TOLERANCING:
ITS DISTRIBUTION, ANALYSIS, AND SYNTHESIS

WOO-JONG LEE
TONY C. WOO

Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109-2117

Technical Report No. 86-30

DECEMBER 1986

<Revised on August 30, 1987>

ABSTRACT

Tolerance, representing a permissible variation of a dimension in an engineering drawing, is synthesized by considering assembly stack-up conditions based on manufacturing cost minimization. A random variable and its standard deviation are associated with a dimension and its tolerance. This probabilistic approach makes it possible to perform trade-off between performance and tolerance rather than worst case analysis as it is commonly practiced. Tolerance (stack-up) analysis, as an inner loop in the overall algorithm for tolerance synthesis, is performed by approximating the volume under the multivariate probability density function constrained by nonlinear stack-up conditions with a convex polytope. This approximation makes use of the notion of reliability index [10] in structural safety. Consequently, the probabilistic optimization problem for tolerance synthesis is simplified into a deterministic nonlinear programming problem. An algorithm is then developed and is proven to converge to the global optimum through an investigation of the monotonic relations among tolerance, the reliability index, and cost. Examples from the implementation of the algorithm are given.

1. INTRODUCTION

Dimensions in engineering drawings specify ideal geometry for size, location, and form [1,2]. Since dimensions are subject to variability inherent in the manufacturing process, some variations, such as $\pm$ 0.001, from the nominal value are allowed. The permissible amount, in this example 0.002, is called *tolerance.*

As a design variable, tolerance should be as near zero as possible. But, because of practical considerations such as an increase in cost, tolerance as a manufacturing variable is often larger than ideal. While larger tolerances are less costly to realize, they are usually associated with poor performance. This trade-off between specification and realization illustrates the traditional conflict between design and manufacturing.

As a design-manufacturing variable, tolerance has more than a local effect in the decision process. Parts are "in-spec" if they are functionally equivalent and interchangeable in assembly. Even though individual tolerances are in-spec, the sum of the individual tolerances in an assembly may not be. For example, in Figure 1, suppose the dimension D consists of nominal dimensions A, B, and C with tolerances of $\pm$a, $\pm$b, and $\pm$c, respectively. Now, the variations a, b, and c represent the worst case for the components. Does the entire assembly whose nominal dimension is D need a tolerance of $\pm(a + b + c)$? The study of the aggregate behavior of given individual variations is referred to as *tolerance analysis* or, more commonly, as *stack-up analysis*. In practice, a designer starts with some initial values for tolerances. If the result of the analysis turns out to be "out-of-spec," the designer reassigns some of the tolerances and iterates the analysis procedure. The process of deciding which tolerances are to be changed and by how much, is referred to as *tolerance distribution.* When performed manually, tolerance distribution is often guided by experience. Without a rigorous procedure, it is difficult to ensure that local changes in tolerances reflect global criteria such as functionality and cost. Distributing tolerances such that the result of tolerance analysis is reflected is referred to as *tolerance synthesis*. This paper presents the development of such a procedure.

<Insert Figure 1>

Tolerance synthesis is formulated here as an optimization problem by treating cost minimization as the objective function and the stack-up conditions as the constraints. Probabilistic concepts are used. Since tolerance implies randomness, a random variable and its standard deviation are associated with a dimension and its tolerance. Such a probabilistic approach enables the partial satisfaction of the stack-up conditions. By permitting a small fraction of the assemblies, say 0.3%, to be out-of-spec, an increase in tolerances may be obtained and in turn a reduction in cost may be achieved. This probabilistic approach is considered to be advantageous over the deterministic approach. Since the deterministic approach [3,4,16] handles only the 100% in-spec case, the resulting tolerances are often more conservative than necessary.

In the probabilistic approach, tolerance analysis involves computing the probability of satisfying the stack-up conditions, given the standard deviations (tolerances). Suppose an inequality $F(\mathbf{X}) \geq 0$ represents a certain stack-up condition, where $\mathbf{X}$ is a random vector composed of dimensions. The probability of satisfying this stack-up condition, i.e., $P(F(\mathbf{X}) \geq 0)$, is then described by the following multiple integral:

$$\int_{F(\mathbf{X}) \geq 0} f(\mathbf{X})\, d\mathbf{X} \tag{1}$$

where $f(\mathbf{X})$ is the multivariate probability density function (p.d.f.) for $\mathbf{X}$. $F(\mathbf{X})$, the function for stack-up condition, is nonlinear if non-rectangular shapes and/or angular dimensions are in an engineering drawing. Consider Figure 2−(b). Suppose the vertical distance between points A and B is to be less than 5.2000. The stack-up condition is $F_2(\mathbf{X}) \geq 0$, where

$$F_2(\mathbf{X}) = -\, x_2 \sin x_1 - x_4 \sin (x_1 + x_3) + 5.2000. \tag{2}$$

The linear case [5,11,13,21] offers simplicity in representation and in processing. As another example, suppose the clearance between two components is to be greater than 0.0001. As illustrated in Figure 2−(a), the stack-up condition is $F_1(\mathbf{X}) \geq 0$, where

$$F_1(\mathbf{X}) = x_1 - x_2 - 0.0001. \quad (3)$$

Processing for the linear case is made simple by the following property: under the assumption of independence, the variance of a linear function can be expressed as the linear sum of the variances of the constituting dimensions. Hence, tolerance synthesis becomes the problem of distributing the given sum of variances into the constituting variances. This distribution can be done by using linear programming [5,21] provided that the tolerance-cost relation takes the form of a piecewise convex function. Unfortunately, this approach cannot be used for the nonlinear case, since there is no general rule for expressing the variance of a nonlinear function such as $F_2(\mathbf{X})$ in terms of variances of its parameters x_1, x_2, x_3, and x_4.

<Insert Figure 2>

To compute the multiple integral of (1) without knowledge of the variance of F(**X**), two approaches to tolerance analysis have been considered − simulation and approximation. Monte Carlo simulation [9,21], while powerful, is computationally intensive. As tolerance analysis is an inner loop in a procedure for tolerance synthesis, a faster method is sought. Now, approximation is practical if it also gives a reasonably accurate solution. In this paper, the volume under the multivariate p.d.f. constrained by nonlinear stack-up conditions is approximated by a convex polytope. The distance of each face of the polytope from the nominal dimension point, which is the origin of a transformed coordinate system, is computed through a notion called the "reliability index" introduced by Hasofer and Lind [10] in their civil engineering work. Such an approximation yields a pleasant surprise to the computation as it converts a probabilistic optimization problem (with nonlinear cost function as objective function and nonlinear stack-up conditions as constraints) to a deterministic one (with distances as constraints).

A nonlinear programming (NLP) algorithm for tolerance synthesis is developed. The steps in the algorithm are shown in Figure 3. For a given probability of yield, it computes optimal assignments of tolerances to individual dimensions while minimizing the manufacturing cost. The algorithm is shown to converge to the global optimum through an investigation of the monotonic relations among tolerances, the reliability index, and cost. (Conceptually, monotonicity is easy to understand: as tolerance increases, performance, which is measured by the reliability index, and cost decrease.) This theoretical basis is considered to have an advantage over others [12,16,17,18] whose algorithms do not necessarily guarantee convergence.

<Insert Figure 3>

This paper is organized as follows. Section 2 presents the probabilistic concepts for tolerance analysis and synthesis. Section 3 reviews tolerance analysis in connection with the reliability index and provides the basis for simplification. Section 4 formulates tolerance synthesis as a probabilistic optimization problem and then converts the problem into an NLP problem. Section 5 develops an algorithm of ensuring convergence to global optimum. Section 6 illustrates the algorithm with examples.

2. BASIC CONCEPTS FOR TOLERANCE ANALYSIS AND SYNTHESIS

To capture the randomness of manufacturing processes, a random vector $\mathbf{X}=(x_1,\cdots,x_n)^T$ is used to represent the n dimensions in an engineering drawing. Mean and standard deviation vectors for **X**, denoted by $\overline{\mathbf{X}}=(\overline{x}_1,\cdots,\overline{x}_n)^T$ and $\Sigma=(\sigma_1,\cdots,\sigma_n)^T$, are associated with nominal dimensions and tolerances, respectively. The notation in this paper is given in Table 1.

<Insert Table 1>

Each dimension variable x_i is assumed to follow the normal distribution since errors in the processing of x_i are due to many small independent sources (operator, material, machine, etc). Indeed, Mansoor [15] showed that most manufacturing processes produce dimensions having a normal distribution. The interval between symmetric tolerance limits is then considered as a confidence interval. Symbolically, if the confidence coefficient for a dimension x_i is γ_i, the corresponding confidence interval is $\overline{x}_i - \gamma_i\sigma_i \leq x_i \leq \overline{x}_i + \gamma_i\sigma_i$ and the tolerance is $2\gamma_i\sigma_i$. Tolerance is then determined by standard deviation and confidence coefficient.

In the n dimensional space, the region bounded by the tolerance limits is called a *tolerance region* $\mathbf{R}_T$. It is the region within which dimensions are to be manufactured. The tolerance region for the two dimensions x_1 and x_2 of $F_1(\mathbf{X})$ in (3) is illustrated in Figure 4−(a). Because of the probabilistic nature, this manufacturing objective can be represented by a confidence level. For example, $\pm 3\sigma_i$ indicates a confidence level that 99.73% of the dimensions will be in-spec. The confidence level α_i for each dimension is assumed to be given. Then, the confidence coefficients γ_i's can be derived strictly from the normal distribution table since $\alpha_i = 1 - 2\Phi(-\gamma_i)$. Hence, tolerances can be determined by standard deviations only.

<Insert Figure 4>

Stack-up conditions are assumed to be given and are represented by *requirement functions* $F_j(\mathbf{X})$, $1 \leq j \leq m$, that are inequalities. Each inequality divides the space into a *safe region* $\mathbf{R}_S = \{ \mathbf{X} \mid F_j(\mathbf{X}) \geq 0 \text{ for } j=1,\cdots,m\}$ and a *fail region* $\mathbf{R}_F = \{ \mathbf{X} \mid F_j(\mathbf{X}) < 0 \text{ for } j=1,\cdots,m\}$. These two regions are illustrated in Figure 4−(b), again using $F_1(\mathbf{X})$ as an example. The hypersurface denoted by $F_j(\mathbf{X})=0$ is called the *limit-state surface* and the value of $F_j(\mathbf{X})$, denoted by M_j, is called the *safety margin* of the j-th requirement function.

