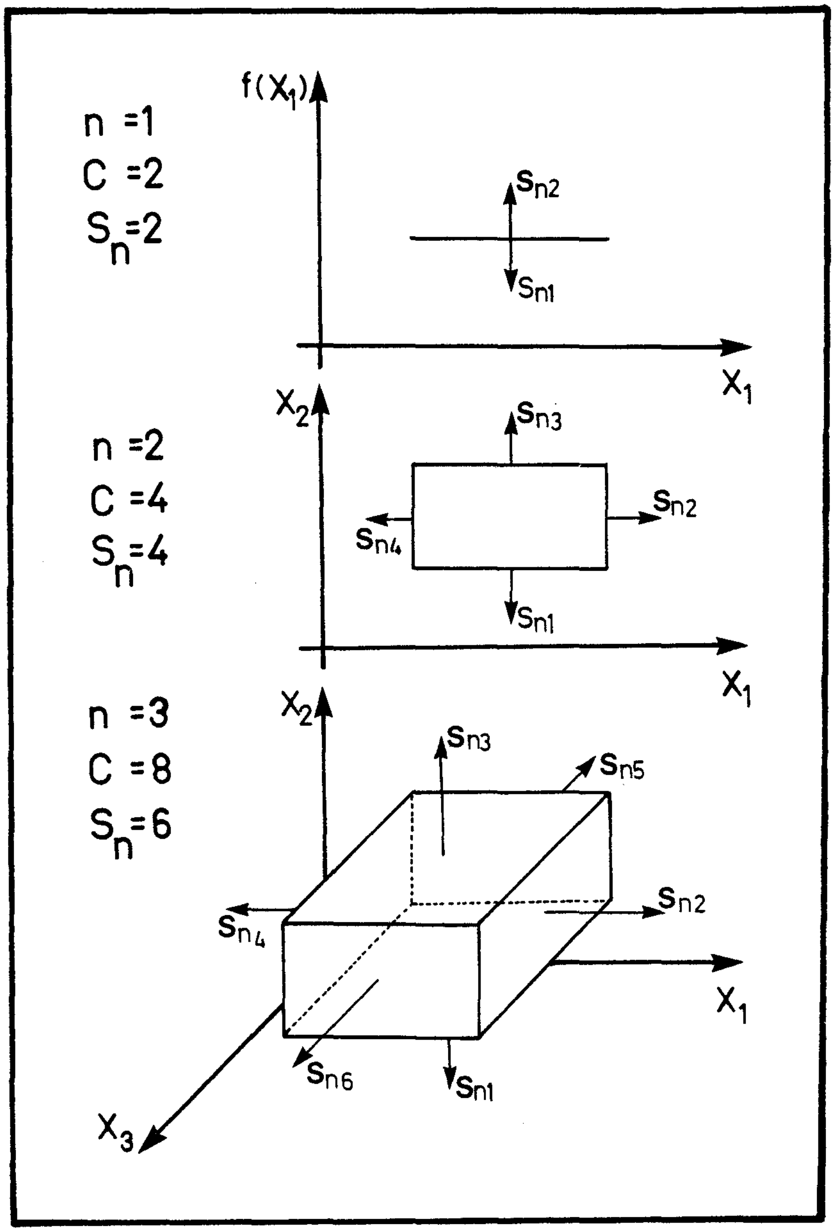


Figure 4.2 Graphical representation of one, two and three dimensional polytope and their corresponding corners and norms.

$$R^U = \{\underline{X} \mid [\underline{X}^o + \underline{t}_U^+] \leq \underline{X} \leq [\underline{X}^o - \underline{t}_U^-]\} \tag{4.3}$$

$$R_c \subset R^U$$

where $\underline{t}_U^+$ and $\underline{t}_U^-$ are the minimum values of the positive and negative tolerances that guarantee full confinement of the feasible region.

d) The regionalization region, R^R: It is the n-dimensional region that can be partitioned into cells while computing the system design scrap. It is the net outcome from subtracting the upper bound region R^U from the lower bound region R^L, as illustrated in Figure 4.1.

$$R^R = \{\underline{\varepsilon} \mid \underline{0} \leq \underline{\varepsilon} \leq \underline{\varepsilon}_U\}$$

$$R^U \subset R^R \tag{4.4}$$

$$R^L \not\subset R^R$$

where $\underline{\varepsilon}_U$ is the upper limit of the regionalization region defining variable. It is defined below.

e) Regionalization region defining variable, $\underline{\varepsilon}$: It defines the region between the upper and the lower bounds in the domain of the system random variables, as illustrated in Figure 4.1. Since the nominal vector $\underline{X}^o$ is assumed fixed throughout the search for the tolerances that define R^U, $\underline{X}$ can be transformed into the $\underline{\varepsilon}$ domain, where $\underline{\varepsilon}$ contains the defining variables $\underline{\varepsilon}^+$ on the upper side and $\underline{\varepsilon}^-$ on the lower side, with corresponding maximum values $\underline{\varepsilon}_U^+$ and $\underline{\varepsilon}_U^-$.

Thus we have

$$\underline{\varepsilon} = [\underline{\varepsilon}^{+}, \underline{\varepsilon}^{-}]^T$$

$$\underline{\varepsilon}_U^{+} = \underline{t}_U^{+} - \underline{t}_L^{+} \qquad (4.5)$$

$$\underline{\varepsilon}_U^{-} = \underline{t}_U^{-} - \underline{t}_L^{-}$$

$$\underline{\varepsilon}_U = [\underline{\varepsilon}_U^{+}, \underline{\varepsilon}_U^{-}]^T$$

$$\underline{0} \leq \underline{\varepsilon} \leq \underline{\varepsilon}_U$$

where $\underline{t}_L$ and $\underline{t}_U$ were defined previously in Equations (4.2) and (4.3), respectively.

f) The scrap slack value, q_i (x_i): It is a measure of how close a given value of x_i is to the apparent vertex or the extreme boundary of the feasible region. It is graphically illustrated in Figure 4.3 for two random variables, for which

$$q_1 = f_2/t_2$$

$$q_2 = f_1/t_1$$

However, for an n-dimensional system, it must be defined as follows:

$$\underline{q} = [\underline{q}^{+}, \underline{q}^{-}]^T \qquad (4.6)$$

$$q_i^{(\pm)} = \prod_j^n f_j^{(\pm)} / \prod_j^n t_j$$

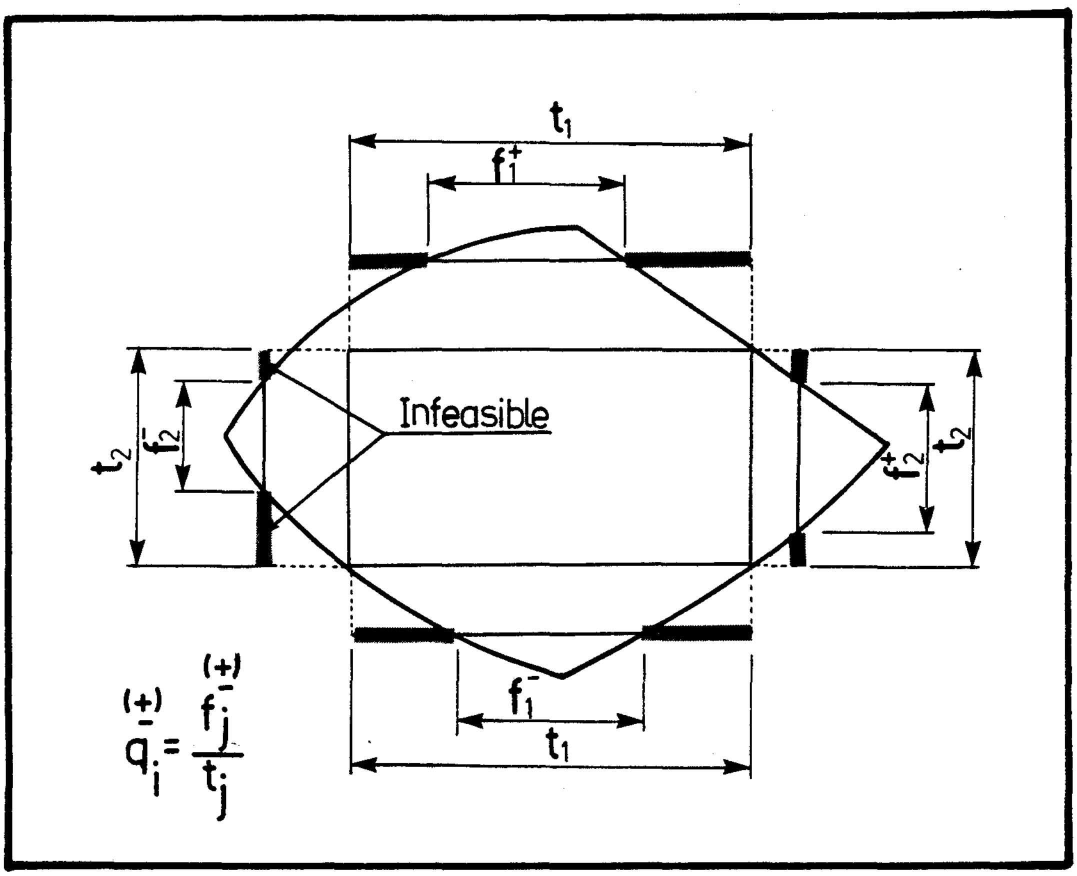


Figure 4.3 The scrap slack value, q.

where

$j = 1, 2, \ldots, n; \quad j \neq i$

$t_j = t_j^+ + t_j^-$

$f_j^+ =$ the feasible portion of t_j on the positive side of t_i.

For the purpose for which it is defined, q_i is a function only of x_i; and n is the number of those variables having tolerances.

g) Line regionalization technique, LR: It follows the same concept as the conventional space regionalization technique described in Chapter 3, where the domain under study is partitioned into cells. The feasibility of each of the cells is checked against the system constraints. Line regionalization technique is used in the ensuing strategies to estimate the scrap slack value q_i. Line and space regionalization techniques differ, however, in two aspects. The first is that in LR the random variables $\underline{\varepsilon}$ are assumed to follow a uniform distribution because the probability of occurrence of events in a cell is not important and does not affect the technique outcomes. In the space regionalization, however, the random variables $\underline{X}$ may follow any probabilistic distributions, which depend on the corresponding manufacturing process. The second variation of LR from the space regionalization technique is that a cell representative is not located in the center of the cell, as shown in Figure 3.3, but in the center of the cell's nearest edge to the lower bound region R^L, as shown in Figure 4.4. It is only necessary to check the boundaries of the regions closest to the R^L region. Employing the LR technique will reduce the number of feasible cells considered to be infeasible.

4.3 Strategy

The basic concept of the method is to determine R^U by minimizing the scrap slack value $\underline{q}$ for each variable, thus forcing the bounds defining R^U to fall on the extreme boundaries of the feasible region R_c. The algorithm adopted is divided into three strategies to be followed successively. They are, the primary upper limit strategy, the upper limit strategy with checking of sides, and a strategy to determine the acceptable upper bound. The first strategy uses the R^L region tolerance bounds in estimating $\underline{q}$ for various iterations of the strategy. The optimum outcome of the primary strategy is then considered as a starting region for the strategy which checks the sides. Here the tolerance bounds, that have to be used in estimating $\underline{q}$, are updated in each iteration of the strategy. These bounds ought to be beyond the R^L limits. To insure that the upper bound region limits lie completely within the joint probability density function of the system, a final check has to be made.

The primary and the checking of sides strategies fall into the category of direct search optimization techniques. They rely on the sequential examination of trial solutions in

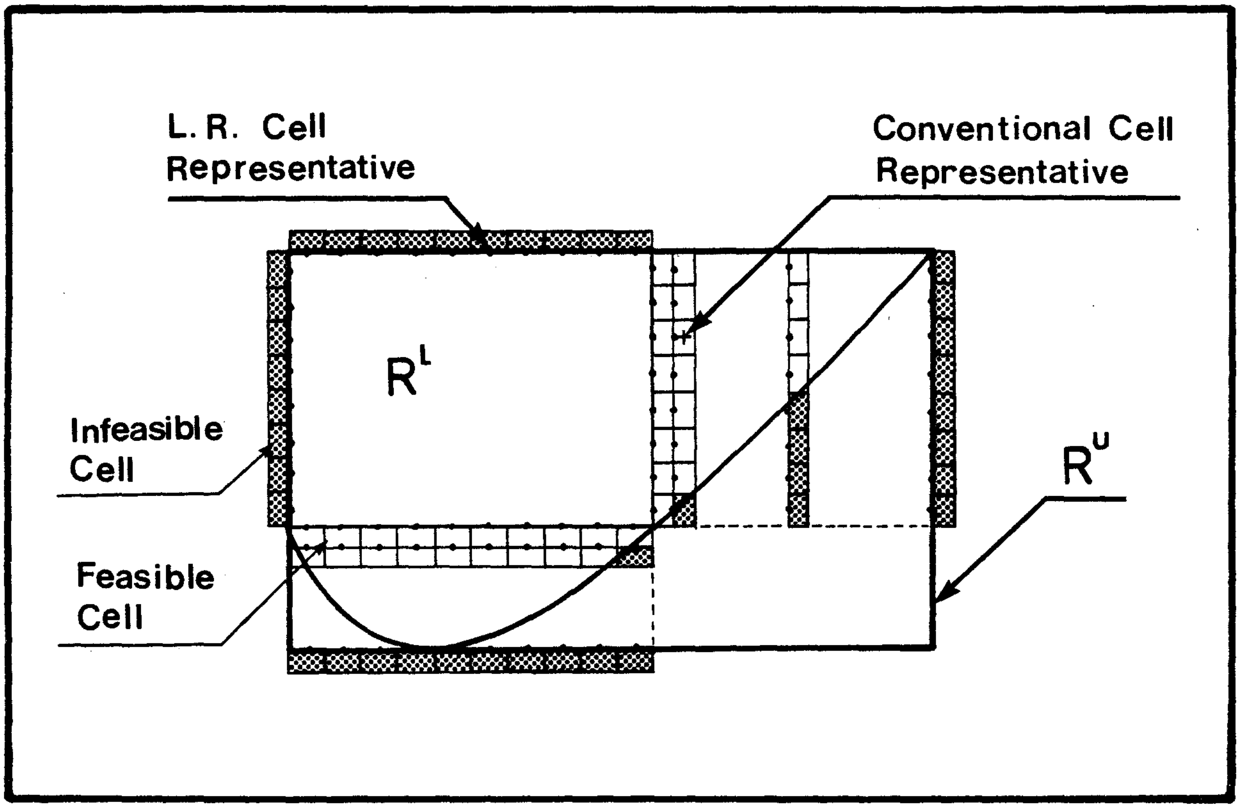


Figure 4.4 Line regionalization.

which each solution is compared with the best obtained up to that point. This is done without evaluating either the objective or the constraint's derivatives.

