NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:
The copyright law of the United States (title 17, U.S. Code) governs the making of photocopies or other reproductions of copyrighted material. Any copying of this document without permission of its author may be prohibited by law.
A New Perspective on Tolerance Analysis Based on Tolerance Zones

P. Padmanabhan, T. Kurfess

EDRC 24-77-91
A NEW PERSPECTIVE ON TOLERANCE ANALYSIS BASED ON TOLERANCE ZONES

by
Prakash Padmanabhan
Department of Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

Thomas R. Kurfess
Assistant Professor
Departments of Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
8 January 1991

This work has been supported by the the Engineering Design Research Center, a NSF Engineering Research Center
ABSTRACT

Universal interchangeability of parts or sub-assemblies is one of the goals of most designs. The knowledge of all possible variations in an assembly due to the specified tolerances is essential to ensure universal interchangeability. Tolerance zone, as defined in [1], is a virtual region formed around a true feature. Tolerance zone propagation can be effectively used to observe variations due to individual part tolerances in an assembly. Although the concept of tolerance zones is not explicitly used in most existing tolerance analysis methods, the underlying principles can be readily used to generate a tolerance zone. However, it is seen that tolerance zones generated by Worst Case or Root Sum Square (RSS) models do not encompass all possible variations in an assembly made of parts that are within tolerance specification, because the effect of non-worst case stacking situations is not considered. The new perspective on tolerance analysis, discussed in this paper, addresses the effect of non-worst case stacking situations by studying angular accumulation in addition to linear accumulation of tolerances. In a general plus/minus tolerated part there is a high probability that the part conforms to the specified tolerances with a surface that is not at the true angular position with respect to other features. Such parts when assembled together lead to angular stack-up. In this paper, angular accumulation is addressed using the concept of tolerance zones and a mathematical basis is developed for the representation and analysis of angular accumulation based on classical kinematic theories. Possible applications of this analysis are also discussed.

INTRODUCTION

As tolerances affect both cost and quality of a product, tolerancing is now considered a critical engineering design function. As such, tolerance allocation is a significant task that deserves considerable attention and consideration. Designers usually specify tight tolerances to ensure high quality while the manufacturing community prefers loose tolerances to reduce manufacturing cost, as well as to ease manufacture. However, in order to both achieve the desired performance and aid the manufacturing community designers must specify reasonable tolerances.
Only in the most simple cases can one generate the necessary tolerances from the desired result. In general, there are no specific guidelines for allocating tolerances for any component. The most common practice is to allocate some tolerance that seems appropriate on the basis of experience or intuition, and then conduct an analysis to ensure that the allocated tolerance suits the desired design function. In order to do this, the designer must be able to realize all possible effects of the tolerances specified, especially if universal interchangeability is one of the design goals.

The effects of specified tolerances are generally analyzed by creating an analytical model that can predict the accumulation of tolerances in an assembly. Prediction of tolerance accumulation is necessary because critical fits, clearances, etc. are usually controlled by the accumulation of several component tolerances. Some of the common ways to analyze tolerance accumulation are the Worst Case Model, statistical methods like Root Sum Squares (RSS) and the Estimated Mean Shift Model.

**Worst Case Analysis**

This analysis, one of the earliest developed techniques, is used mainly to ensure that all assemblies made from various interchangeable components meet the assembly tolerance requirements. It is based on the idea that when all components in an assembly occur at their worst (max. or min.) limit simultaneously the worst possible assembly is obtained. Hence the sum of the individual tolerances represent the variation in the final assembly due to linear stack-up. A tolerance zone can be obtained based on this analysis.

Mathematically,

\[
dU = \sum_{i=1}^{n} \delta_i
\]

where \(\delta_i\)'s are the individual component tolerances and \(dU\) is the assembly tolerance. As long as \(dU\) is less than the assembly tolerance for all cases, all components within the tolerance limit will meet the specified assembly requirements [Fortini 1967].

This is a simple model and is useful where there are no complex form tolerances. It makes no assumption as to how the parts are distributed in the
tolerance zone. When tolerances are allocated using this method, the individual component tolerances will be greatly reduced in assemblies containing many components and this leads to increased production cost. This method and the tolerance zone obtained from it are based on linear stack-up. However, in a later section, it is shown that the tolerance zone obtained on the basis of linear stack-up is insufficient to encompass all possible assemblies made from "good" parts (parts that conform to the specified tolerances).