The intersection of $\mathbf{R}_T$ and $\mathbf{R}_S$ is referred to as the *reliable region* $\mathbf{R}_R$. Symbolically, $\mathbf{R}_R = \{ \mathbf{X} \mid \mathbf{R}_T \cap \mathbf{R}_S\} = \{ \mathbf{X} \mid \{\bigcap_{i=1}^{n} (\bar{x}_i - \gamma_i\sigma_i \leq x_i \leq \bar{x}_i + \gamma_i\sigma_i)\} \cap \{F_1(X) \geq 0, \cdots, F_m(X) \geq 0\}\}$. It is formed by m+2n functions: m of them are from the requirement functions for $\mathbf{R}_S$ and 2n functions are from the upper and lower tolerance limits for the n dimensions. To simplify the explanation on reliable region, hereafter, the functions defined by tolerance limits will also be referred to as requirement functions. That is, the tolerance limit for x_i, i.e., $\bar{x}_i - \gamma_i\sigma_i \leq x_i \leq \bar{x}_i + \gamma_i\sigma_i$, will be treated as two different requirement functions $x_i - \bar{x}_i + \gamma_i\sigma_i$ and $\bar{x}_i + \gamma_i\sigma_i - x_i$. A reliable region is illustrated in Figure 4−(c) for the $\mathbf{R}_T$ and $\mathbf{R}_R$ from Figures 4−(a) and (b). It is noted that the $\mathbf{R}_R$ depends on σ_i since the $\mathbf{R}_T$ varies with σ_i. Figure 5−(a) shows two tolerance regions having the same $P(\mathbf{R}_T)$. These two regions are, however, different in area because of the difference in the standard deviation of the two density functions. But, for a given set of stack-up conditions, $\mathbf{R}_S$ is fixed as illustrated in Figure 5−(b).

An important concept called *yield* is computed as the probability of $\mathbf{X}$ being in $\mathbf{R}_R$, i.e.:

$$P(\mathbf{R}_R) = \int_{\mathbf{X} \in \mathbf{R}_R} \phi(\overline{\mathbf{X}}; \mathbf{V}) d\mathbf{X}. \tag{4}$$

This probability indicates the degree to which the manufactured dimensions satisfy the given stack-up conditions. Note that $\mathbf{R}_T$ and $\mathbf{R}_S$ correspond to the region for the manufactured dimensions and for the desired stack-up conditions, respectively. The yield

$P(\mathbf{R}_R)$ increases by being given smaller tolerances (standard deviations). Figure 5−(c) illustrates this concept. The one on the right offers a higher yield than the one on the left, since it has smaller tolerances. But, smaller tolerances may incur higher manufacturing costs. To resolve the trade-off between yield and cost, due to tolerances, a synthesis procedure is needed. But, the basis for synthesis should be established first.

<Insert Figure 5>

3. SIMPLIFICATION OF TOLERANCE ANALYSIS

Tolerance analysis is to compute the yield $P(\mathbf{R}_R)$ from a set of tolerances (standard deviations) that constitute the multivariate normal p.d.f. $\phi(\overline{\mathbf{X}};\mathbf{V})$ in equation (4). The integration is to be taken in the reliable region $\mathbf{R}_R$ bounded by m+2n functions. As an inner loop in the overall algorithm for tolerance synthesis (shown in Figure 3), tolerance analysis demands speed and accuracy.

For speed, a two-step approximation of $\mathbf{R}_R$, illustrated in Figure 6, is taken: first as a convex polytope and then as an inscribed hypersphere. While the computational advantage of replacing m+2n functions by m+2n hyperplanes and subsequently by a single radius may be obvious, the locations at which linear approximation is to be taken may not be. For accuracy, the consideration of preserving the probabilistically densest area should be taken as illustrated by the column of figures on the right. Now, suppose $\mathbf{R}_R$ has been suitably transformed such that the dimensions z_i are independent and that they follow the standard normal distribution. Refer to Figure 7 and consider two expansion points for linearization, $\mathbf{Z}_1{}^*$ and $\mathbf{Z}_2{}^*$, with distances d_1 and d_2, respectively. Because of normality, the densest area is in the vicinity of the origin. Furthermore, the density decreases exponentially in distance squared, i.e., $(1/2\pi)\left(e^{-d_1{}^2/2} - e^{-d_2{}^2/2}\right)$. It becomes clear then, by choosing the point $\mathbf{Z}^*$ closest to the origin, as illustrated in Figure 7−(b), both speed and accuracy can be achieved.

<Insert Figures 6 and 7>

The reliability index β is defined as the minimum distance from the origin to a limit-state surface formed by a requirement function in an independent standardized coordinate system, called the *standard system.* (The transformation from the dependent vector space $\mathbf{X}$ to the standard system $\mathbf{Z}$ is explained in the Appendix.) The point $\mathbf{Z}^*$ on the limit-state surface with the minimum distance to the origin is referred to as the *design point.* Linearization at the design point is performed by finding the tangent hyperplane in the standard system. The requirement function $G(\mathbf{Z})$ is thus linearized by the tangent hyperplane $L(\mathbf{Z}^*)$ at $\mathbf{Z}^*$ such that $\mathbf{R}_F$ is approximated by $\mathbf{R}_F^*$. The remainder of this section is devoted to the relations between $P(\mathbf{R}_R)$ and β.

Consider the simple case of only one requirement function. Because of the rotational symmetry in the standard system, the probability of covering one side of the tangent hyperplane can be computed from the univariate normal distribution. Hence, the approximated yield $P(\mathbf{R}_R^*)$ only involves looking up the standard normal distribution table.

__Lemma 1__. In the case of a single requirement function, $P(\mathbf{R}_R)$ can be approximated by

$$P(\mathbf{R}_R) \simeq \Phi(\beta). \tag{5}$$

Note that, for a linear requirement function, $P(\mathbf{R}_R) = \Phi(\beta)$. The accuracy of (5) depends on the curvature of the requirement function. As long as the radius of curvature at the design point is large compared to the reliability index, (5) has been shown to be quite accurate in most practical cases [14].

Now, consider the general case of multiple requirement functions. The approximated reliable region $\mathbf{R}_R^*$ after the linearization is always convex and $P(\mathbf{R}_R^*)$ can be obtained by the following lemma:

__Lemma 2__. The yield $P(\mathbf{R}_R)$ is approximated after the linearization by:

$$P(\mathbf{R}_R) \simeq P(\mathbf{R}_R^*) = \int_{-\infty}^{\beta_{m+2n}} \cdots \int_{-\infty}^{\beta_1} \phi(\mathbf{0}; \mathbf{C}_M)\, dz_1 \cdots dz_{m+2n} \tag{6}$$

where the correlation matrix of z_j's, denoted by $\mathbf{C}_M$, is the correlation matrix of the safety margins M_j [6].

Bounds are useful since equation (6) cannot be evaluated for a general $\mathbf{C}_M$ [6,7].

<u>**Lemma 3**</u>. The probability of covering the convex polytope $\mathbf{R}_R^*$ is bounded by:

$$\chi_n^2\left((\min_{j=1}^{m+2n} \beta_j)^2\right) \leq P(\mathbf{R}_R^*) \leq \min_{j=1}^{m+2n}\{\Phi(\beta_j)\} \qquad (7)$$

The lower bound is based on the observation that the hypersphere with a radius of the minimum of β_j always lies inside $\mathbf{R}_R^*$ (and $\mathbf{R}_R$) and the probability covered by this n-dimensional hypersphere is $\chi_n^2\left((\min_{j=1}^{m+2n} \beta_j)^2\right)$. The upper bound of (7) holds since the probability of the intersection is less than or equal to its component probability.

4. FORMULATION OF TOLERANCE SYNTHESIS

The objective of tolerance synthesis is to determine tolerances by minimizing the manufacturing cost C(**T**). The tolerances t_i are constrained to satisfy the stack-up conditions with a certain probability level such that at least a given yield, $1 - \delta$, should be guaranteed. The problem can be then formulated as:

$$\text{Min } C(\mathbf{T})$$

subject to (8)

$$P(\mathbf{R}_R) \geq 1-\delta$$

$$\text{where } t_i \geq 0 \quad \text{for } i=1,\cdots,n.$$

Associating tolerance t_i with standard deviation σ_i, problem (8) becomes the following probabilistic optimization problem, in which all the parameters are described by random variables and their first and second moments:

$$\text{Min } C(\Sigma)$$

subject to (9)

$$\int_{X \in \mathbf{R}_R} \phi(\overline{\mathbf{X}};\mathbf{V})\, d\mathbf{X} \geq 1 - \delta$$

$$\text{where } \sigma_i \geq 0 \quad \text{for } i=1,\cdots,n.$$

Since (9) involves the computation of yield in the constraint, the formulation can be simplified by using the reliability index. This simplification converts (9) into a deterministic optimization problem.

The discussion begins with the simple case of a single requirement function. Based on Lemma 1, the constraint of (9) can be modified such that the reliability index should be greater than β^* where the constant β^* comes from the equation $1-\delta=\Phi(\beta^*)$. Then, the formulation becomes:

$$\text{Min } C(\Sigma)$$

subject to (10)

$$\beta \geq \beta^*$$

$$1-\delta = \Phi(\beta^*).$$

Formulation (10) is next extended to the general case if the values of δ_j for each of the $F_j(\mathbf{X})$, $1 \leq j \leq m+2n$, are given.

Suppose the δ_j's are same, e.g., $\delta_j = \delta$, for $1 \leq j \leq m+2n$. This implies that each stack-up condition is satisfied with at least $1-\delta$ level. Then, the resultant yield $P(\mathbf{R}_R)$ is upper bounded by $1-\delta$ due to Lemma 3. This approach, referred to as MULTI-1, is formulated as follow:

$$\text{Min } C(\Sigma)$$

subject to (11)

$$\beta_j \geq \beta^* \qquad \text{for } j=1,\cdots,m$$

$$1-\delta = \Phi(\beta^*).$$

Notice that the constraints for the tolerance limits are omitted since they have already been considered when setting the confidence coefficients. That is, for $m+1 \leq j \leq m+2n$, $\beta_j = \beta^* = \gamma_j$. MULTI-1 gives loose tolerances since the desired yield is reflected by its upper limit.