Solution of the following nonlinear optimization problem will determine the upper bound region.

$$U_{UL} = \sum_{i}^{2n} q_i \ (\underline{\varepsilon}) = \text{minimum}$$

subject to (4.7)

$$\Phi(\underline{\varepsilon})_{UL} = \sum_{i}^{2n} q_i \ (\underline{\varepsilon}) \geq 0$$

where the scrap slack value $\underline{q}$ is minimized until it reaches its zero limit. The constraint ϕ is defined in order to bound the expansion of the upper limit region to the minimum.

4.3.1 Primary Upper Limit Strategy

The R^L region is assumed to be known, and consequently the $\underline{t}_L$ vector is defined. The strategy is schematically illustrated in Figure 4.5, and proceeds as follows:

Step 1 Set i = 1, where i is the number of the tolerance variables. It varies between one and 2n.

Step 2 Set j = 1, where j is a counter of the number of iterations for each variable. Compute $\underline{S}_i$, the direction of inflating the upper bound region away from R^L.

Step 3 Define the limits of the regionalization

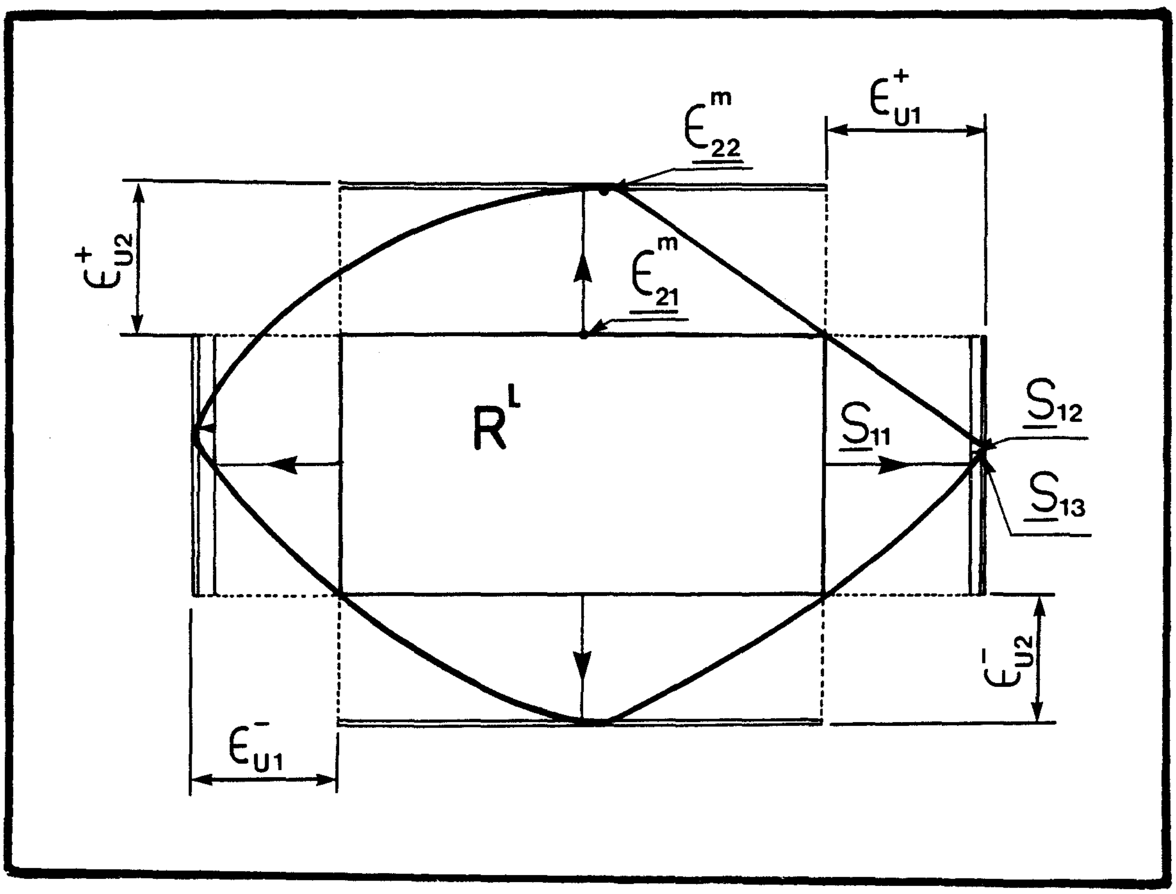


Figure 4.5 Upper limit strategy.

checking domain. They equal the corresponding limits of the R^L region for the 2n tolerance variables of the system. For j greater than one, the checking limit for the i variable has to be updated with the $\varepsilon_{i,j-1}$ value from the previous iteration.

Step 4 Partition the checking domain into R_{mi} cells defined in Equation (3.11).

Step 5 Using the LR technique, explained in sub-section 4.2.g, estimate $q_{ij}(\varepsilon_i)$.

Step 6 If q_{ij} equals zero, go to Step 9. Otherwise, compute $\underline{\varepsilon}^m_{ij}$, the mid point of the feasible portion of the checking domain.

Step 7 Determine the optimum step size α^*_{ij}. Starting from the feasible point $\underline{\varepsilon}^m_{ij}$ and proceeding in the given direction $\underline{S}_i$ until one of the system constraints approaches zero, as explained in sub-section 4.3.2.

Step 8 Update ε_i

$$\varepsilon_{ij} = \varepsilon_{i,j-1} + \alpha^*_{ij} S_i \tag{4.8}$$

Check the improvement in α^*_{ij} by computing p

$$p = \frac{\alpha^*_{ij} - \alpha^*_{i,j-1}}{\alpha^*_{ij}} \tag{4.9}$$

If p is greater than a predetermined improvement factor, p_s, increase the counter j by one and go to Step 3.

Step 9 $\varepsilon^U_i = \varepsilon_{ij}$

If i equals 2n, then stop, otherwise set i = i+1 and go to Step 2.

4.3.2 The One-dimensional Search Strategy

Consider the m constraint functions

$$\phi^{\ell} (\underline{X}^o, \underline{t}) \geq 0 \; ; \quad \ell = 1,2,\ldots m \tag{4.11}$$

of the n-variable vector $\underline{X}^o$ and the 2n-variable vector $\underline{t}$. If $\underline{X}^o$ is assumed fixed throughout the search, Equation (4.11) could be transformed to the $\underline{\varepsilon}$ domain by using Equation (4.5). Therefore, the system set of constraints could be expressed as

$$\phi_{\ell} (\underline{\varepsilon}) \geq 0 \quad ; \quad \ell = 1,2,\ldots m \tag{4.12}$$

of the 2n regionalization region defining variables vector $\underline{\varepsilon}$. For any feasible point, $\underline{\varepsilon}^m$ say, Equation (4.12) will be satisfied. A point on the boundary of the feasible region R_c is defined as a point for which

$$\begin{aligned} &\phi_k (\underline{\varepsilon}) = 0 \\ \text{and} & \\ &\phi_{\ell} (\varepsilon) \geq 0 \qquad ; \; \ell = 1,2,\ldots m \\ &\qquad\qquad \text{and} \quad \ell \neq k \end{aligned} \tag{4.13}$$

for some k ε[1,m]. During the procedure of defining the upper limit of the regionalization region R^U; we wish to find a point on the infeasible region starting from some feasible point $\underline{\varepsilon}^m$ and proceeding in a given direction $\underline{S}$.

More specifically, we wish to determine the step size α^* for which

$$\phi_k (\underline{\varepsilon}^m + \alpha^* \underline{S}) = 0 \tag{4.14}$$

$$\phi_\ell (\varepsilon^m + \alpha^* \underline{S}) \geq 0$$

where $\ell = 1,2,\ldots,m$; $\ell \neq k$

for some k. The one-dimensional search procedure which has been implemented is based on the secant method and proceeds as follows:

Step 1 Evaluate $\phi_\ell (\underline{\varepsilon}^m)$, $\ell = 1,2,\ldots,m$.
Define the initial values for both the base and the lower limits of an auxiliary variable $\underline{h}$ as follows

$$h_i = h_\ell^L = \phi_\ell \quad ; \quad \ell = 1,2,\ldots m. \tag{4.15}$$

and also

$$\alpha^o = \alpha^L = 0$$

Set α to a predetermined initial value.

Step 2 Evaluate $\phi_\ell (\varepsilon^m + \alpha \underline{S})$. If all the m constraints are greater than zero double the value of α and repeat, otherwise define the updated and the upper value of both α and $\underline{h}$ as follows:

$$\alpha^1 = \alpha^U = \alpha$$

and (4.16)

$$h_\ell^1 = h_\ell^U = \phi_\ell$$

for all $\phi_\ell < 0$

Let $L = \{\ell \mid \phi_\ell(\underline{\varepsilon}^m + \alpha\, \underline{S}) < 0\}$ (4.17)

Step 3 If $\dfrac{\alpha^U - \alpha^\ell}{\alpha^U} > \Delta\, \varepsilon_i$; (4.18)

then proceed if not stop and take the current value of α as α^*, where $\Delta\, \varepsilon_i$ is the minimum allowable change in ε_i.

Step 4 Choose the successive value of α by selecting the minimum out of L calculated values as follows:

$$\alpha = \min_{\ell \varepsilon L} \frac{\alpha^1 - h_\ell^1\,(\alpha^1 - \alpha^o)}{h_\ell^1 - h_\ell^o} \tag{4.19}$$

If $\alpha < \alpha^L$

then $\alpha = \alpha^\ell$

If $\alpha > \alpha^U$

then $\alpha = \alpha^U$

Set $\alpha^o = \alpha^1$

$\alpha^1 = \alpha$

$h_\ell^o = h_\ell^1$; $\ell \varepsilon L$

$h_\ell^1 = \phi_\ell(\underline{\varepsilon}^m + \alpha^1 \underline{S})$

This step is essentially the secant method for determining α.

Step 5 If $h_\ell^1 > 0 \qquad \ell \varepsilon L$

then set $\alpha^L = \alpha^1$ (4.21)

$$h_\ell^L = h_\ell^1$$

and repeat from step 3. Otherwise

$$\alpha^U = \alpha^1$$

$$L = \{\ell \mid \phi_\ell(\underline{\varepsilon}^m + \alpha^1 \underline{\varepsilon}) < 0\} \qquad (4.22)$$

$$h_\ell^U = h_\ell^1 \quad ; \quad \ell \varepsilon L$$

and repeat from step 3.

At every step of the procedure there is an upper and lower bound on the step size α, and when the interval $[\alpha^L, \alpha^U]$ gets small enough, the procedure terminates. Moreover, the minimum value of the constraints at any step is equivalent to the value of the unconstrained objective function in the conventional optimization.

4.3.3 Upper Limit Strategy With Checking of Sides

The primary upper limit strategy, described in subsection 4.3.1, utilized the lower bound region tolerance limits, $\underline{t}_L$, in estimating the scrap slack value using a line regionalization technique. The primary upper limit strategy is a necessary but not a sufficient algorithm to identify the R^U region, since a part of the feasible region

R_c could still be outside R^U (primary), as indicated by the example of Figure 4.6. To guarantee full enclosure of the feasible region, a complementary checking strategy is necessary. It is illustrated in Figure 4.6 and proceeds as follows:

Step 1 Execute the primary upper limit strategy and define the R^U (primary) region boundary by $\underline{\varepsilon}_1$. The ε subscript indicates the number of the iteration.