**Root Sum Square (RSS) Method**

Another common analytical model for the analysis of tolerances is RSS. This method is based on the assumption that the probability of all components occurring at their worst case simultaneously is rather low when component variations occur in the form of normal distribution. As such, the assembly variation in most cases will be less than that predicted by worst case model. The assembly variation is given by [Fortini, 1967]

\[ dU = \sqrt{\sum_{i=1}^{n} \delta_i^2} \]  

(2)

However, by considering the distribution of tolerances, 100% assemblability is lost. Even in the case where a small percentage of rejections are acceptable it cannot account for any naturally occurring shifts in production processes. Though it is often used to predict the likely rejections in a mass production and it predicts less rejections than that would normally occur in real processes [Greenwood and Chase, 1987]. It cannot be used to validate a design completely, especially when only a few products are manufactured (i.e., when the central limit theorem cannot be invoked).

**Estimated Mean Shift Model**

This is an improvement on the above two methods [Greenwood and Chase, 1987]. It incorporates a balanced mix of both worst case model and RSS model in order to account for naturally occurring shifts in production processes. The method is based on resolving the component tolerance into two parts, mean shift and variability about the mean. This is done by
selecting a shift factor $f_i$ (0 ≤ $f_i$ ≤ 1) for each component. The resulting assembly tolerance is given by

$$dU = \sum_{i=1}^{n} f_i \delta_i + \sqrt{\sum_{i=1}^{n} (1-f_i)^2 \delta_i^2}$$

(3)

The first summation is composed of estimated mean shifts, which are treated as a worst case model. The second summation represents component variability, which is treated as a sum of squares. Note that if the shifts are 100% (i.e., $f_i = 1.0$), the model results in a Worst Case Model and if the shifts are 0% (i.e., $f_i = 0.0$), it results in a simple RSS method.

This method is useful when one part in an assembly has poor control during its manufacture. The effect of this part on the assembly tolerance can be taken into account by varying only the shift factor concerned with it. This accounts for bias or naturally occurring shifts realistically. This method has the same disadvantages as RSS model except that it can account for shifts. It is useful for tolerance allocation of parts that are mass produced, but not for analysis of tolerances for small batch productions or for a thorough evaluation of a design. It is shown that the tolerance zones obtained by the above methods are insufficient as they do not include some of the assembly possibilities that arise due to angular stack-up.

Statistical Methods and Monte Carlo Analysis

We can also study the effects of tolerances without actually analyzing accumulation of tolerances, by using statistical methods like the Method of Moments and the Method of Integration, or by sampling techniques like Monte Carlo analysis.

Statistical methods obtain the system output scatter that results from random variables (tolerances) whose means are the nominal design values and whose variances are related to tolerances [Cox, 1979]. The procedure is to obtain a probability density function (pdf) of this combination of variables and predict the variability of system performance. Based on the system output scatter one can obtain the fraction of the assemblies that do not function correctly due to the allocated tolerances. Most methods are based on the policy that as long as this fraction of the mechanisms that do not function properly due to tolerances is less than or equal to the fraction due to other
causes (material defect, mis-assembly, etc.) the assigned tolerances are adequate [Evans, 1974].

Some of the disadvantages are that the precision of most statistical methods is too limited for some stack tolerancing analyses and variations in manufacturing processes can make the describing pdf not very suitable for accurate predictions [Cox, 1979]. Also statistical methods usually give an estimate of what percentage of assemblies will be acceptable and this is not what is required for assuring universal interchangeability.

However, Monte Carlo analysis, a powerful tool for both linear and non-linear tolerance analysis and for distributions that are not normal, does not have the above disadvantages. This method operates by generating large number of assembly instances each of which corresponds to a point in the tolerance region. Each instance is then checked to determine whether it lies in the design tolerance region. The main disadvantage of Monte Carlo analysis is that it is highly time consuming and expensive to perform. Therefore, it cannot be used iteratively in the early design stages where tolerances must be allocated, analyzed and changed repeatedly to come up with a good design. It also requires a large sample if its results are to be sufficiently accurate.

Tolerance analysis using the above methods does not give a complete picture of the permissible variations due to the specified tolerances as the effect of angular stack-up is not considered. The tolerance zone obtained from the new perspective includes the effect of angular variations in inter-related part features and the effect of stacking non-worst case models.

THE NEW PERSPECTIVE.