Another approach MULTI-2 is designed here to reflect the desired yield with the lower limit so that tighter tolerances are produced. As indicated by Lemma 3, the lower limit of $P(\mathbf{R}_R)$ can be obtained through the largest hypersphere inscribed in the approximated reliable region $\mathbf{R}_R^*$, which is a convex polytope. The center of such a hypersphere must be coincident with the origin in the standard system. This approach can be summarized as follows:

$$\text{Min } C(\Sigma)$$

subject to (12)

$$\beta_j \geq \beta^* \quad \text{for } j=1,\cdots,m$$

$$1-\delta=\chi_n^2(\beta^{*2}).$$

As in (11), the constraints for the tolerance limits are omitted since $\beta_j = \beta^* = \gamma_j$ for $m+1 \leq j \leq m+2n$. MULTI-2 produces tolerances that ensure at least $1-\delta$ yield.

From these two approaches, the limits of tolerances for the convex polytope representing the yield $1-\delta$ can be obtained as follows:

$$\sigma_{i,2} \leq \sigma_i \leq \sigma_{i,1} \qquad \text{for } 1 \leq i \leq n \tag{13}$$

where $\sigma_{i,1}$ and $\sigma_{i,2}$ are the i-th standard deviations from MULTI-1 and MULTI-2, respectively.

To aid the decision of choosing σ_i, a simple variation of MULTI-1, referred to as MULTI-1.5, is also developed by setting β^* based on the equations $\Phi(\beta^*)=(1-\delta)^{1/m}$ MULTI-1.5 has a special meaning if the requirement functions are nonnegatively correlated, i.e., if every element of $\mathbf{C}_M$ is nonnegative. Under this condition, $P(\mathbf{R}_R^*)$ of equation (6) is greater than $\prod_{j=1}^{m+2n} \Phi(\beta_j)$, which is the approximated yield when the requirement functions are independent, i.e., when $\mathbf{C}_M$ is an identity matrix [7]. In other words,

$$\prod_{j=1}^{m+2n} \Phi(\beta_j) \leq P(\mathbf{R}_R^*) \leq \min_{j=1}^{m+2n} \{\Phi(\beta_j)\}. \tag{14}$$

By distributing δ equally into the m stack-up conditions through $(1-\delta)^{1/m}$, the lower and upper bounds of (14) become $(1-\delta)\times P(\mathbf{R}_T)$ and $(1-\delta)^{1/m}$, respectively.

5. ALGORITHM

An algorithm for the nonlinear programming problems (11) and (12) is developed. To show convergence, emphasis is given to the monotonicity among the reliability index, cost, and tolerances (standard deviations). A function $g(\mathbf{X})$ is said to be monotone nonincreasing in every elements x_i if $\mathbf{X}_1 \geq \mathbf{X}_2$ results in $g(\mathbf{X}_1) \leq g(\mathbf{X}_2)$.

5-1. Monotonicity of Cost in Tolerance

Monotonicity is generally understood in practice: the more tolerance the less cost. Indeed, most tolerance-cost models [19,20,22] describe this relation with inverse or exponential functions. A common characteristic of these models is that the first derivative of the cost function $\partial C_i(\sigma_i)/\partial \sigma_i$ is always negative and monotone increasing with respect to tolerance (standard deviation). These properties imply that $C_i(\sigma_i)$ is strictly decreasing convex. This paper uses the following additive cost function to represent the total manufacturing cost:

$$C(\Sigma) = \sum_{i=1}^{n} C_i(\sigma_i) \tag{15}$$

Then, the following lemma holds since the sum of convex functions is also convex:

Lemma 4. $C(\Sigma)$ is a convex function.

5-2. Monotonicity of Reliability Index in Tolerance

Observe the effect of a change in tolerance on the change of distance $d_{j,0}$ of a point $\mathbf{X}_0$ on the limit-state surface $F_j(\mathbf{X})=0$ to the origin of the standard system.

Theorem 1. If the dimensions are independent,

$$\frac{\partial d_{j,0}}{\partial \sigma_i} \leq 0 \qquad \text{for } 1 \leq i \leq n \text{ and } 1 \leq j \leq m.$$

Proof. The distance $d_{j,0}$ can be expressed by $\{(\mathbf{X}_0 - \overline{\mathbf{X}})^T \mathbf{V}^{-1} (\mathbf{X}_0 - \overline{\mathbf{X}})\}^{1/2}$ (For details, refer to the Appendix). This expression corresponds to the equation of an ellipsoid with center $\overline{\mathbf{X}}$, the semi-axes of which are expressed as the product of $d_{j,0}$ and the standard deviations. Then,

$$d_{j,0}^{\;2} = \sum_{k=1}^{n} \frac{(x_{0k} - \bar{x}_k)^2}{\sigma_k^2} \tag{16}$$

As a standard deviation σ_k decreases, the distance $d_{j,0}$ from a point chosen arbitrarily on $F_j(\mathbf{X})$ to the origin increases. ●

From this theorem, it follows immediately that the minimum distance, i.e., the reliability index, also increases if a standard deviation decreases.

Lemma 5. If the dimensions are independent,

$$\frac{\partial \beta_j}{\partial \sigma_i} \leq 0 \qquad \text{for } 1 \leq i \leq n \text{ and } 1 \leq j \leq m.$$

This lemma not only shows the nonincreasing monotonicity of the reliability index, but also reconfirms the trade-off between tolerance and performance: the tolerance represented by standard deviation has an inverse relationship with performance implied by the reliability index.

5-3. Convergence to Global Optimum

From Lemma 5, the functional form of β_i can be established.

Lemma 6. β_j, which is a function of Σ, is a quasi-concave function provided that the dimensions are independent.

Proof. A function is quasi-concave if the functional value of the convex combination of any two arbitrarily chosen points is greater than or equal to the lesser of the functional values of the two chosen points. From Lemma 5, β_j is quasi-concave. •

Lemmas 4 and 6 provide the basis for the development of the algorithm. With the additional aid of Lemma 7, the local optimum satisfying the Kuhn-Tucker conditions [23] becomes the global optimum.

Lemma 7. In the NLP problem with $C(\Sigma)$ and β_j, for $1 \le j \le m$, differentiable to σ_i for $1 \le i \le n$, let the minimization-based objective function $C(\Sigma)$ be convex and the constraints β_j be quasi-concave. Suppose Σ^* satisfies the Kuhn-Tucker conditions. Then Σ^* is optimal for the NLP problem.

Proof. The proof can be found in pp. 43–44 of [23].

5-4. Development of the Algorithm

Algorithm TOL-M is developed for problems (11) and (12). The algorithm starts with fairly large tolerances. This initial assignment incurs a low cost but it may not satisfy the stack-up conditions. To satisfy the conditions, some or all of tolerances need to be decreased. TOL-M provides a tolerance reduction procedure to the level that the resulting reliability index is equal to β^*. For clarity, it is described in pseudo-PASCAL.

Algorithm TOL-M

```
    {Input : Requirement functions F_j(X),   j=1,···,m
    Output: Optimum standard deviation vector Σ*
    Note : β* comes from the equation 1−δ=Φ(β*) for (11) or 1−δ=χ_n^2(β*^2) for (12).}
Begin
    StopFlag:= false;
    Σ*:= fairly large standard deviations;
    Repeat
        For j:= 1 to m do
            Calculate β_j under the current Σ*;
            {Refer to the iterative scheme in the Appendix for computing β_j}
        β_min:= min(β_1,···,β_m);
        If |β_min − β*| ≤ ε then StopFlag:= true
        Else
            If β_min > β* then
                Obtain new Σ* by increasing some standard deviations
                {Under the current solution, the system is over-reliable}
            Else
                Obtain new Σ* by decreasing some standard deviations;
                {Under the current solution, the system is under-reliable}
    Until StopFlag = true;
End;
```

The increase or decrease of standard deviations at every iteration is done by the following three steps:

(i) Check each variable in the given order to see if it can be changed to make the constraints more feasible.

(ii) If so, modify the variable. The amount is determined by bisection because of the monotonicity of the reliability index.

(iii) Repeat steps (i) and (ii) from the next variable.

This local search procedure (i)~(iii) guarantees the convergence to the global optimum due to Lemma 7.

Theorem 2. Algorithm TOL-M converges to the global optimum of problems (11) and (12), under the assumption of independence of dimensions.

6. EXAMPLES

The algorithm has been implemented in PASCAL and it runs on the University of Michigan IBM 3090–400/VM under the MTS operating system. Two examples, in which the requirement functions are linear and nonlinear, are examined.

Example 1. Linear Stack-Up Conditions

The linear requirement functions for this example are listed in Figure 8. The first requirement function $F_1(\mathbf{X})$ is for the size condition of the base part, that is, the length x_4+x_5 is to be less than or equal to 5.005 . The other three functions reflect the clearance conditions between the two parts.

<Insert Figure 8>

The nominal dimensions are given as $\overline{\mathbf{X}}^T$ = (1.0, 2.0, 3.0, 4.0, 1.0, 0.998, 2.0, 2.998). And, the additive cost function of (15) is used for total manufacturing cost where each tolerance-cost function $C_i(\sigma_i)$ is assumed to follow

$$C_i(\sigma_i) = \frac{a_i \times 10^{-3}}{(6\sigma_i)^{b_i}} \qquad (17)$$

The coefficients in (17) are set as: $a_1=a_2=1.0$, $a_3=a_4=1.5$, $a_5=0.8$, $a_6=0.9$, $a_7=0.8$, and $a_8=0.6$; and $b_1=2.0$, $b_2=1.8$, $b_3=1.7$ $b_4=2.0$, $b_5=3.0$, $b_6=2.0$, and $b_7=b_8=1.9$. With these data, three approaches (MULTI-1, MULTI-1.5, and MULTI-2) are tested for a 95% yield, i.e., $1-\delta=0.95$.