Step 2 Check the 2^n corners of the R^U (primary) region against the system constraints. If all the corners are infeasible, stop

$$R^U = R^U \text{ (primary)}$$

$$\underline{\varepsilon}^U = \underline{\varepsilon}_1 \tag{4.23}$$

otherwise, identify the feasible corners, C_f, and proceed.

$$C_f < C = 2^n \tag{4.24}$$

Number of the side checking iterations equals

$$J = n.\ C_f \tag{4.25}$$

While the number of iterations I, for each of the J side checking iterations depends on the closed feasible region bounds.

$$j \ \varepsilon\ [1,J] \quad ; \quad i \ \varepsilon\ [1, I] \tag{4.26}$$

Set j = 1 and i = 1.

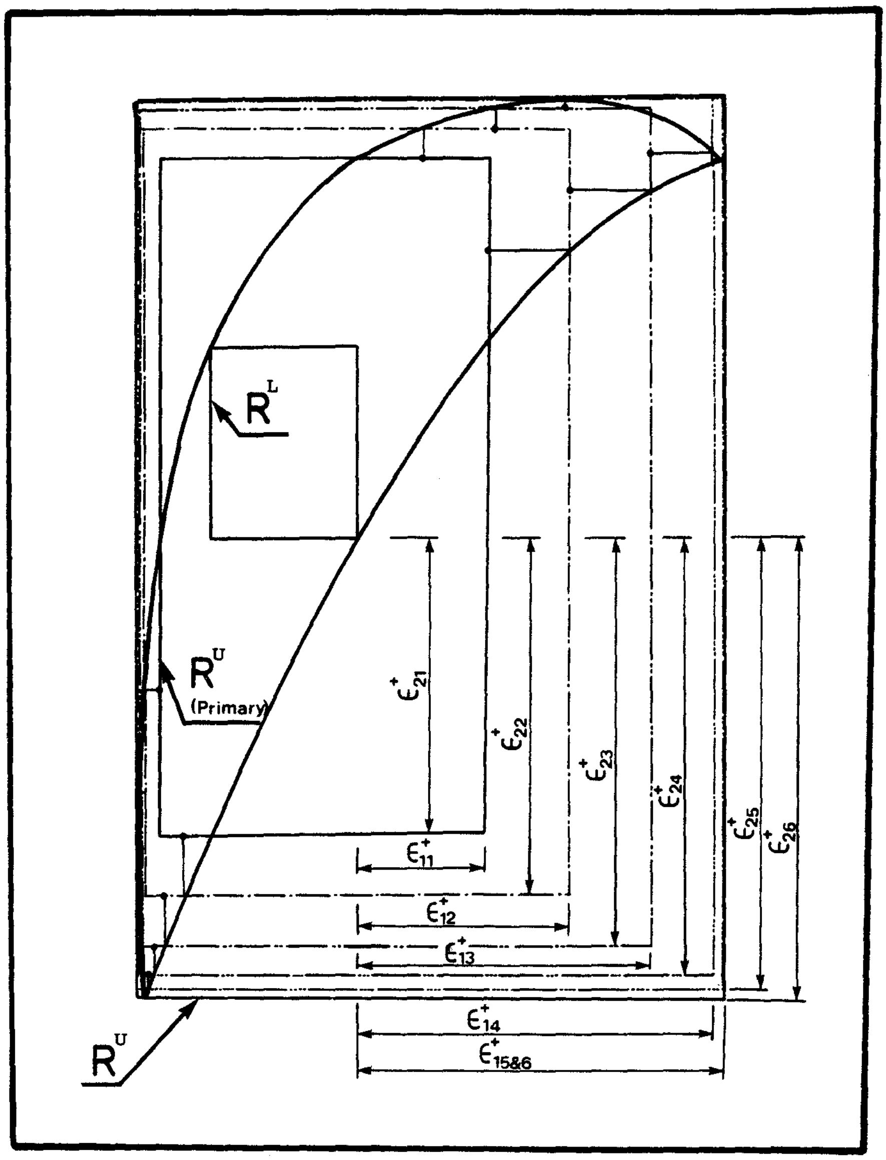


Figure 4.6 Upper limit strategy with checking of sides.

Step 3 Define the line regionalization checking domain. It covers the area between the current upper limit corner for the i iteration and the projected corresponding corner for the (i-1) iteration. This defines the j checking side. Note that for the starting condition, when (i-1) equals zero, the corresponding location of the feasible corner of the lower bound region is utilized to define the LR domain.

Step 4 Estimate $q_{ij}(\underline{\varepsilon})$, where $\underline{S}_j$ is constant for the I iterations. Follow the same procedure as that explained in the primary strategy, until no improvement is eminent.

Step 5 If $j = J$ stop, otherwise set

$j = j + 1$

4.4 Acceptable Upper Bound

It was assumed throughout both the primary and the checking sides strategies that the upper limit bound region is a sub-region or fully contained within the manufacturing tolerance region for each random variable. That is to say,

$$R^U \subset R_{mt}$$

where,

$$R_{mt} = \{\underline{t} \mid t_i \geq t_{min,i},\ t_{i+n} \leq t_{max,i}\ ;\ i \in [1,n]\} \quad (4.27)$$

In other words, the manufacturing scrap percentage, estimated separately for each of the n independent design variables, is greater than or just equal to zero.

This is a hypothetical assumption which in practice is not usually fulfilled. Therefore, a sufficient check has to be done after the one dimensional search is carried out for both the primary and the checking of sides strategies. If Equation (4.27) does not apply for any of the intermediate bounds, the corresponding lower or upper manufacturing tolerance limit should be considered as the acceptable upper limit bound, as illustrated in Figure 4.7.

Therefore,

$$R_a^U \subset R_{mt}$$

$$R_a^U \subset R^U \qquad (4.28)$$

but $$R^U \not\subset R_{mt}$$

or $$R_{mt} \not\subset R^U$$

4.5 Discussion and Conclusions

Having described mathematically, and with the aid of two dimensional graphical representations, a strategy to define the upper regionalization bound, R^U of a system, it is important to classify some dependency relationships between various regions that have been defined earlier. R^U does not depend on R^L and it is unique for each system. On

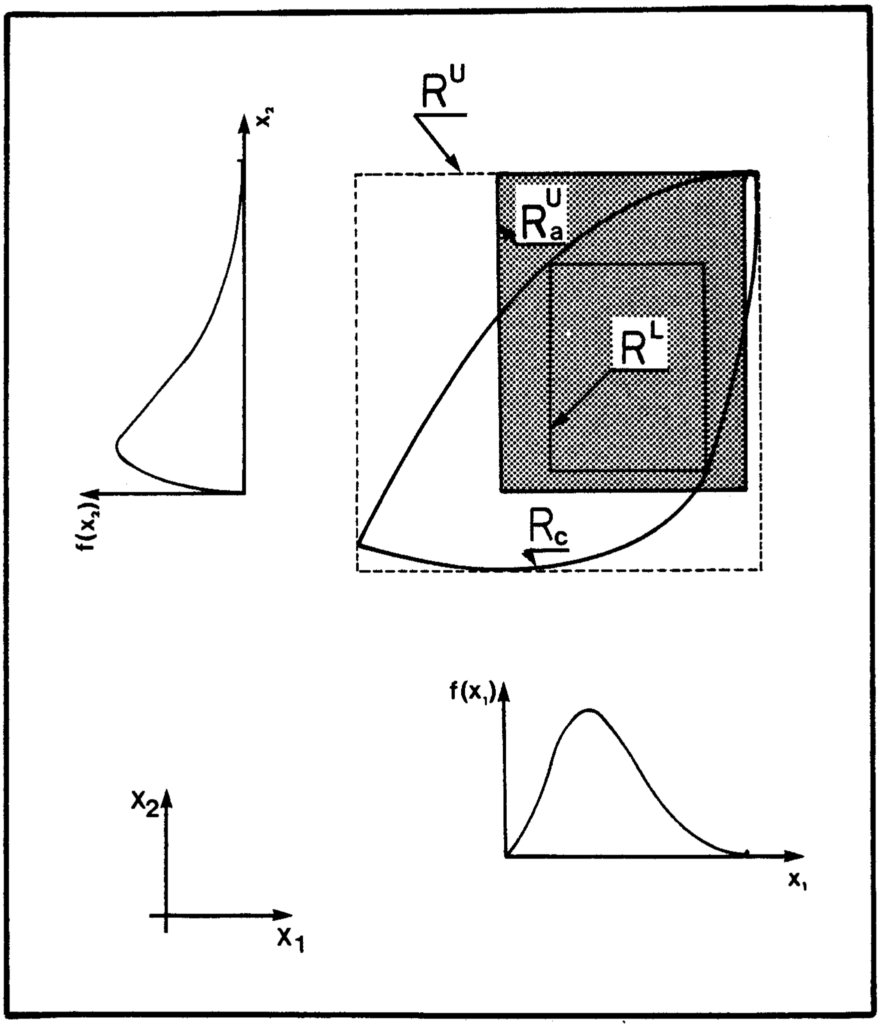


Figure 4.7 Acceptable upper limit.

the other hand, R^L depends upon the location of the nominal optimum point $\underline{X}^o$, however it does not depend on the shape of the objective function as long as it maximizes the system tolerances, as has been shown in Chapter 2. Therefore, it may be necessary to define several R^L regions during the course of allocating the optimum performance condition of a system. However, it is only necessary to define R^U once.

The main advantage of transforming the t-domain into ε-domain, throughout the previously illustrated strategy, is to avoid specifying an additional set of inequality constraints to guarantee that the optimum upper limit tolerances are greater than or equal to their corresponding lower limit tolerances. While using the ε-domain, however, it is only sufficient to use the absolute values of the defining variables $\underline{\varepsilon}$, since this ensures that their values are always positive irrespective of the optimization outcome.

The R^U region does not depend on the convexity condition of R_c. However, if the estimate of a scrap slack value q_i was due to a non-continuous domain of f_j, then the mid point vector $\underline{\varepsilon}_i^m$ should be computed to center the largest continuous feasible portion of the f_j domain.

The strategy has been utilized in solving the examples in Section 3.5, where a substantial computational savings up to 30% was achieved. A primary check was carried out for each random tolerance variable t_i at the beginning of every optimization iteration, to detect whether t_i lies beyond the upper regionalization limit or not. Therefore, if it happens to be outside the R^U region no cells will be generated and no checking will be carried out while the probability of scrap of the deleted infeasible portion of the inflated tolerance region is calculated directly knowing the corresponding distributions.

CHAPTER 5

IMPLEMENTATIONS

5.1 Introduction

The tolerance assignment problem for full and less than full acceptance design conditions was introduced. Strategies used to mathematically define the regionalization region with its lower and upper limits were consequently described. It is important, before reaching research conclusions, to clarify the possible sources of errors in estimating the optimum design variables and their corresponding upper and lower tolerances. And also, to discuss some proposed remedies to overcome them. The convexity assumption of the feasible region will be discussed, including an algorithm to detect the non-convexity, and to go around it to validate the allocation of the optimum regions. The limitations created due to the partitioning of the system domain using regionalization techniques are discussed. A sensitivity analysis of the errors in estimating the system scrap is also done.

Finally, several additional applications for the proposed methodology are stated to show a sample of the many design systems that could be investigated.

5.2 Convexity

The exactness of R^L or R_{FA} regions, defined in

Equation (4.3), depends on the validity of the convexity assumption of the R_c region, Equation (4.1). Therefore, it is necessary to verify the correctness of R^L before going any further in either estimating the system scrap or defining the R^U region.

The line regionalization technique, introduced in Chapter 4, will be utilized in conjunction with the one dimensional search, described in Section 4.3.2, to define the acceptable lower bound region, R_a^L. The correctness of R_a^L does not depend upon the complexity or the convexity of the constraints region, R_c.

The problem of defining R_a^L could be mathematically formulated as a nonlinear optimization problem as follows, in a manner similar to the formulation for obtaining R^U.

$$U_{LL} = \sum_{c}^{2n} q_i \; (\underline{\varepsilon}') \tag{5.1}$$

subject to

$$\Phi(\underline{\varepsilon})_{LL} = 2n - \left(\sum_{i}^{2n} q_i \; (\varepsilon') \right) \geq 0$$

where the scrap slack value q_i, defined in Equation (4.2), is maximized till it reaches unity. This has to be guaranteed for all of the 2n values of q_i. The constrain ϕ is defined in order to bound the contraction of the lower limit region to the mimimum. The acceptable lower region defining variable, $\underline{\varepsilon}'$, is similar to $\underline{\varepsilon}$ defined in Equation (4.5). And it could be mathematically expressed as follows:

$$\underline{\varepsilon}' = [\underline{\varepsilon}^{1+}, \underline{\varepsilon}^{1-}]^T$$

$$\underline{\varepsilon}_L'^{+} = \underline{t}_L^{+} - \underline{t}_{La}^{+} \qquad (5.2)$$

$$\underline{\varepsilon}_L'^{-} = \underline{t}_L^{-} - \underline{t}_{La}^{-}$$

$$0 \leq \underline{\varepsilon}' \leq \underline{\varepsilon}_L$$

where $\underline{t}_L$, $\underline{t}_{La}$ are two vectors that contain the tolerances between the system nominals and the R^L and R_a^L limits, respectively.

5.2.1 Primary Acceptable Lower Limit Strategy

The primary acceptable lower limit strategy is similar to the primary upper limit strategy introduced and explained in sub-section 4.3.1 with one exception, that the one dimensional search direction $\underline{s}_i'$ for any of the 2n variables is in the opposite direction of $\underline{s}_i$ defined in Equation (4.8).