The initial variances are assigned in accordance with the loose fit case of ANSI-Y14.5M [2] for hole and shaft tolerance determination. The result of executing TOL-M is given in Table 2. MULTI-1 generates loose tolerances as prescribed by the upper bound in Lemma 3. For a total cost of 946.83, each stack-up condition would be satisfied with a

95% confidence if the dimensions x_1 to x_8 are manufactured within the tolerances obtained by MULTI-1. MULTI-1.5 gives the result when the 95% yield is equally distributed into the given four stack-up conditions by $(0.95)^{1/4}=0.98726$. With the tolerances from MULTI-1.5, each stack-up condition can be satisfied with 98.726% confidence. MULTI-2 generates tight tolerances by inserting a hypersphere covering 95% into the reliable region, for a total cost 6383.17. The four stack-up conditions are satisfied simultaneously with at least the 95% confidence. The resulting yield from the tolerances of MULTI-1 and MULTI-1.5 are upper bounded by 95% and 98.726%, respectively. MULTI-2 ensures the given 95% yield while MULTI-1 and MULTI-1.5 may not.

<Insert Table 2>

Example 2. General Stack-Up Conditions

Six requirement functions for Example 2 are shown in Figure 9. The first and second functions, $F_1(\mathbf{X})$ and $F_2(\mathbf{X})$, are for the vertical and the horizontal clearance conditions of the two parts. And, the third and the fourth functions show that the difference between angles θ_1 and θ_2 should be within $\pm\pi/180$ radians for successful assembly. The last two functions show that the size difference of the two parts should be within 0.01.

<Insert Figure 9>

The nominal dimensions are given as $\overline{\mathbf{X}}^T$ = (50.0, 40.00125, 20.05, 9.9985, 9.9985, 30.0, 10.0, 30.0, 10.05, 30.0, 40.0, 50.0). Each tolerance has the cost function of (17). The coefficients in (17) are set as: $a_1=0.2$, $a_2=1.0$, $a_3=a_4=0.015$, $a_5=0.008$, $a_6=0.009$, $a_7=0.008$, $a_8=0.006$, $a_9=1.0$, $a_{10}=0.01$, $a_{11}=0.015$, and $a_{12}=0.2$; and $b_1=\cdots=b_{12}=2.0$. With these data, three approaches (MULTI-1, MULTI-1.5, and MULTI-2) are tested for a 95% yield. The result of executing TOL-M is given in Table 3.

<Insert Table 3>

7. SUMMARY

This paper presents a unified procedure for tolerance synthesis by distributing tolerances so as to satisfy the stack-up conditions. As a global criterion, cost minimization is used.

Probabilistic concepts for tolerance analysis and synthesis are introduced. In terms of dimensions and tolerances, areas of interest to manufacturing and to design are defined as tolerance region and safe region, respectively. The intersection of the two regions, i.e., the reliable region $\mathbf{R}_R$, is investigated in detail.

Tolerance analysis for computing $P(\mathbf{R}_R)$ is expedited through an approximation of $\mathbf{R}_R$ with a convex polytope $\mathbf{R}_R^*$. Bounds of $P(\mathbf{R}_R^*)$ is examined by using the reliability index β. For the upper and lower bounds, the probabilistic optimization problem for tolerance synthesis is converted into two NLPs. Then, effort is devoted to developing an algorithm, which is an iterative method of ensuring convergence.

This iterative method demonstrates the potential for automatic tolerance synthesis, especially for the general nonlinear case. The concepts in this paper contributes to the understanding of parameters in design, manufacturing, and assembly by investigating:

(a) the relation between tolerance and nominal dimension,

(b) the relation between tolerance and desired yield, and

(c) the relation between tolerance and stack-up condition.

REFERENCES

1 *ISO System of Limits and Fits, General Tolerances and Deviations*, R286–1962, American National Standards Institute, New York, 1962.

2 *Dimensioning and Tolerancing*, ANSI Standard Y14.5M-1982, American Society of Mechanical Engineering, New York, 1982.

3 Balling, R.J., Free, J.C., and Parkinson, A.R., "Consideration of Worst-Case Manufacturing Tolerances in Design Optimization," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 108, Dec. 1986, pp. 438–441.

4 Bandler, J., "Optimization of Design Tolerances Using Nonlinear Programming," *Journal of Optimization Theory and Applications*, Vol. 14, No. 1, July 1974, pp. 99–114.

5 Bjorke, O., *Computer Aided Tolerancing*, Tapir, Norway, 1978.

6 Ditlevsen, O.D., "Narrow Reliability Bounds for Structural Systems," *ASCE Journal of Structural Mechanics*, Vol. 7, No. 4, Dec. 1979, pp. 453–472.

7 Ditlevsen, O.D. and Bjerager, "Methods of Structural System Reliability," *Structural Safety*, Vol. 4, 1986, pp. 195–229.

8 Graybill, F.A., *An Introduction to Linear Statistical Model -Volume 1*, McGraw-Hill Book Co., Inc., New York, 1961.

9 Grossman, D.D., "Monte-Carlo Simulation of Tolerancing in Discrete Parts Manufacturing and Assembly," Report No. STAN-CS-76-555, Computer Science Department, Stanford Univ., May 1976

10 Hasofer, A.M. and Lind, N.C., "Exact and Invariant Second-Moment Code Format," *ASCE J.Eng.Mech.Div.*, Vol. 100, EM1, 1974, pp. 111–121.

11 Huang, J. and Ostwald, P.F., "A Method for Optimal Tolerance Selection," ASME Paper No. 76-WA/DE-23.

12 Johnson, R.H., "How to Evaluate Assembly Tolerances," *Product Engineering*, Jan. 1953, pp. 179–181.

13 Kirkpatrick, E.G., *Quality Control for Managers and Engineers*, John Wiley and Sons, Inc., New York, 1970.

14 Madsen, H.O., Krenk, S., and Lind, N.C., *Methods of Structural Safety*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1986.

15 Mansoor, E.M., "The Application of Probability to Tolerances Used in Engineering Designs," *Proc.Inst.Mech.Eng.*, Vol. 178, No. 1, 1963, pp. 29–51.

16 Michael, W. and Siddall, J.N., "The Optimization Problem with Optimal Tolerance Assignment and Full Acceptance," *ASME Journal of Mechanical Design*, Vol. 103, Oct. 1981, pp. 842–848.

17 Michael, W. and Siddall, J.N., "The Optimal Tolerance Assignment with Less Than

Full Acceptance," *ASME Journal of Mechanical Design*, Vol. 104, Oct. 1982, pp. 855-860.

18 Parkinson, D.B., "The Application of Reliability Methods to Tolerancing" *ASME Journal of Mechanical Design*, Vol. 104, July 1982, pp. 612-618.

19 Peat, A.P., *Cost Reduction Charts for Designers and Production Engineers*, The Machining Publishing Co. Ltd., Brighton, 1968.

20 Spotts, M.F., "Allocation of Tolerances to Minimize Cost of Assembly," *ASME Journal of Engineering for Industry*, Aug. 1973, pp. 762-764.

21 Turner, J.U., "Tolerances in Computer-Aided Geometric Design," Ph.D. Dissertation, Rensselaer Polytechnic Institute, Troy, New York, 1987.

22 Wilde, D. and Prentice, E., "Minimum Exponential Cost Allocation of Sure-Fit Tolerances," ASME Paper No. 75-DET-93.

23 Zangwill, W.I., *Nonlinear Programming – A Unified Approach*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1969.

APPENDIX. RELIABILITY INDEX

The solution scheme for the reliability index is composed of the following two steps: first, transform the correlated variable space into the standard system; second, find a point which is of the minimum distance from origin to the limit-state surface of a requirement function.

The transformation of the correlated vector **X** is accomplished by taking the following three steps [8]: first, translation $\mathbf{X}^o = \mathbf{X} - \overline{\mathbf{X}}$; second, orthogonal transformation $\mathbf{Z}^o = \mathbf{P}^T\mathbf{X}^o$, where **P** is the orthogonal matrix to diagonalize **V** through $\mathbf{P}^T\mathbf{V}\mathbf{P} = \mathbf{V}_z$; third, standardization $\mathbf{Z} = \mathbf{D}^{-1}\mathbf{Z}^o$ where $\mathbf{D}\mathbf{D}^T = \mathbf{V}_z$. Figure 10 shows the changes of a given function before and after the transformation. Then, the requirement function in the standard system is expressed symbolically by G(**Z**) as:

$$G(\mathbf{Z}) = F(\mathbf{D}^{-1}\mathbf{P}^T(\mathbf{X}-\overline{\mathbf{X}}))$$

<Insert Figure 10>

The next step is to find a point on the limit-state surface G(**Z**) having the minimum distance from the origin. Since the Euclidean distance from the origin to a point **Z** is expressed as $(\mathbf{Z}^T\mathbf{Z})^{1/2} = ((\mathbf{X}-\overline{\mathbf{X}})^T\mathbf{V}^{-1}(\mathbf{X}-\overline{\mathbf{X}}))^{1/2}$, the following NLP can be used for finding β:

$$\text{Min } \beta^2 = (\mathbf{X}-\overline{\mathbf{X}})^T\mathbf{V}^{-1}(\mathbf{X}-\overline{\mathbf{X}})$$

subject to (18)

$$F(\mathbf{X}) \leq 0.$$

Here, the objective function is expressed as β^2 instead of β since the positive definiteness of the covariance matrix **V** always guarantees the same solution.

As a solution scheme for (18), the iterative method based on the Newton-Raphson method [7,14] is used:

$$\mathbf{X}^{(k+1)} = \overline{\mathbf{X}} + \mathbf{V}\ \nabla F(\mathbf{X}^{(k)})\ \frac{(\mathbf{X}^{(k)} - \overline{\mathbf{X}})^T\ \nabla F(\mathbf{X}^{(k)}) - F(\mathbf{X}^{(k)})}{\nabla F(\mathbf{X}^{(k)})^T\ \mathbf{V}\ \nabla F(\mathbf{X}^{(k)})}$$

where $\mathbf{X}^{(k)}$ denotes the solution after the k-th iteration and $\nabla F(\mathbf{X})$ is the $n \times 1$ gradient vector of $F(\mathbf{X})$ at $\mathbf{X}$.