The algorithm is schematically illustrated in Figure 5.1, and the optimum primary acceptable lower region is defined as R_a^L (primary).

5.2.2 Optimum Acceptable Lower Limit

The primary acceptable region confines only feasible design outcomes within its boundaries, although it is not necessarily the global optimum for Equation (5.1). Therefore, it is mandatory to go a step further and reoptimize the allocation of the lower bound while taking into account the

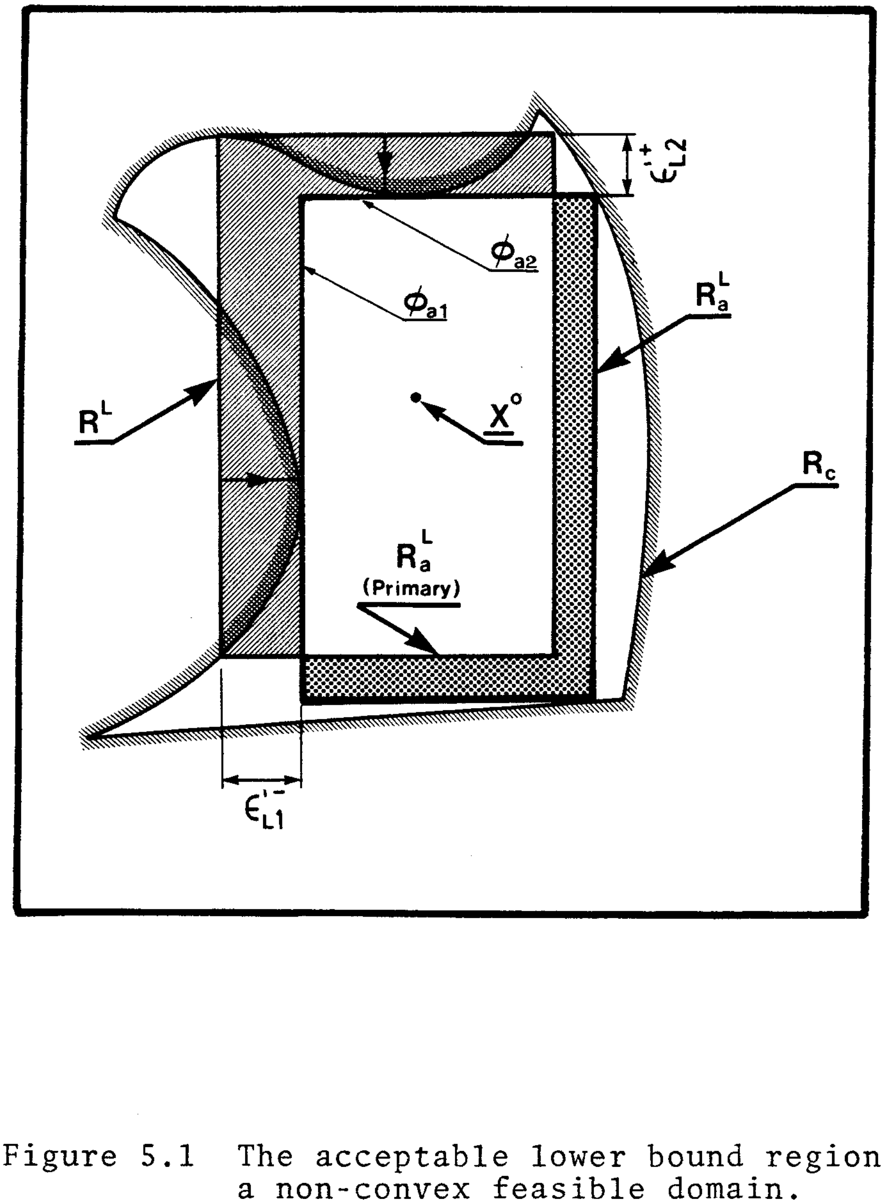


Figure 5.1 The acceptable lower bound region in a non-convex feasible domain.

updated maximum feasible value for the system tolerances. The problem could be mathematically expressed as follows:

$$\begin{aligned} & \text{minimize } U(\underline{X}) \\ \text{subject to} \quad & \underline{\Phi}(\underline{X}) \geq 0 \qquad (5.3) \\ & \underline{\phi}_a(\underline{T}) \geq 0 \end{aligned}$$

where U and $\underline{\Phi}$ are the original system objective and set of constraints. While $\underline{\Phi}_a$ is an additional set of linear constraints that confines the acceptable lower bound in the feasible regions by preventing its sides, that have been primarily contracted using the strategy introduced in sub-section 5.2.1, from re-expanding. It typically could be expressed as follows:

$$\phi_{ak} = t^{(\pm)}_{La,i} - t^{(\pm)}_{2} \geq 0$$

$$t^{(\pm)}_{La,i} = t^{(\pm)}_{Li} - \varepsilon'^{(\pm)}_{Li} \qquad (5.3)$$

where $k = 1, 2, \ldots, K$

$i = 1, 2, \ldots, 2n$

K is the number of successful contractions performed on the R_L region. It could equal zero, which proves that the system feasible domain is at least one dimensional convex between the bounds of the lower limit region. The optimum acceptable limit region R^L_a is shown in Figure 5.1, where K equals two.

5.3 Limitations

The space regionalization technique and the LR technique defined and described in Chapters 3 and 4 rely on dividing the regionalization region R^R, Equation (4.6), into a finite number of nonoverlapping cells. Each of these cells covers a sector of the joint density space of the system random variables. Depending upon the size of the cell and on its location in the density space, a weight is assigned to it. Since each cell occupies a portion of the system space, it consists of an infinite number of distinct designs. However, it is only represented by a single point, located in the center of the cell n-dimensional space - in the space regionalization technique - or in the center of the cell (n-1) - dimensional space and towards the R^L region.

To estimate the system scrap almost every cell in the R^R region - excluding those by the R^L active corners - must be tested against the system constraints. Since only the cell representative point, C Figure 5.2, is checked for feasibility, there is a probability of under or over estimating the system scrap. As illustrated in Figure 5.2, a cell might be considered feasible even if more than half of its volume is actually infeasible - i.e., outside the R_c region - or vice versa. Similar errors might be encountered while defining the R^U region or checking the convexity within either the R^L or R_{FA} regions.

Two separate solutions could be adopted to increase the confidence in the regionalization strategy outcomes.

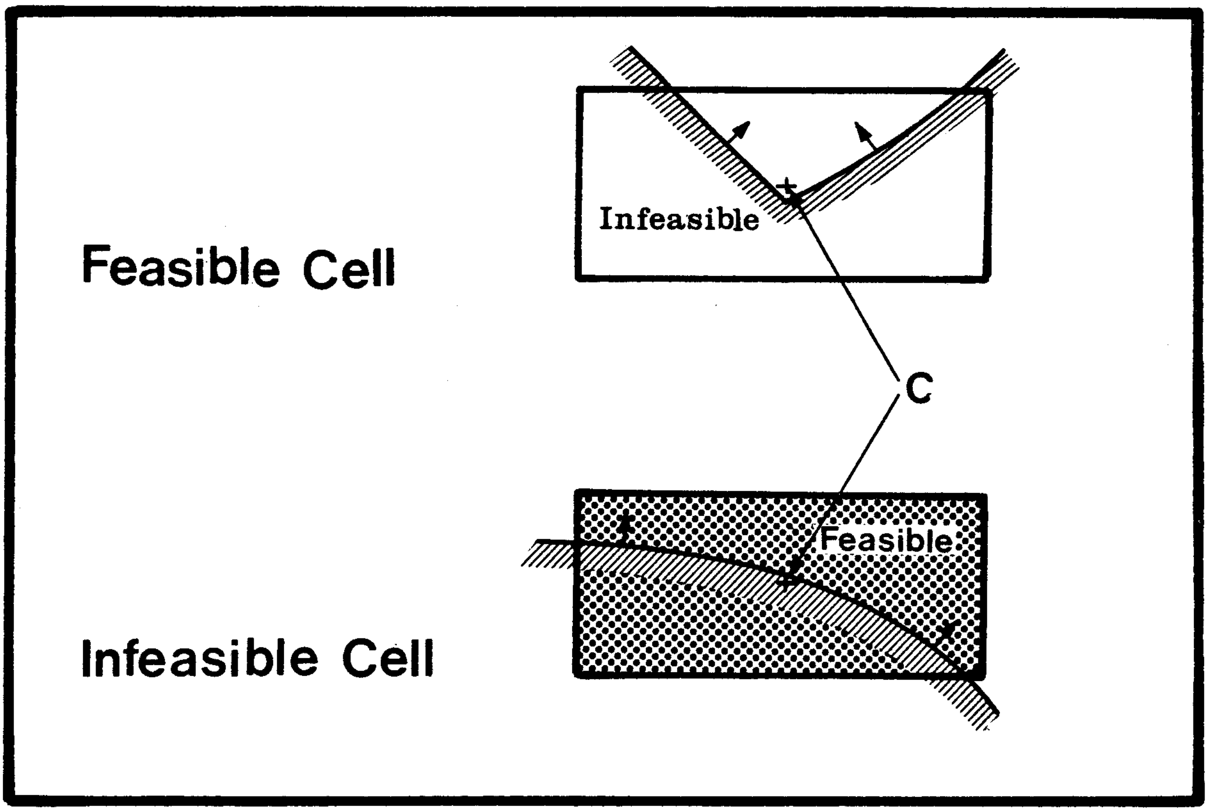


Figure 5.2 False estimate of cell's feasibility.

The first method would be to explore additional points within a cell at the final iteration, to see if they agree with the representative point C.

The second method would be to increase the total number of cells so that the system characteristics and feasibility will have less change within any of the cells.

5.3 Sensitivity

The accuracy of the system scrap estimates, that are derived by utilizing the space regionalization technique, depends mainly upon the location of the cell's representative points and on the total number of cells, N_{cell}. If a cell center is taken as its representative, and the system feasible domain as well as the system input specifications and objective are assumed fixed, then the only variable which affects the regionalization outcomes is N_{cell}. The absolute proportionality of the number of cells for each of the system random variables depends upon the range of each variable and its probability distribution. Three scalling factors $\underline{\lambda}_1$, $\underline{\lambda}_2$ and $\underline{\lambda}_3$ have been defined in Equation (3.11). They describe the various weights assigned for the various sectors of the system random variables. The exact proportion between $\lambda_{i,j}$ and $\lambda_{i+1,j}$, for $i\varepsilon[1,(n-1)]$ and $j\varepsilon[1,2,3]$, could be calculated as explained in Equation (3.12.a). Also the relation between $\lambda_{i,1}$ and $\lambda_{i,3}$ could be defined as explained in Equation (3.12.b). Therefore, irrespective of the system complexity, two reference values of $\lambda_{1,1}$ and $\lambda_{1,2}$ have to be independently chosen. In this section we will study the effect of the choice of λ's on the accuracy of the system estimated design scrap, i.e., the sensitivity of S_d with respect to the variation in the λ's.

The feasible region R_c and the lower regionalization bound region R^L defined in problem $[P]_4$ will be taken as a base for the sensitivity analysis. The problem upper bound region R^U is defined as the upper value of the regionalization region defining variable $\underline{\varepsilon}^U$. Therefore,

$$\underline{\varepsilon}^L = \underline{0}$$

$$\underline{\varepsilon}^U = [\underline{\varepsilon}^{+U}, \underline{\varepsilon}^{-U}]^T$$

$$= [\varepsilon_1^{+U}, \varepsilon_2^{+U}, \varepsilon_1^{-U}, \varepsilon_2^{-U}]^T \tag{5.4}$$

$$= [0.035, 0., 0., 0.015]^T$$

$$\underline{X}^L = [0.1544, 0.20, 0.11, 0.1661]^T$$

$$\underline{X}^U = [0.1896, 0.20, 0.11, 0.1508]^T$$

The problem random variables r_1 and r_2 are assumed to follow a uniform distribution that covers the upper limits. The exact design scrap percentage S_{de} calculated by integration, through the R^U limits equals

$$S_{de} = 34.87\% \tag{5.5}$$

Instead of choosing both λ_{11} and λ_{12} the latter could be chosen as a percentage of the former as follows:

$$\lambda_{12} = f.\lambda_{11} \frac{\int_{z_1^+}^{z_1^-} f(z_1)\, d\, z_1}{\int_{z_1^{+*}}^{z_1^+} f(z_1)\, d\, z_1} \tag{5.5}$$

where $0 < f \leq 1$

When f equals one, both λ's will be proportional to their corresponding probability of occurrence. However, if f is chosen to be less than one, the intensity of the number of cells that partition the middle sector decreases accordingly. By utilizing Equations (3.11), (3.12), and (5.5) the various number of cells, for each of the three sectors that divides each of the system random variables, could be mathematically expressed as follows:

$$R_{t1}^+ = \text{Int}\, [\lambda_{11}\ \varepsilon_1^{+U}/(x_1^{+L} - x_1^{-L})] + 1$$

$$= \text{Int}\, [0.792\ \lambda_{11}] + 1 \tag{5.6}$$

$$R_{m1} = \text{Int}\, [\lambda_{12}] + 1$$

$$= \text{Int}\, [f\ \lambda_{11}\ (x_1^{+L} - x_1^{-L})/\varepsilon_1^{+U}] + 1$$

$$= \text{Int}\, [1.263\ f\ \lambda_{11}] + 1$$

$$R_{t1}^- = R_{t2}^+ = 0$$

$$R_{m2} = \text{Int}\, [\lambda_{22}] + 1$$

$$= \text{Int}\, [1.233\ \lambda_{12}] + 1$$

$$= \text{Int}\, [1.557\ f\ \lambda_{11}] + 1$$

$$R_{t2}^{-} = \text{Int}\,[\lambda_{23}\ \varepsilon_2^{-U}/(x_2^{+L} - x_2^{-L})] + 1$$

$$= \text{Int}\,[0.4534\ \lambda_{23}] + 1$$

$$= \text{Int}\,[0.320\ \lambda_{11}] + 1$$

For this particular example where all the four corners of the lower region are active, the total number of cells needed to partition the regionalization region equals

$$N_{cell} = R_{m1} \times R_{t2}^{-} + R_{m2} \times R_{t1}^{+} \qquad (5.7)$$

The exponential increase in the total number of cells due to the increase in either the value of λ_{11} or f is shown in Figure 5.3. Figure 5.4 displays the convergence of the estimated design scrap using the regionalization technique toward the exact design scrap as the value of λ_{11} increases. Figure 5.5 magnifies the absolute difference between the exact and the estimated design scrap.