Table 1. Notation

Notation	Description
$C(\cdot)$	manufacturing cost function
$F_j(\mathbf{X})$	the j-th requirement function where $F_j(\mathbf{X}) \geq 0$ is the desired condition
m	number of given requirement functions, $1 \leq j \leq m$
M_j	safety margin of $F_j(\mathbf{X})$
n	number of tolerances to be decided, $1 \leq i \leq n$
$P(\cdot)$	probability
$\mathbf{R}_R$	reliable region
$\mathbf{R}_S$, $\mathbf{R}_F$	safe region, fail region
$\mathbf{R}_T$	tolerance region
$\mathbf{R}^*$	approximated region
$\mathbf{T}$	tolerance vector
$\mathbf{V}$	$n \times n$ covariance matrix
$\mathbf{X}$	random vector for dimensions
$\overline{\mathbf{X}}$	mean vector of $\mathbf{X}$
$\mathbf{Z}$	dimension vector transformed into the standard system
β_j	reliability index for $F_j(\mathbf{X})$
β^*	reliability index derived from the desired yield
γ_i	given confidence coefficient for x_i
δ	permissible dissatisfaction of assembly ,i.e., $1-\delta$ is the desired yield
δ_j	permissible dissatisfaction of $F_j(\mathbf{X})$
Σ	standard deviation vector of $\mathbf{X}$
$\Phi(\cdot)$	cumulative standardized normal distribution function
$\phi(\overline{\mathbf{X}};\mathbf{V})$	multivariate normal p.d.f. having mean $\overline{\mathbf{X}}$ and covariance $\mathbf{V}$
$\chi_n^2(\cdot)$	cumulative chi-squared distribution with n degree of freedom

Table 2. Result of Example 1

Dimension x_i	Tolerance t_i (*)		
	MULTI-1	MULTI-1.5	MULTI-2
x_1	0.00446	0.00328	0.00186
x_2	0.00168	0.00124	0.00070
x_3	0.00238	0.00176	0.00100
x_4	0.00547	0.00403	0.00229
x_5	0.01740	0.01281	0.00727
x_6	0.00168	0.00123	0.00070
x_7	0.00142	0.00104	0.00059
x_8	0.00371	0.00273	0.00155
Normalized Cost (Actual Cost)	1.00 (946.83)	1.92 (1,816.38)	6.74 (6,383.17)
Run Time in CPU seconds	0.137	0.141	0.151

(*) set at $6\sigma_i$

Table 3. Result of Example 2

Dimension x_i	Tolerance t_i (*)		
	MULTI-1	MULTI-1.5	MULTI-2
x_1	0.0265	0.0187	0.0093
x_2	1.7907	1.2331	0.6411
x_3	0.0840	0.0579	0.0301
x_4	0.1023	0.0705	0.0367
x_5	0.0280	0.0019	0.0010
x_6	0.0032	0.0022	0.0011
x_7	0.0028	0.0019	0.0010
x_8	0.0021	0.0015	0.0008
x_9	1.7947	1.2385	0.6451
x_{10}	0.1035	0.0714	0.0371
x_{11}	0.0840	0.0579	0.0301
x_{12}	0.0250	0.0168	0.0092
Normalized Cost (Actual Cost)	1.00 (4.93)	2.11 (10.38)	7.77 (38.30)
Run Time in CPU seconds	0.416	0.418	0.417

(*) set at $6\sigma_i$

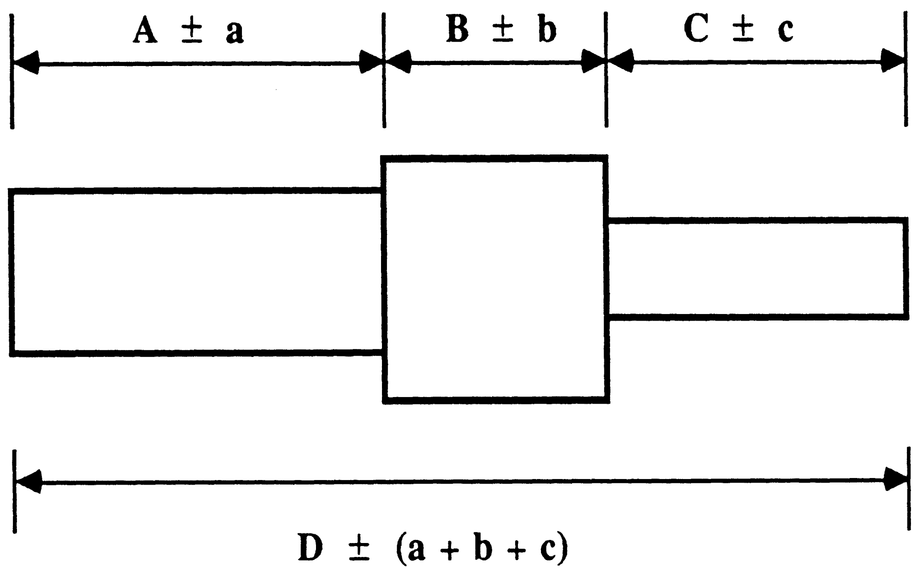
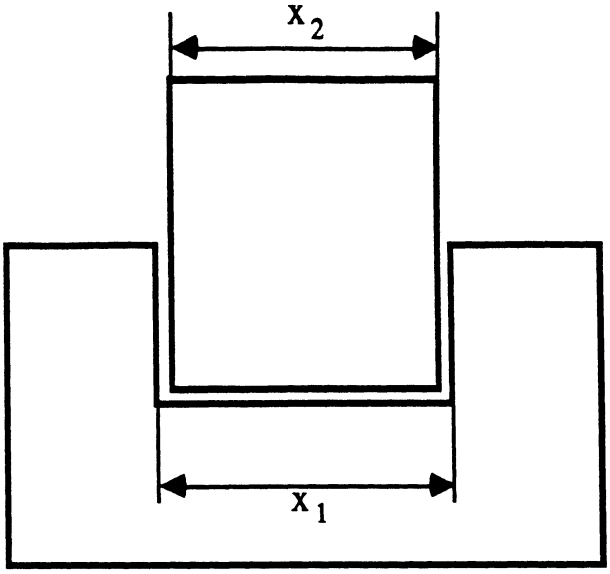


Figure 1. Tolerance Stack-Up

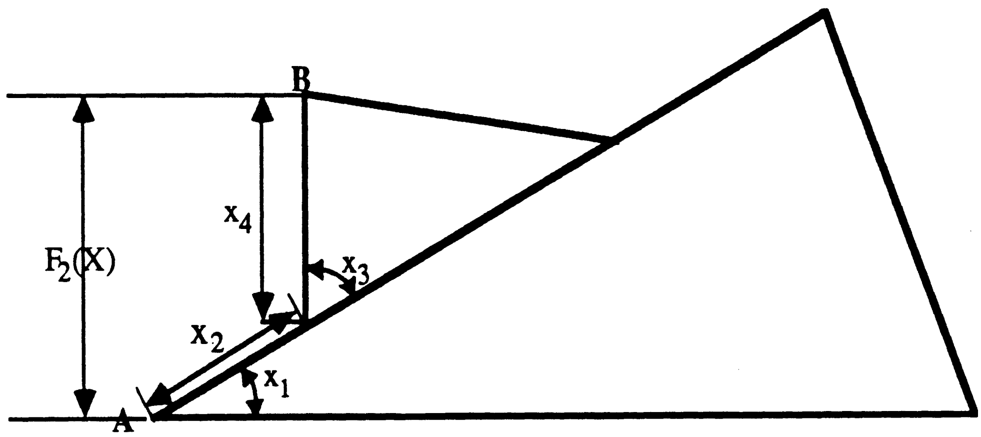
(a)



clearance between x_1 and x_2 $\geq$ 0.0001

$F_1(X) \geq 0$ where $F_1(X) = x_1 - x_2 - 0.0001$

(b)