Some important observations can be concluded from the previous two figures. First, depending upon the value of f, there is a transition region that lasts until a critical value of λ_{11}. Throughout this transition region a vast fluctuation in the value of the estimated S_d is experienced. This is due to the cancellation of the over estimated cells with the under estimated ones as illustrated in Figure 5.6 for f equals one and λ_{11} equals two. The critical λ_{11} could be defined as the minimum value of λ_{11} at which the relative error in the estimated value of S_d will be less than the value

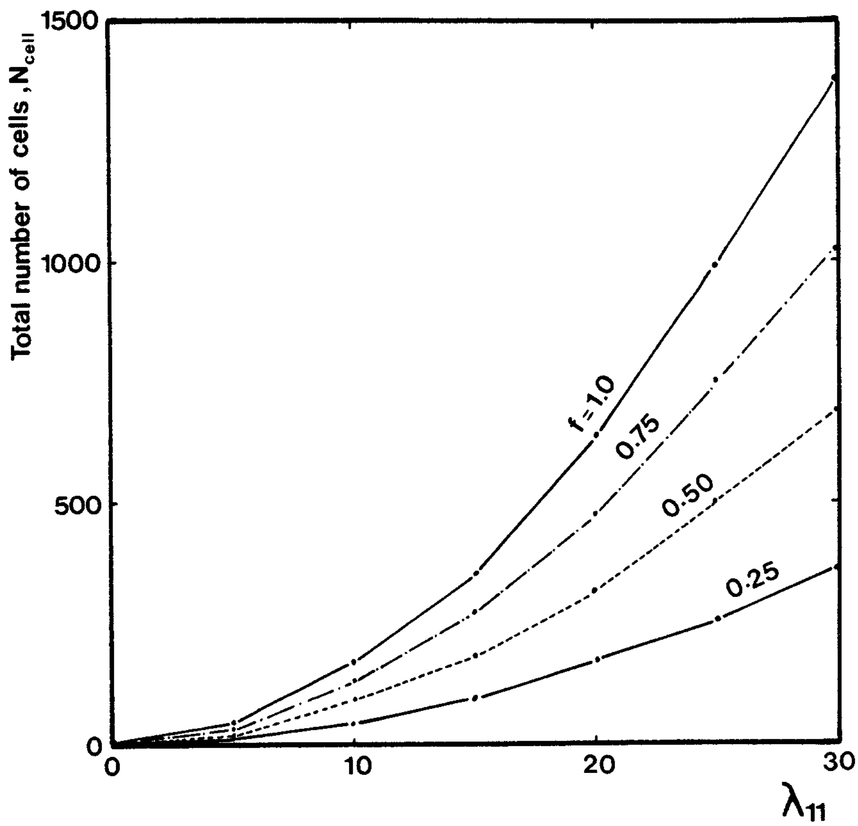


Figure 5.3 The exponential increase in the number of cells.

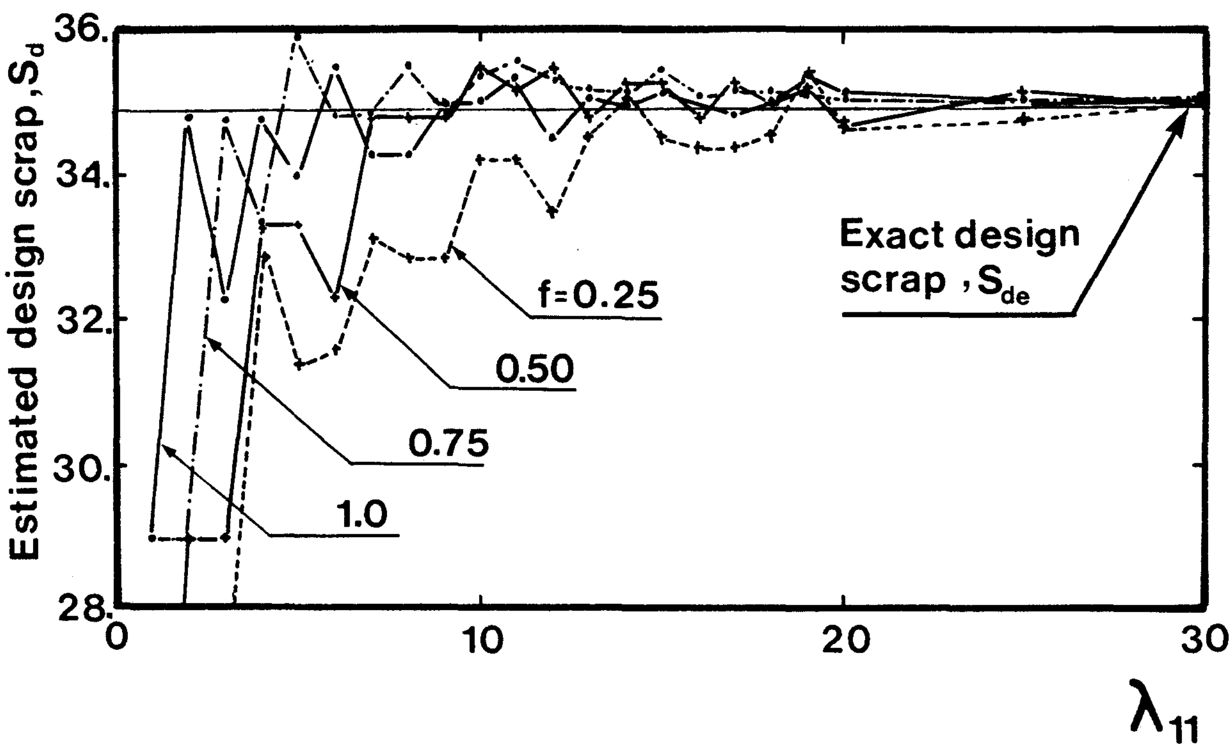


Figure 5.4 Convergence of the estimated design scrap.

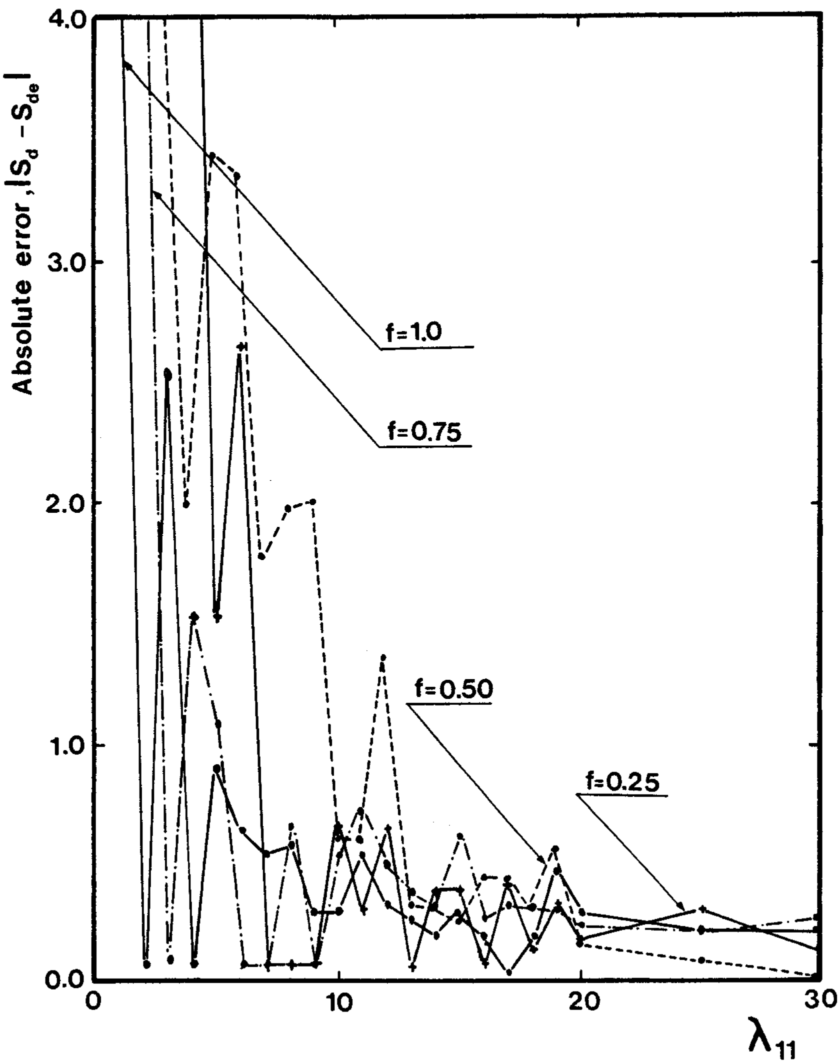


Figure 5.5 The error in the estimated design scrap.

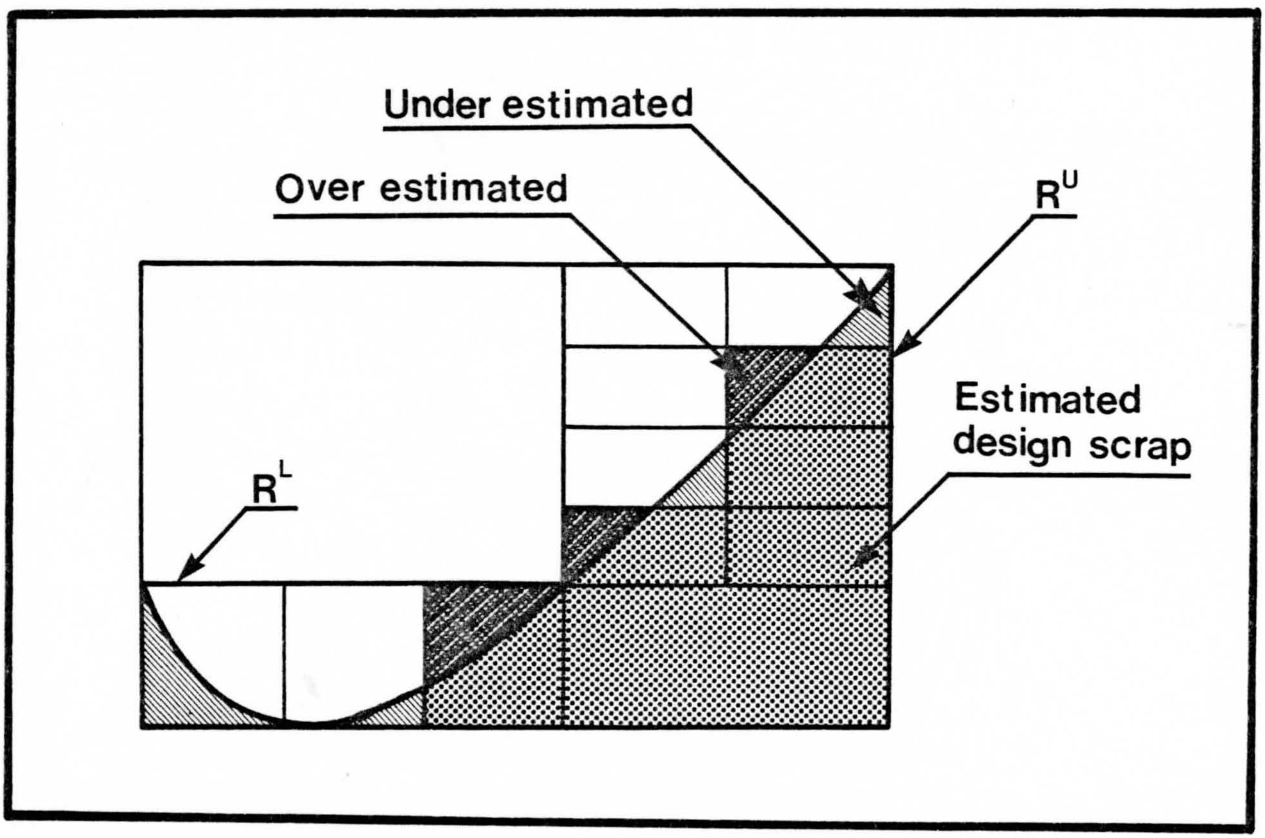


Figure 5.6 A balance between the over estimated and under estimated cells.

of a predetermined maximum acceptable computational relative error. Secondly, the absolute error of the estimated S_d never approaches zero due to the computational accummulated truncated errors. Finally, for the same number of cells, it has been observed that by using smaller f and higher λ_{11}, more accurate estimates ought to be acquired.