vertical height from A to B $\leq$ 5.2000

$F_2(X) \geq 0$ where

$$F_2(X) = - x_2 \sin x_1 - x_4 \sin(x_1 + x_3) + 5.2000$$

Figure 2. Examples of Linear and Nonlinear Stack-Up Conditions

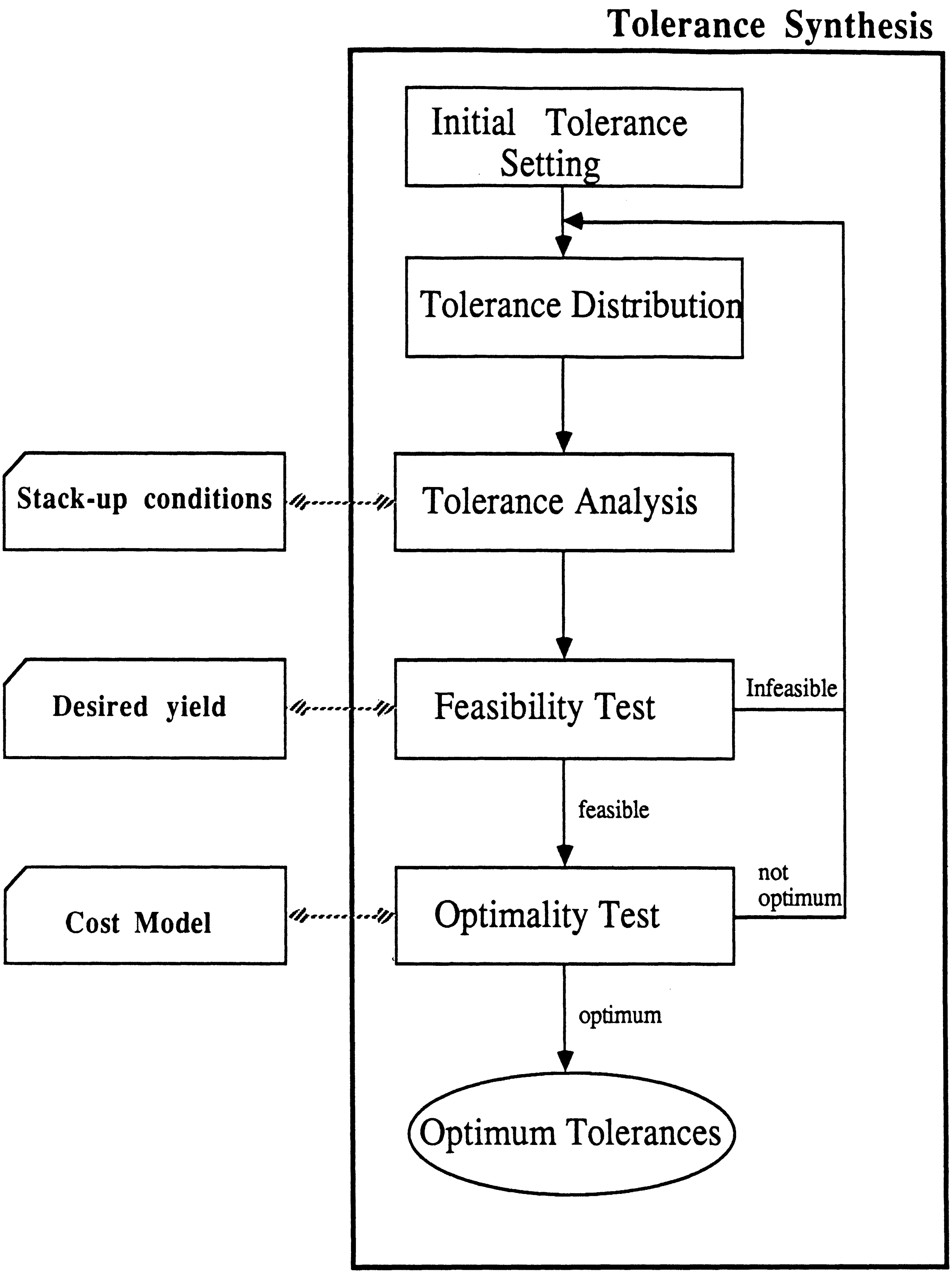


Figure 3. Basic Scheme of Tolerance Synthesis

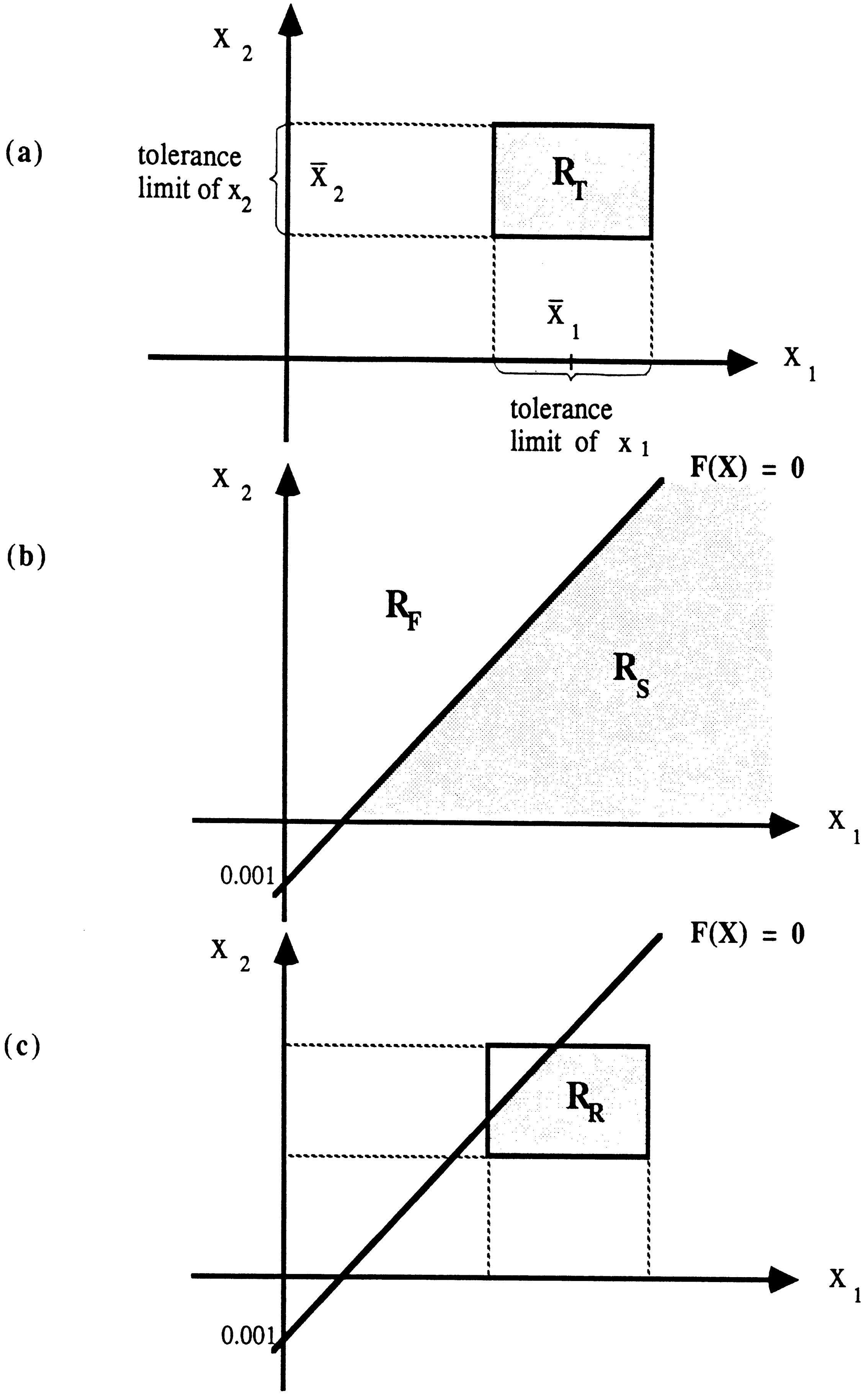