5.5 Additional Applications

Techniques and algorithms introduced in Chapters 2, 3 and 4 could be applied to various realistic engineering design systems. Throughout the thesis only the example of the two shrink fitted cylinders has been adapted to illustrate the various concepts and algorithms. This example was chosen, however, because of both its simplicity and its limited number of variables. The following general step by step procedure could be used as a guideline to determine the optimum design variables, and their associated tolerances as well as the optimum scrap percentages of an engineering system.

1. Choose the system design variables. Formulate the system inequality constraints, and determine the system specifications and the probabilistic distributions of the random variables.

2. Construct the system objective function, e.g., cost model. It is normally a function of both the system variables and its scrap.

3. Determine the system optimum nominal variables, $\underline{X}^{o*}$, where the system scrap and tolerances are set to be zeros.

This is used as a starting value to the preceding steps.

4. Determine the system regionalization acceptable lower and upper limit regions. This is utilized in the saving of computational effort.

5. Determine the system optimum nominals and associated tolerances.

The following is a sample of additional systems from various engineering disciplines, that the foregoing tolerance assignment procedure could be applied to.

5.5.1 Mechanical Systems

Bennett [55], solved a car suspension design problem, shown in Figure 5.7, using the classical optimization technique. Both handling and ride performance constraints as well as a reasonable location of the car centre of gravity are considered. The problem design variables are

$$\underline{X}^o = [a,\ K_F,\ K_r,\ K_{fs},\ K_{rs}]^T \qquad (5.7)$$

where, a is the longitudinal location of the car centre of gravity from the front wheels, K_F and K_r are the front and rear suspension stiffness, and K_{fs} and K_{rs} are the front and rear stabilizer bar stiffness, respectively. All of the five design variables should be considered random. And the optimum nominal variables with their associated tolerances could be taken as an input specification in designing each of the sub-systems, i.e., front suspension, rear suspension, front stabilizer and rear stabilizer, separately.

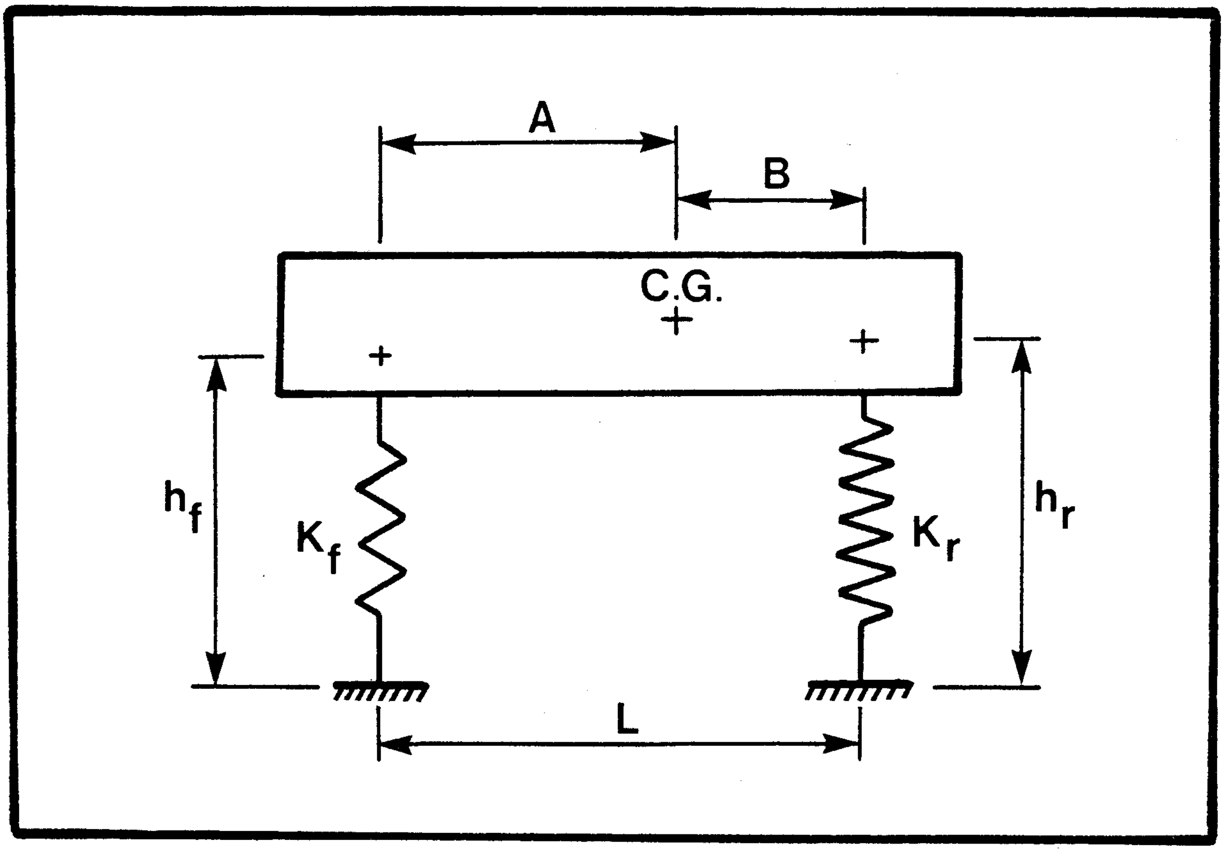


Figure 5.7 Suspension design problem.

Dhande et al. [6] analyzed a four-bar linkage mechanism, shown in Figure 5.8, by allocating equally spaced tolerances and clearance to the four members of the linkage for a specified maximum allowable mechanical error. It is the author's opinion that two main unrealistic assumptions have been adopted by Dhande et al. The link lengths are assumed to be normally distributed and nonlinear constraints have been linearized using Taylor series up to the first order terms. However, by using the general algorithm presented in this thesis there will be no need to adapt the foregoing assumptions and moreover the mechanism expected error could be optimally identified instead of the assumed three sigmas band of confidence level.

5.5.2 Chemical Systems

The Williams-Otto process [57], shown in Figure 5.9, represents a simple chemical plant. The system could be decomposed into six subsystems, each of which could be optimized separately [53]. The system also might be simplified and then optimized in total. The performance characteristics of the various subsystems are either directly specified or mathematically expressed as a function of the system variables. Irrespective of the way of optimizing the system, the tolerance assignment technique could be employed not only to allocate the various tolerances associated with the physical dimensions of the process component but also to determine the permissible

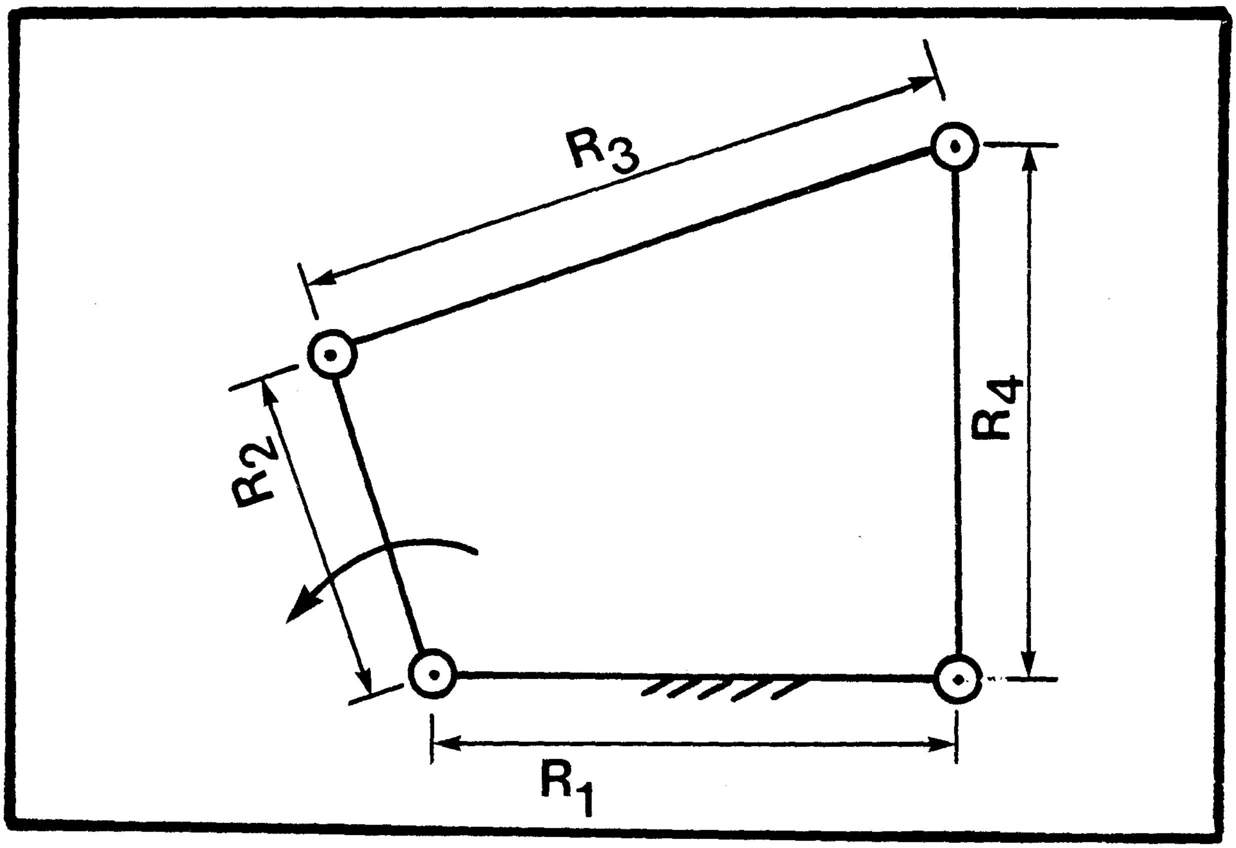


Figure 5.8 Four-bar linkage mechanism.

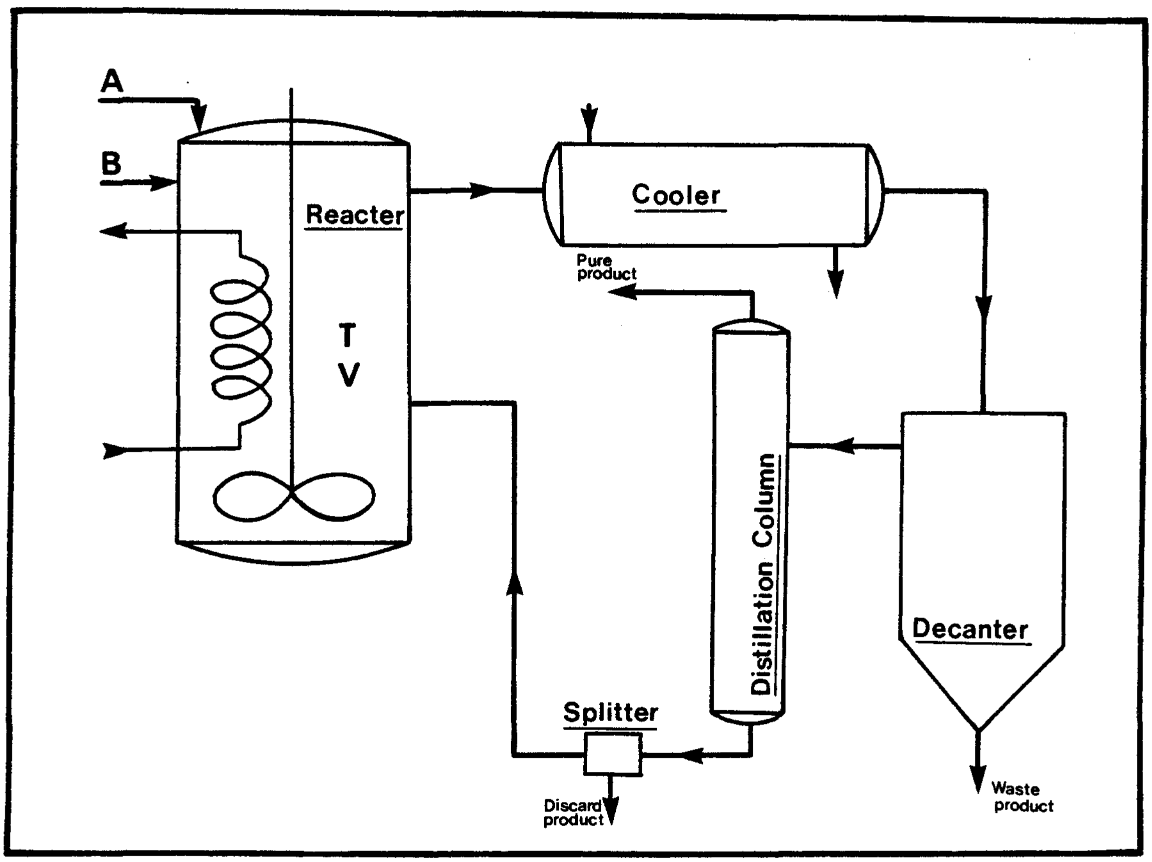


Figure 5.9 The Williams-Otto process.

optimum upper and lower limits of the system environment control variables, e.g., the reactants flow rates, the reaction temperature, the wasted and the discarded product specifications.