Figure 4. Tolerance, Safe, and Reliable Regions

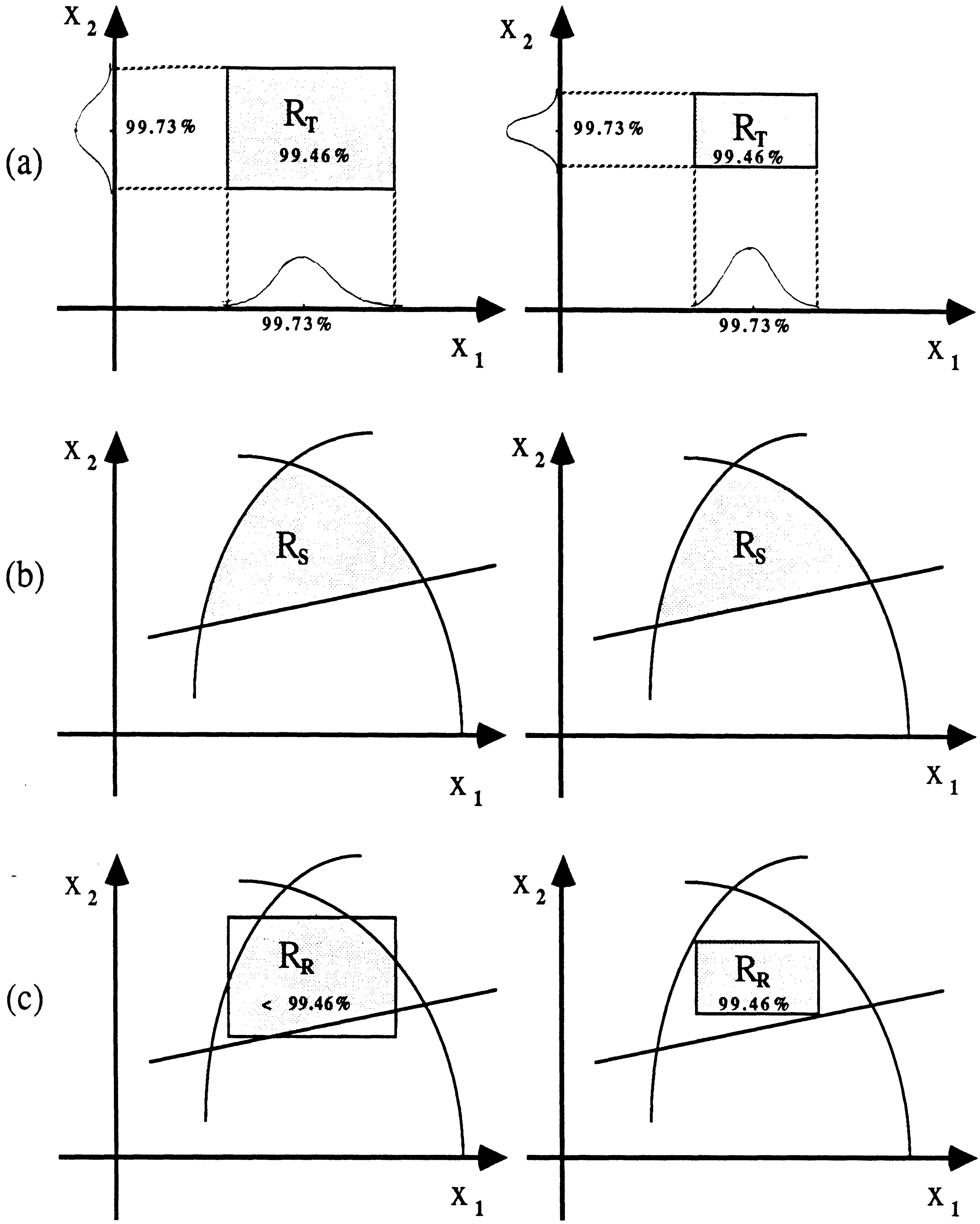


Figure 5. Effect of Tolerances on Reliable Region

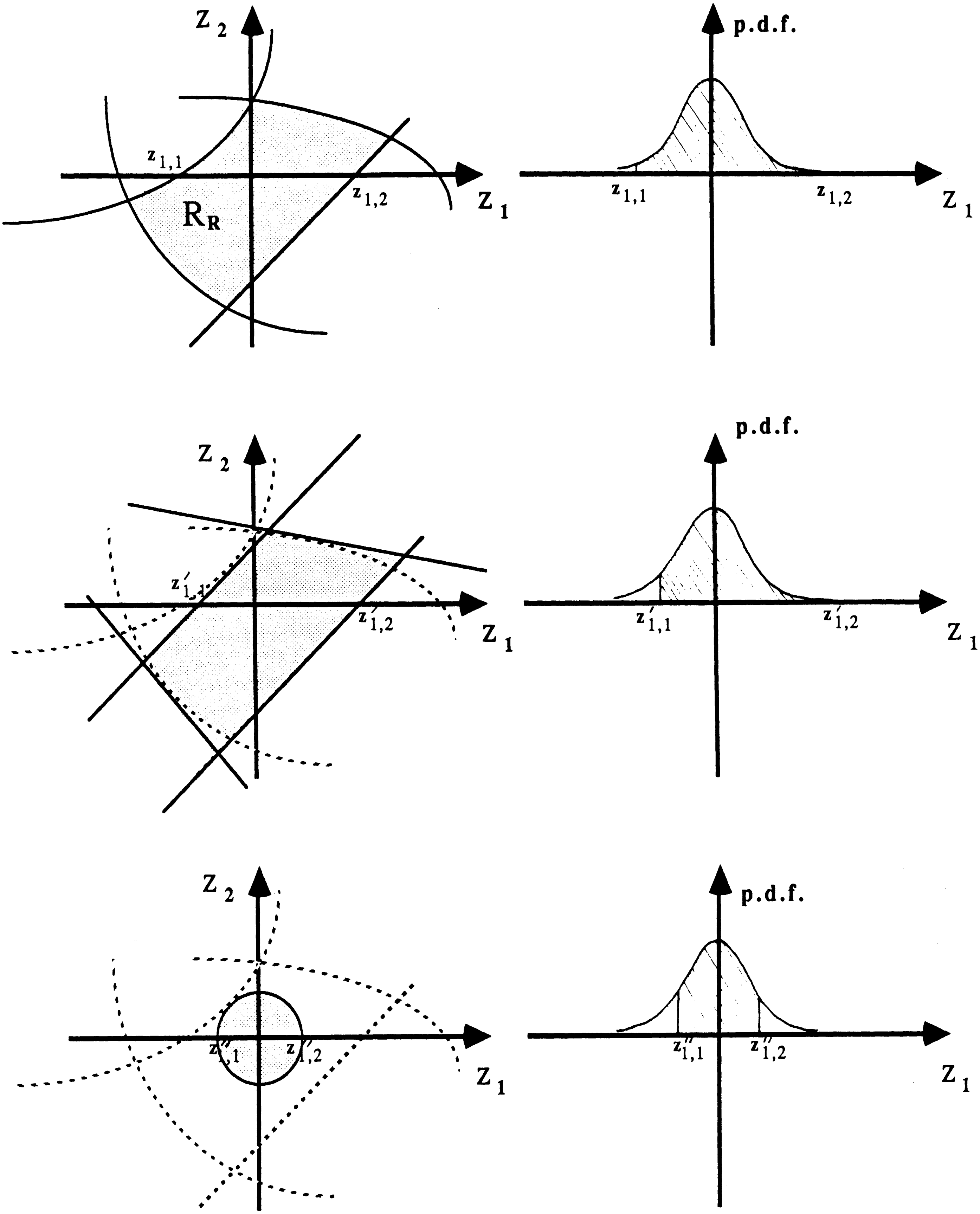
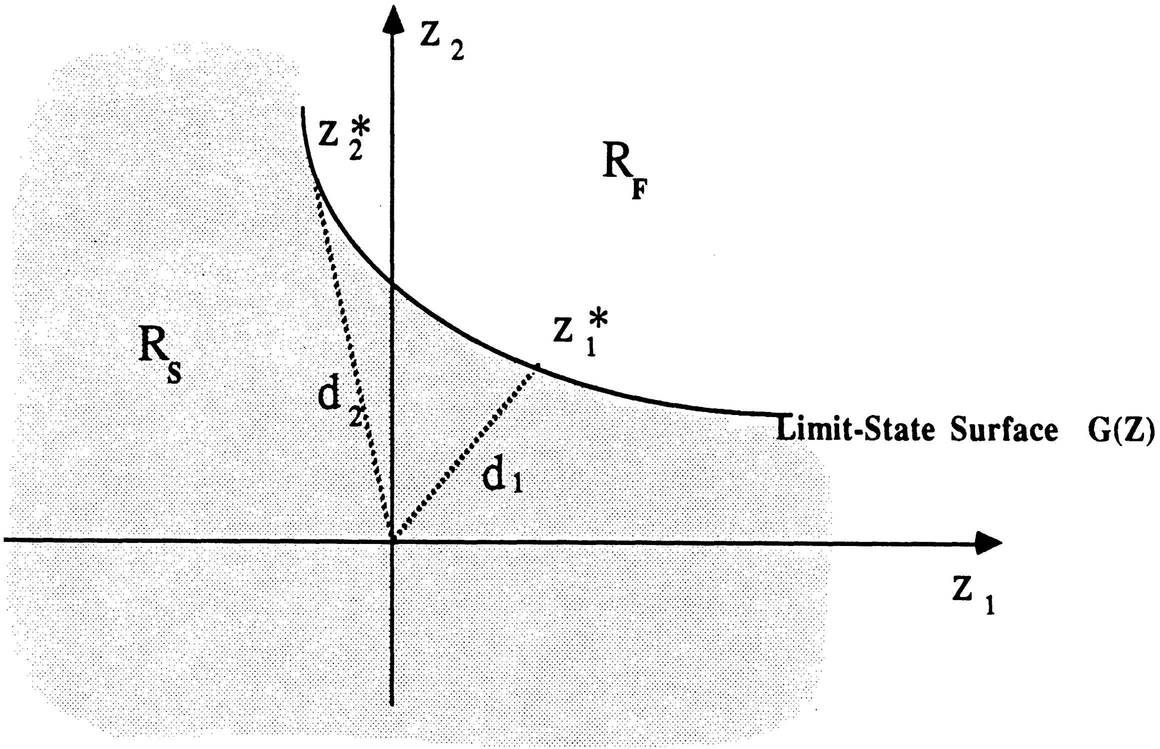
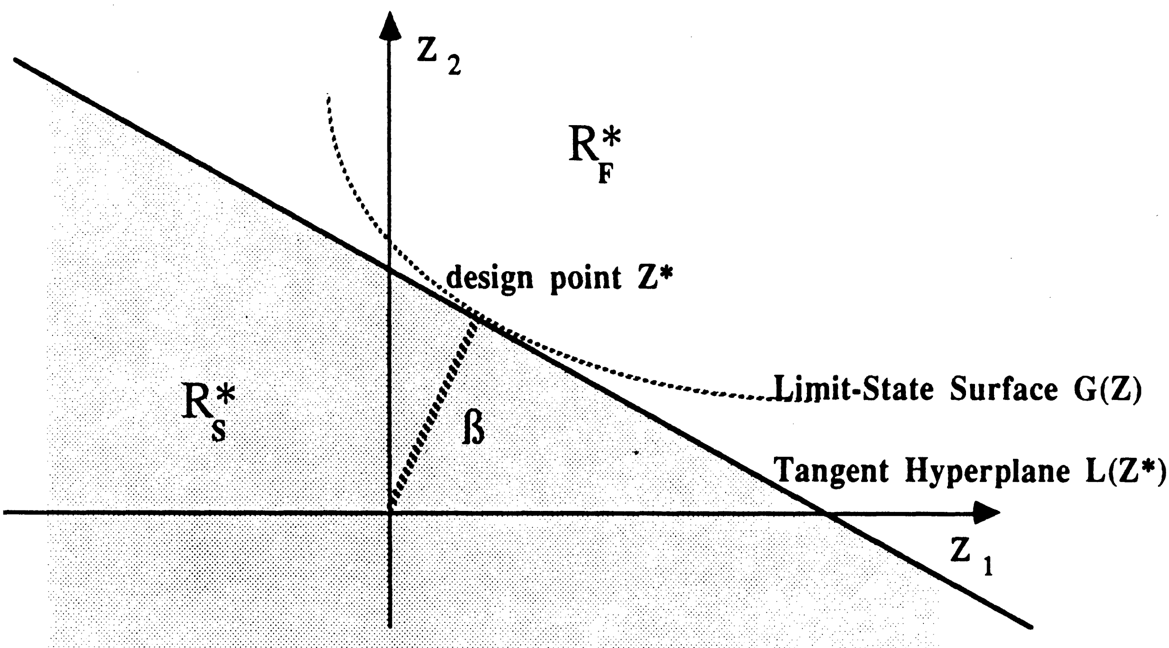


Figure 6. Approximation of R_R

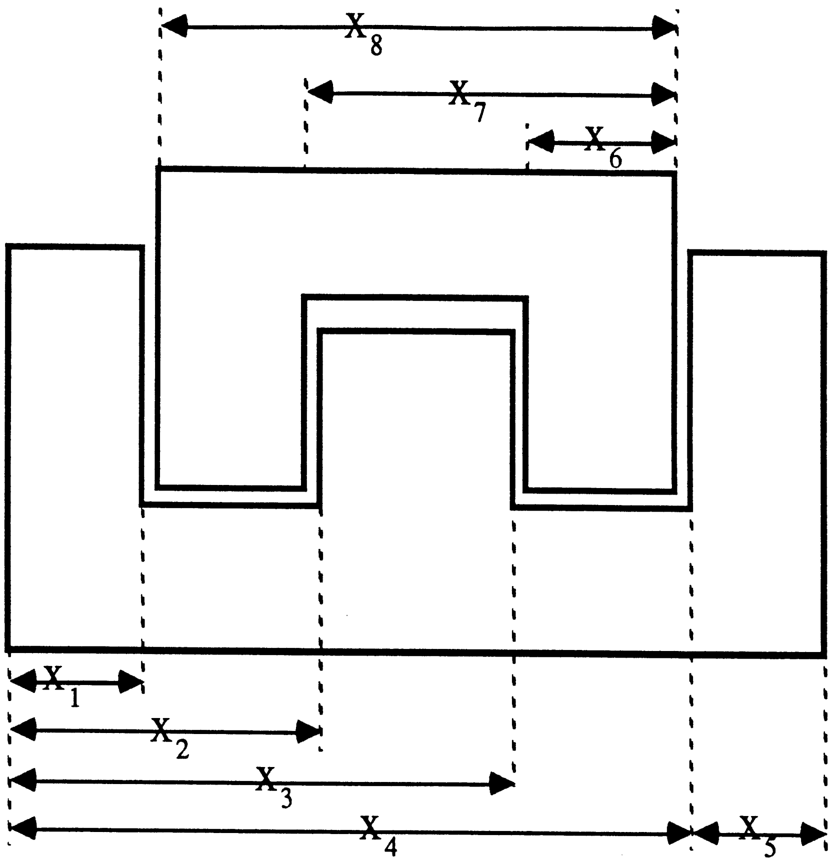


(a)