5.5.3 Civil Systems

Smith and Hinton [56] described a simple water supply system, shown in Figure 5.10, where they chose the tank storage volume, V, and the water pump horsepower, HP, and water head, h_f, as the system variables. The in-flow discharge, Q, and the pipeline inner diameter, D, were chosen to be the system input specifications. The problem, however, could be tackled from a different prospective using the tolerance assignment algorithm. The required pump head may be considered as a state variable that is a function of the pipe's length, diameter and surface characteristics. The problem objective would be either to minimize the total cost or to maximize the system value, where customers unsatisfaction due to some shortage in the water supply could be directly expressed. This shortage, however, is proportional to the system design scrap percentage. The plant expected life span would either be specified by the designer or treated as an additional random variable.

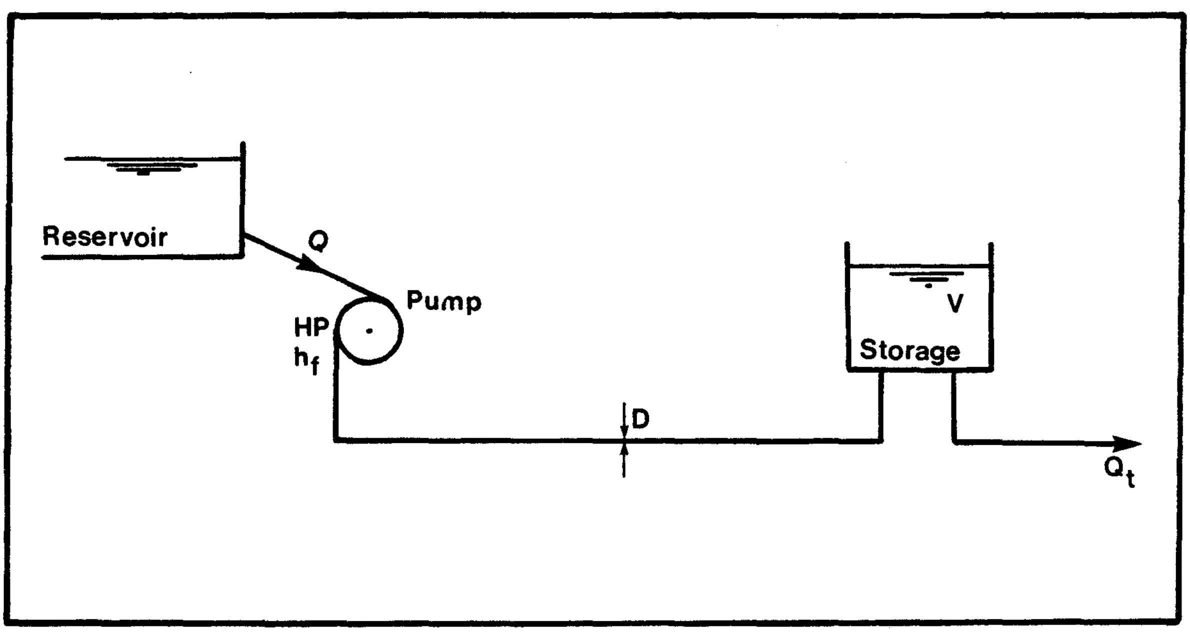


Figure 5.10 Water supply system.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Concluding Summary

The prediction of the behaviour of an engineering system under a variety of conditions is the central issue in applied science work. The basis of such predictions is a model of the relevant phenomenon under study. The more realistic and complete the model, the more accurate is the prediction. A major source of realism in engineering models is the recognition of the random nature of some of their input variables as well as the expected (predicted) outcomes for the design variables. The general design objective, therefore, is to create a system that not only performs the desired function but also represents a solution that is optimal with respect to a design objective function. The system objective may feature components such as: cost, reliability, compatibility with other systems, or in general system value.

In this thesis, the problem of tolerance assignment for both the full and less than full acceptance design conditions has been considered. The design scrap is the percentage of those design outcomes that violate the system constraints. The utilization of the space regionalization technique permits the estimation of the design scrap by calculating weighted infeasible cells.

The distinction between design and manufacturing scrap has been thoroughly defined and elaborated. As far as the author is aware, Chapter 3 provides the only available algorithm which permits allocation of the optimum scrap percentages of a system accurately and without relying on the evaluation of the system partial derivatives. The method is general enough to be applied with any statistical distribution and not necessarily for mechanical systems.

An algorithm to define the acceptable upper regionalization bound has been presented. The algorithm is divided into three strategies to be followed successively. The strategies depend upon one dimensional search and line regionalization techniques, which were designed to suit the tolerance problem. The primary strategy defines a necessary but not sufficient bound that should contain the feasible region of the system. The upper limit strategy with checking of sides, on the other hand, guarantees a sufficient but not necessarily statistically acceptable bound. Finally, the acceptable upper bound strategy adjusts the region limits to be fully contained within the bounds of the system joint probability density function. The procedure is not only efficient but also leads to a considerable computational saving; and provides a quick means of detecting the tolerance limits of a system beyond which all the design outcomes will be totally scrapped.

The inaccuracy in estimating the system scrap due to the violation of the one dimensional convexity assumption has been indicated. For a fixed nominal point, an efficient procedure is suggested for defining the acceptable lower region that bounds within its orthogonal planes the maximum feasible design outcomes irrespective of any non-convexity in the feasible domain.

The precision of the regionalization procedure mainly depends on the total number of cells to be checked against the system feasibility. The regionalization domain for each of the random design variables is divided into three sectors. The number of cells in each sector depends upon its length and relative probability of occurrence. Two scaling factors have to be chosen, depending on the required accuracy in the estimated system scrap.

6.2 Recommendations for Further Research

Promising directions for further research have been revealed by this work. Some of these directions are proposed below:

1. The design outcomes that happen to fall outside the optimum tolerance region are considered scrap. However, some of these designs could be repaired to meet the system performance specifications. Repair usually involves an additional cost that must be included. A portion of the scrap region, therefore, could be defined as a repair region that contributes to decreasing the system scrap percentage.

2. A user-oriented computer program package could be developed to allocate the optimum tolerances associated with system variables. The package should have the capability of

handling general engineering systems in conjunction with a reliable optimization package like OPTIVAR, [50]. It is anticipated that direct search strategies will be most successful.

3. An investigation is desirable to collect the probability distributions of real random variables. Techniques for determining appropriate and precise cost functions are also needed, in order to describe the relationship between the system nominals, their associated tolerances, and the system scrap. Using this information, the algorithms developed in this thesis can be applied for effective statistical optimum design.

4. A method to estimate the critical values of λ_{11} and f would be desirable. The method should guarantee the minimum number of regionalization cells and satisfy the requirement of a maximum permissible error in computing the system scrap percentage.

5. There is a strong correlation between reliability theory and the tolerance assignment problem with less than full acceptance. Further study is needed to explore and define this relation. Time dependent problems could then be used for a possible application.

REFERENCES

1. Siddall, J. N., "Analytical Decision-Making in Engineering Design", Prentice-Hall, Englewood Cliffs, J.N., 1972.

2. Bracken, J. and G. R. McCormick, "Selected Applications of Nonlinear Programming", Wiley, N.Y., 1968.

3. Siddall, J. N., "An Orderly Approach to Design Decisions", ASME paper No. 76-DE-12.

4. Corlew, G. T. and F. Oakland, "Monte Carlo Simulation for Setting Dimensional Tolerance", Machine Design, May 6, 1976, pp. 91-95.

5. Sutherland, G. H. and B. Roth, "Machine Design; Accounting for Manufacturing Tolerances and Costs in Function Generating Problem", J. of Eng. for Ind., Trans. of ASME, Feb. 1975, pp. 283-286.

6. Dhande, S. G. and J. Chakroborty, "Analysis and Synthesis of Mechanical Error in Linkages, A Stochastic Approach", J. of Eng. for Ind., Trans. of ASME, Aug. 73, pp. 672-676.

7. Peters, J., "Tolerancing the Components of an Assembly for Minimum Cost", J. of Eng. for Ind., Trans. of ASME, Series B, Vol. 92, No. 3, Aug. 1970, pp. 677-682.

8. Ligget, J. V., Fundamentals of Position Tolerance, Society of Manufacturing Engineers, Dearborn, Michigan, 1970.

9. Garrett, R. E., and A. S. Hall, "Effect of Tolerance and Clearance in Linkage Design", J. of Eng. for Ind., Trans. of ASME, Series B, Vol. 91, No. 1, Feb. 1969, pp. 198-202.

10. Fortini, E. T., "Dimensioning for Interchangeable Manufacture", Industrial Press Inc., New York, N.Y., 1967.

11. Roth, E. S., "Functional Gaging of Positional Toleranced Parts", Industrail Press Inc., New York, N.Y., 1967.

12. Scarr, A. J. T., "Metrology and Precision Engineering", McGraw-Hill Publishing Co., Ltd., Berkshire, England, 1967.

13. Spotts, M. F., "Mechanical Design Analysis", Prentice-Hall Inc., Englewood Cliffs, N.J., 1961.

14. Wakefield, L. P., "Dimensioning for Interchangeability", Blackie and Son Ltd., Glasgow, 1964.

15. Sandquist, W. L. and N. L. Ensick, "Practical Ways to Apply Statistical Tolerancing", Product Engineering, May 27, 1963, pp. 47-51.

16. Tarello, W. R., "Use Tabled Tolerances", Product Engineering, June 24, 1963, pp. 66-77.

17. Conway, H. G., "Engineering Tolerances", Sir Isaac Pitman and Sons Ltd., 1962.

18. Jaquez, R.,"K- Factors for Computing Tolerance Limits for Normal Distributions", Industrial Quality Control, Vol. 19, Nov. 1962, pp. 27-28.

19. Brooks, K. A., "Statistical Dimensioning Program", Machine Design, Vol. 23, Sept. 14, 1961, pp. 140-145.

20. Spotts, M. F., "Design of Machine Elements", 3rd ed., Prentice-Hall Inc., Englewood Cliffs, N.J., 1961.

21. Grohe, W., "Precision Measurement and Gaging Techniques", Tudor Publishing Co., New York, N.Y., 1960.

22. Tuttle, S. B., "Error Analysis", Machine Design, Vol. 32, No. 12, June 9, 1960, pp. 153-158.

23. Gladman, C. A., "Techniques for Applying Probability to Tolerancing of Machined Dimensions", NSL Tech. Paper No. 11, CSIRO, Melbourne, 1959.

24. Spotts, M. F., "An Application of Statistics to the Dimensioning of Machine Parts", J. of Eng. for Ind., Trans. of ASME, Series B, Vol. 81, Nov. 1959, pp.317-322.

25. Thoen, R. L., "Statistical Tolerancing With Punched Card Computers", Machine Design, June 13, 1957, pp. 121-123.

26. Meltrecht, K. H. and R. M. Caddell, "How to Determine Production Tolerances; Part One - Complete Interchangeability and Part Two - Statistical Methods and Selective Assembly", The Tool Engineer, Oct. 1957, pp. 81-89, and Nov. 1957, pp. 85-89.

27. Buckingham, E., "Dimensions and Tolerances for Mass Production", The Industrial Press, N.Y., 1954.

28. Pike, E. W. and T. R. Silverberg, "Assigning Tolerances for Maximum Economy", Machine Design, Vol. 25, Sept. 53, pp. 139-146.

29. Acton, F. S. and E. G. Olds, "Tolerances Additive or Pythagorean?", Industrial Quality Control, Nov., 48, pp. 6-12.

30. Ostwald, P. F., and J. Huang, "A Method for Optimal Tolerance Selection", J. of Eng. for Ind., Trans. of ASME, Aug. 77, pp. 558-565.

31. Groover, M. P., "Monte Carlo Simulation of the Machining Economics Problem", J. of Eng. for Ind., Trans. of ASME, Aug. 75, pp. 931-938.

32. Spotts, M. F., "Allocation of Tolerances to Minimize Cost of Assembly", J. of Eng. for Ind., Trans. of ASME, Aug. 73, pp. 762-764.

33. Speckhart, F. H., "Calculation of Tolerance Based on a Minimum Cost Approach", J. of Eng. for Ind., Trans. of ASME, May 72, pp. 447-453.

34. Hillier, M. J., "A Systematic Approach to the Cost Optimization of Tolerances in Complex Assemblies", Bull. Mech. Engng. Educ., Vol. 5, pp. 157-161, 1966.

35. Latta, L. W., "Least-cost Tolerancing", Product Engineering, Setp. 16, 1963, pp. 111-113.

36. Abdel-Malek, H. L., " Unified Treatment of Yield Analysis, Worst-Case Design and Yield Optimization", Ph.D. Thesis, McMaster University, Hamilton, Ontario, 1977.