(b)

Figure 7. Reliability Index as Distance



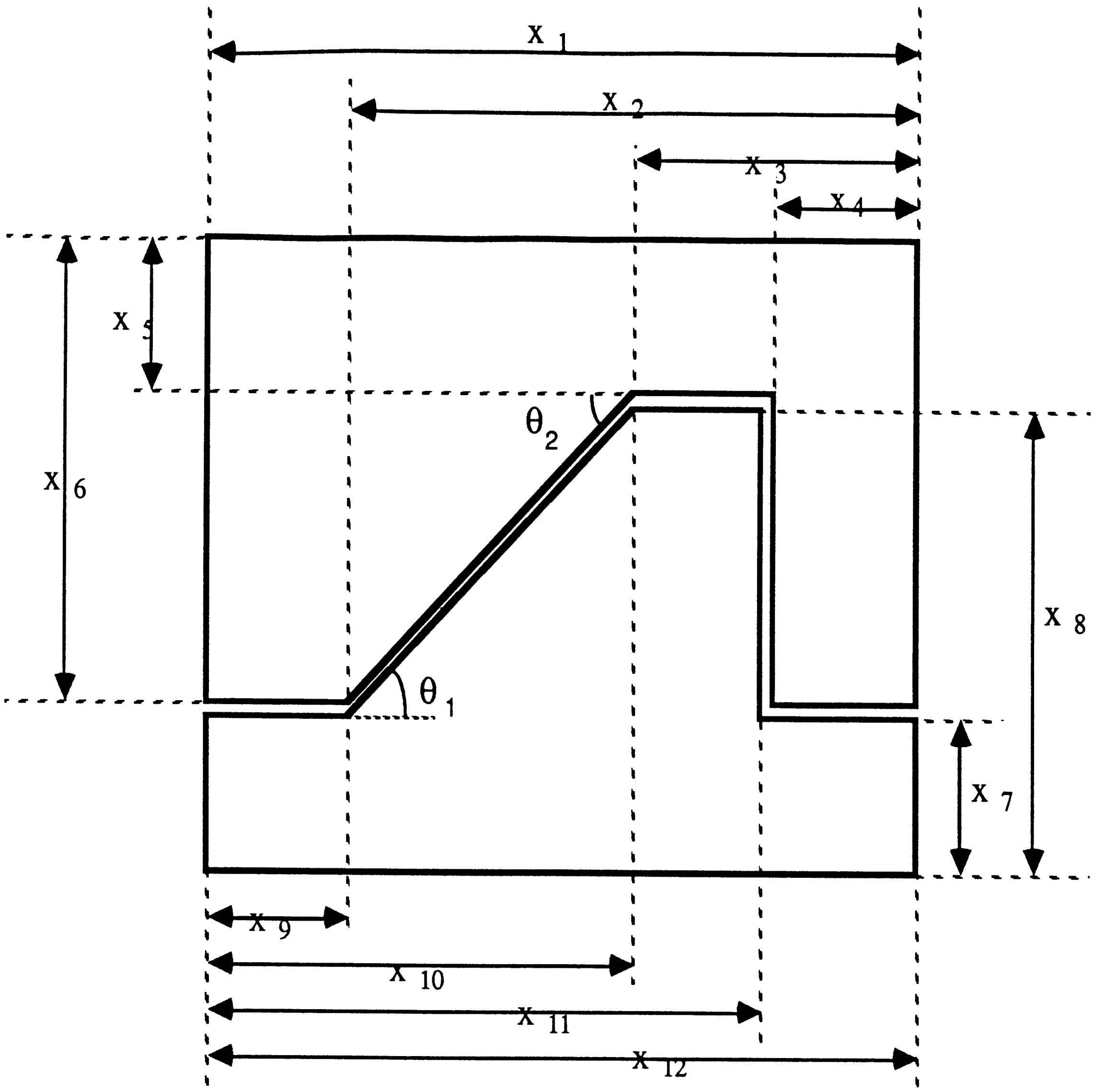
$$F_1(X) = - x_4 - x_5 + 5.005$$

$$F_2(X) = x_2 - x_1 - x_8 + x_7 - 0.0003$$

$$F_3(X) = x_7 - x_6 - x_3 + x_2 - 0.001$$

$$F_4(X) = x_4 - x_3 - x_6 - 0.0003$$

Figure 8. Example of Linear Requirement Functions



$F_1(X) = (x_6 - x_5) - (x_8 - x_7)$

$F_2(X) = (x_3 - x_4) - (x_{11} - x_{10})$

$$F_3(X) = (x_8 - x_7)*(x_2 - x_3) - (x_6 - x_5)*(x_{10} - x_9) + \tan(\pi / 180) * \{(x_{10} - x_9)*(x_2 - x_3) + (x_8 - x_7)*(x_6 - x_5)\}$$

$$F_4(X) = (x_6 - x_5)*(x_{10} - x_9) - (x_8 - x_7)*(x_2 - x_3) + \tan(\pi / 180) * \{(x_{10} - x_9)*(x_2 - x_3) + (x_8 - x_7)*(x_6 - x_5)\}$$

$F_5(X) = - x_1 + x_{12} + 0.01$

$F_6(X) = x_1 - x_{12} + 0.01$

Figure 9. Example of Linear and Nonlinear Requirement Functions

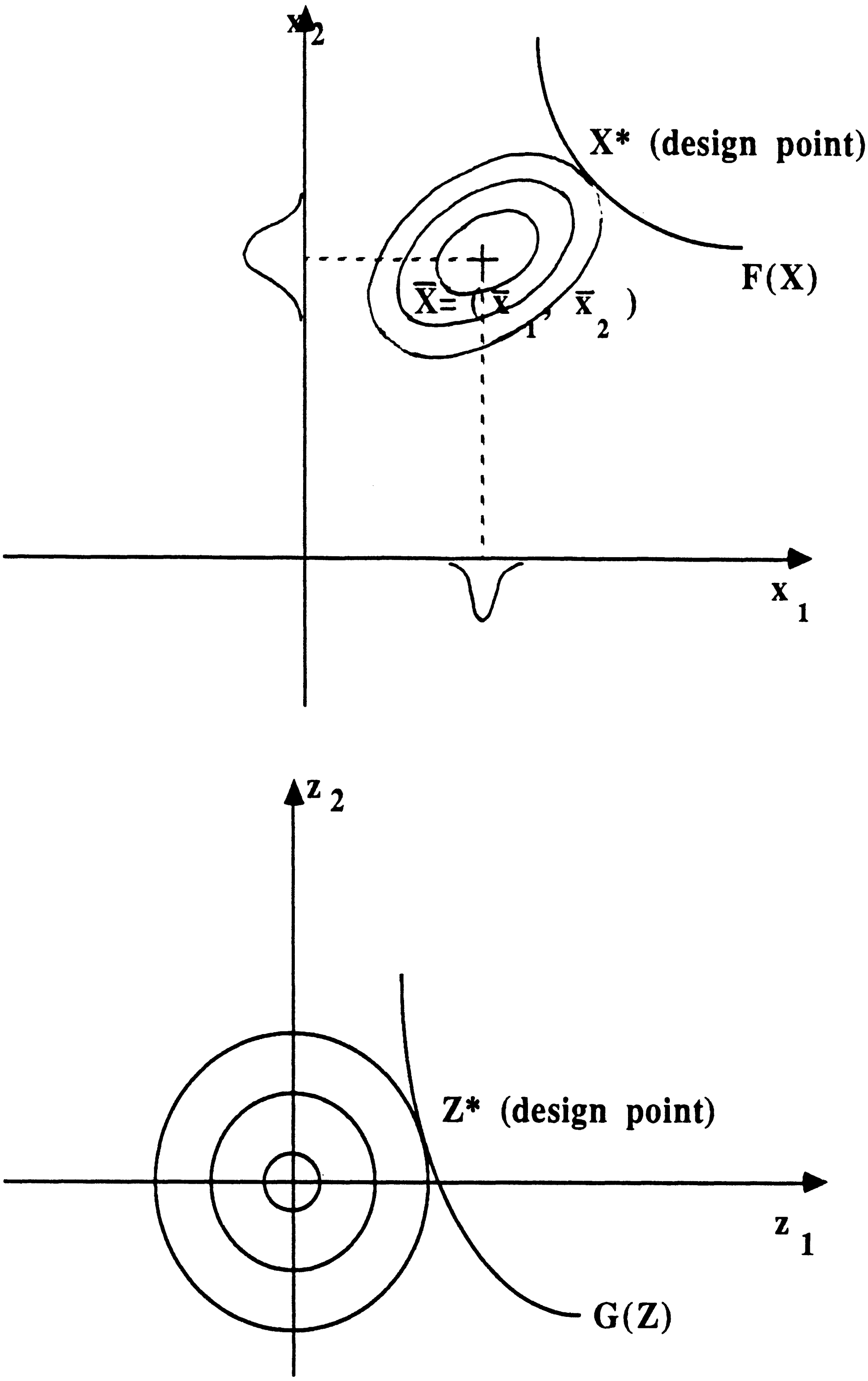


Figure 10. Transformation to the Standard System