37. Bandler, J. W., P. C. Lin and H. Troup, "A Nonlinear Programming Approach to Optimal Design Centering, Tolerancing and Turning", IEEE Trans. Circuits and Systems, Vol. CAS-23, 1976, pp. 155-165.

38. Director, S. W. and G. D. Hachtel, "The Simplical Approximation Approach to Design Centering and Tolerance Assignment, "IEEE Trans. of Circuits and Systems, Vol. CAS-24, No. 7, July 77, pp. 363-371.

39. Director, S. W., and G. D. Hachtel, "Yield Estimation Using Simplical Approximation", Proc. IEEE Int. Symp. Circuits and Systems, Phoenix, 1977, pp. 579-582.

40. Karafin, B. J., "The General Component Tolerance Assignment Problem in Electrical Networks", Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA, 1974.

41. Scott, T. R. and T. P. Walker, Jr., "Regionalization: A Method for Generating Joint Density Estimates", IEEE Trans. Circuits and Systems, Vol. CAS-23, No. 4, April 1976, pp. 229-234.

42. Gopal, K., "Efficient Statistical Analysis and Design of Systems Using Regionalization", Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, 1978.

43. Elias, N. J., "New Statistical Methods for Assigning Device Tolerances", Proc. IEEE Int. Symp. Circuits and Systems, Newton, MA, 1975, pp. 329-332.

44. Pinel, J. F. and K. Singhal, "Efficient Monte Carlo Computation of Circuit Yield Using Importance Sampling", Proc. IEEE Int. Symp. Circuits and Systems, Phoenix, AZ, 1977, pp. 575-578.

45. Pinel, J. F. and K. A. Roberts, "Tolerance Assignment in Linear Network, Using Nonlinear Programming", IEEE Trans. Circuit Theory, Vol. CT-19, 1972, pp. 475-479.

46. Bandler, J. W., "Optimization of Design Tolerances Using Nonlinear Programming", Proc. 6th Princeton Conf. Informations Sciences and Systems, Princeton, N.J., 1972, pp. 655-659.

47. Bjorke, O., "Computer-Aided Tolerancing", Topic Publishers, 1978.

48. Foupel, J. H., Engineering Design, John Wiley and Sons, 1964.

49. Tromp, H., "The Generalized Tolerance Problem and Worst Case Search", Conf. Computer-Aided Design of Electronic and Microwave Circuits and Systems, Hull, England, 1977, pp.72-77.

50. Siddall, J. N., 'OPTIVAR' - Designers' Optimization Subroutines, Faculty of Engineering, McMaster University, Canada, 1979.

51. Himmelblau, D. M., Process Analysis by Statistical Methods, Wiley, 1970.

52. Peat, A. P., "Cost Reduction Charts for Designers and Production Engineers", The Machinery Publishing Co. Ltd., Brighton, 1968.

53. Siddall, J. N. and W. Michael, "Large System Optimization Using Decomposition With Soft Specification ", ASME paper (78-DAC-21), 1979.

54. Siddall, J. N. and W. Michael, "The Optimization of Large Nonlinear Systems", CANCAM 79, Sherbrooke, Quebec, Canada.

55. Bennett, J. A., "Automated Design of Multi-Objective Systems", ASME paper No. 75-DET-111.

56. Husain, A., and K. Ganguish, "Optimization Techniques for Chemical Engineers", The MacMillam Co. of India Ltd., 1976.

57. Smith, A. A. and E. Hinton, "Civil Engineering Systems Analysis and Design", unpublished manuscript, McMaster University, Civil Eng. Dept., 1978.

APPENDIX A

WORST CONDITION CONSTRAINTS

In the optimum toleranced design problem, $[\underline{X}^*]$, each nominal constraint, $\phi(\underline{X}^o)$, would be replaced by 2^n constraints, where n is the number of the independent design variables which accept tolerances. However, a systematic procedure could be followed to reduce the number of the toleranced constraints and consequently the computational time involved. Tromp [49] has derived a similar formulation.

The problem could be stated as follows: Minimize the toleranced problem constraints, $\underline{\Phi}(\underline{X})$, with respect to the normalized positive and negative random tolerances, $\underline{\alpha}^+$ and $\underline{\alpha}^-$, respectively.

$$
\begin{aligned}
&\underset{\underline{\alpha}}{\text{Minimize}} \ \underline{\Phi}[\underline{X}] \\
&\text{subject to} \quad \underline{\alpha}^+ \in R_\alpha^+ \\
&\qquad\qquad\quad\ \ \underline{\alpha}^- \in R_\alpha^- \\
&\text{where,} \quad \underline{\alpha}=[\underline{\alpha}^+,\underline{\alpha}^-]^T \quad , \quad \underline{X}=\underline{X}^o+\underline{T}^+\underline{\alpha}^++\underline{T}^-\underline{\alpha}^- \\
&\qquad\qquad R_\alpha^+=\{\underline{\alpha}^+ |\ 0<\alpha_i^+<1, \quad i=1,2\ldots n\} \\
&\qquad\qquad R_\alpha^-=\{\underline{\alpha}^- |\ -1<\alpha_i^-<0, \quad i=1,2\ldots n\}
\end{aligned}
\tag{A.1}
$$

The tolerance region, R_t, however, should completely lie inside the feasible region, R_c.

$$R_t \subset R_c \tag{A.2}$$

where,

$$R_c = \{\underline{X} \mid \underline{\Phi}(\underline{X}) \geq \underline{0}\}$$

$$R_t = \{\underline{X}(\underline{\alpha}^+, \underline{\alpha}^-) \mid \underline{\alpha}^+ \in R_\alpha^+, \ \underline{\alpha}^- \in R_\alpha^-\}$$

Problem (A.1) could be discretized if we assume that R_c is convex as follows

$$\text{Minimize} \quad \underline{\Phi}[\underline{X}(\underline{\alpha}^+, \underline{\alpha}^-)]$$

$$\begin{aligned} \text{subject to} \quad & 1-\alpha_i^+ \geq 0 \\ & \alpha_i^+ \geq 0 \\ & 1+\alpha_i^- \geq 0 \\ & -\alpha_i^- \geq 0 \qquad ; i=1,2\ldots n \end{aligned} \tag{A.3}$$

Assuming that a minimum exists at $[\underline{\alpha}^{+*}, \underline{\alpha}^{-*}]^T$ and that $\underline{\Phi}$ is differentiable. The Kuhn-Tucker condition of optimality could be applied on the constraint problem (3.A) and the following six sets of equations would be obtained.

$$\underline{\nabla}_{\underline{\alpha}} \underline{\Phi}[\underline{X}(\underline{\alpha})] = \sum_{i=1}^{n} \lambda_{1i} \underline{\nabla}_{\underline{\alpha}^+} (1-\alpha_i^{+*}) + \sum_{i=1}^{n} \lambda_{2i} \underline{\nabla}_{\underline{\alpha}^+} \alpha_i^{+*} + \sum_{i=1}^{n} \lambda_{3i} \underline{\nabla}_{\underline{\alpha}^-} (1+\alpha_i^{-*}) +$$

$$\Sigma \ \lambda_{4i} \underline{\nabla}_{\underline{\alpha}^-} (-\alpha_i^{-*}) \tag{A.4}$$

$$\lambda_{1i}(1-\alpha_i^{+*})=0 \tag{A.5}$$

$$\lambda_{2i}\ \alpha_i^{+*}\ =0 \tag{A.6}$$

$$\lambda_{3i}(1+\alpha_i^{-*})=0 \tag{A.7}$$

$$-\lambda_{4i}\ \alpha_i^{-*}=0 \tag{A.8}$$

$$\lambda_{1i},\ \lambda_{2i},\ \lambda_{31},\ \lambda_{4i} \geq 0 \tag{A.9}$$

where, $\nabla_{\underline{\alpha}} = [\frac{\partial}{\partial\alpha_1^+}, \frac{\partial}{\partial\alpha_2^+}, \ldots \frac{\partial}{\partial\alpha_n^+}, \frac{\partial}{\partial\alpha_1^-}, \frac{\partial}{\partial\alpha_2^-}, \ldots \frac{\partial}{\partial\alpha_n^-}]^T$

$i=1,2,\ldots n$

$\underline{\lambda}_1$, $\underline{\lambda}_2$, $\underline{\lambda}_3$ and $\underline{\lambda}_4$ are Kuhn-Tucker multipliers.

From Equation (A.4) we get

$$\frac{\partial\phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial\ \alpha_i^+} = -\lambda_{1i}+\lambda_{2i} \tag{A.10}$$

and

$$\frac{\partial\phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial\ \alpha_i^-} = \lambda_{3i}-\lambda_{4i} \tag{A.11}$$

where, $j=1,2,\ldots m$

Let $\underline{X}^*\varepsilon\{\underline{X}(\underline{\alpha}^*)|\alpha_i^{+*}\varepsilon\{0,1\},\ \alpha_i^{-*}\ \varepsilon\{-1,0\},\ i=1,2,\ldots n\}$ (A.12)

Then, there are only two possible values for each α_i^{+*} and α_i^{-*}. They are (0,1) and (-1,0), respectively.

a) $\alpha_i^{+*}=1$

from Equation (A.6)

$$\lambda_{2i}=0$$

and from Equations (A.10) and (A.9)

$$\frac{\partial \phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial \alpha_i^+} = -\lambda_{1i} \leq 0 \tag{A.12}$$

b) $\alpha_i^{+*}=0$

from Equation (A.5)

$$\lambda_{1i}=0$$

and from Equations (A.10) and (A.9)

$$\frac{\partial \phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial \alpha_i^+} = \lambda_{2i} \geq 0 \tag{A.13}$$

c) $\alpha_i^{-*}=-1$

from Equation (A.8)

$$\lambda_{4i}=0$$

and from Equations (A.11) and (A.9)

$$\frac{\partial \phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial \alpha_i^-} = \lambda_{3i} \geq 0 \tag{A.14}$$

d) $\alpha_i^{-*}=0$

from Equation (A.7)

$$\lambda_{3i}=0$$

and from Equations (A.11) and (A.9)

$$\frac{\partial \phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial \alpha_i^-} = -\lambda_{4i} \leq 0 \qquad \text{(A.15)}$$

The preceding four states could be summarized in the following rule.

$$\begin{aligned} \alpha_i^{+*} &= 1 \qquad \text{if } -\text{sgn}\left[\frac{\partial \phi_j[\underline{X}(\alpha^*)]}{\partial \alpha_i^+}\right] > 0 \\ &= 0 \qquad \text{otherwise} \qquad \text{(A.16)} \\ \alpha_i^{-*} &= -1 \qquad \text{if } -\text{sgn}\left[\frac{\partial \phi_j[\underline{X}(\underline{\alpha}^*)]}{\partial \alpha_i^-}\right] < 0 \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

Special cases

i) Linear Constraint:

$$\phi(\underline{X}) = A^T\underline{X}+b \geq 0 \qquad \text{(A.17)}$$

$$= a_1x_1+a_2x_2+\ldots \; a_nx_n+b \geq 0$$

$$= a_1(x_1^o+t_1^+\alpha_1^++t_1^-\alpha_1^-)+a_2(x_2^o+t_2^+\alpha_2^++t_2^-\alpha_2^-)+\ldots$$

$$+a_n(x_n^o+t_n^+x_n^++t_n^-\alpha_n^-)+b \geq 0$$

$$= a_1 t_1^+ \alpha_1^+ + a_2 t_2^+ \alpha_2^+ + \ldots a_n t_n^+ \alpha_n^+ +$$

$$a_1 t_1^- \alpha_1^- + a_2 t_2^- \alpha_2^- + \ldots a_n t_n^- \alpha_n^- + b_m \geq 0$$

where, $b_m = a_1 x_1^o + a_2 x_2^o + \ldots + a_n x_n^o + b$

therefore,

$$\alpha_i^{+*} = 1 \qquad \text{if } a_i \text{ is positive} \tag{A.18}$$

$$\alpha_i^{-*} = -1 \qquad \text{if } a_i \text{ is negative}$$

$$\alpha_i^{+*},\ \alpha_i^{-*} = 0 \qquad \text{otherwise}$$

ii) Symmetrical Tolerance $t_i^+ = t_i^- = t_i$

$$\underline{X} = \underline{X}^o + \underline{T}\ \underline{\beta} \qquad -1 \leq \beta_i \leq 1 \qquad i=1,2\ldots n \tag{A.19}$$

$$\underline{X}^* \varepsilon \{\underline{X}(\underline{\beta}) \mid \beta_i \varepsilon \{-1,1\}; \qquad i=1,2\ldots n\}$$

therefore,

$$\underline{\beta}^* = -\text{sgn}\ \underline{\nabla}_{\underline{\beta}}\ \underline{\Phi}[\underline{\Phi}(\underline{\beta})] \tag{A.20}$$

where $\underline{\nabla}_{\underline{\beta}} = [\frac{\partial}{\partial \beta_1}, \frac{\partial}{\partial \rho_2} \ldots\ldots \frac{\partial}{\partial \beta_n}]^T$

For a linear constraint

$$\beta_i^* = -\text{sgn}\ a_i$$

$$\underline{\beta}^* = -\text{sgn}\ \underline{A} \tag{A.21}$$

Therefore the sign of the worst tolerance in a linear constraint is always the opposite of that of the associated nominal variable.