The tolerances that matter most are affected by linear and non-linear stack-up of other tolerances [Chase et al, 1990]. Inadequate tolerance analysis may result in unacceptable assemblies even though all manufactured parts for the assembly are within tolerance limits. The modest investments that are required for a thorough tolerance analysis may and usually do yield substantial benefits [Chase et al, 1990]. Solid modeling and CAD systems are now quite easily accessible and are invariably available in any industry. The visual capabilities of these systems can be used to analyze tolerances more
thoroughly with the proposed new perspective. This new perspective in
tolerance analysis is based on viewing the various possible effects of a
tolerance specification on a solid modeler or CAD system by looking at
tolerance zones. ANSI Y14.5M-1982 [1] states that tolerance zone is a virtual
region formed around the true feature. A visual model with an analytical
basis is considered and tolerance zone representation is emphasized instead of
just an analytical analysis which gives a number that represents the total or
maximum variation in assembly dimensions. Simulation of an assembly on
a solid modeler can be used to give a better idea on the effect of tolerances on
the assembly. An example of a simple assembly is analyzed to elucidate this.

Consider an assembly consisting of a stack of six blocks. The tolerance
specification on a single block is shown below. The tolerance shown is very
simple and straightforward.

![Figure 1. Toleranced Block](image1)

![Figure 2. Tolerance Zone](image2)

The figure on the right represents the tolerance zone (shaded area)
resulting from the specified tolerance. Now let us analyze the effect of
tolerances when six similar blocks are stacked one on top of another to make
an assembly.

The worst case analysis predicts a tolerance zone as shown in Figure 3.
RSS analysis leads to a tolerance zone as shown in Figure 4. Notice this
tolerance zone is smaller than the one predicted by worst case analysis because
the probability of all components occurring at their worst limit is considered
here. Suppose the acceptable assembly tolerance zone is same as the tolerance
zone generated by worst case analysis. (Assembly tolerance zone is the zone
or space in which all acceptable assemblies must lie). Apparently all the assemblies will be acceptable as both analyses lead to a tolerance zone that is within the assembly tolerance zone.

The Estimated Mean Shift Model, by definition, gives a tolerance zone which is in between the above two cases. When angular stack-up is considered it is seen that there are cases where blocks which are individually within specification lead to an assembly which lies outside the tolerance zone obtained by worst case and RSS analysis. Clearly the assembly will also lie outside the tolerance zone obtained from estimated mean shift model. Hence the case of estimated mean shift model is not considered here.

Shown in Figure 5 is a block which is within the specified tolerance zone (represented by the dark area). However, the top face is slightly angular with respect to the bottom face. Consider the assembly of six such blocks. As each of the blocks are individually within the specified tolerance, the assembly should lie within the tolerance zone obtained from the worst case scenario or RSS method. However, this is not the case (Figures 6 and 7) since the angularity can accumulate (or partly cancel) depending on how the blocks are stacked up. The extreme case occurs when all of them are inclined in the same direction and amplify the angularity on the final assembly. This is shown in Figure 6 and Figure 7.
Figure 3. Worst Case Tolerance Zone

Figure 4. RSS Tolerance Zone

Figure 5. Block with angularity
The assembly lies outside the tolerance zone depicted by Worst Case and RSS models even though all of the individual blocks are within tolerance specification and hence is unacceptable. This clearly shows the insufficiency of the two methods in predicting the variations that can occur in the final assembly due to part tolerances.
For simplicity we had considered angularity that can be described as a rotation about only one of the three coordinate axes. In reality the top face of the individual blocks can be tilted in three dimensional space and hence the top surface of the assembly can lie anywhere in a volumetric space. Now the question is how do we get the volumetric tolerance zone or likely space that will contain all assemblies, made of parts that are in tolerance limits, so that this zone can be compared to the assembly tolerance zone to make sure all assemblies will be acceptable. The next section addresses this problem.

Tolerance Zone Determination

Given the tolerance specification in Figure 5 we can define two cases, which we will call angular extreme cases, (Figure 9) in addition to the conventional worst cases (Figure 8).

The stacking of these two conventional cases can be considered as two zones and the difference of these two zones will generate the tolerance zone as defined by worst case model.

The angular extreme cases can be considered as angular worst cases in two dimensions. The stacking of these angular extreme cases will give two new zones. In a two dimensional case, the union of the largest conventional
worst case and the two angular extreme cases differenced with the smallest conventional worst case gives the tolerance zone that contains all assemblies, made from parts within tolerance limits (Figure 10). When the tolerance zone thus obtained lies within the assembly tolerance zone, all assemblies are acceptable.

On a CAD system or a solid modeler it is perhaps quite easy to draw the two angular extreme case parts and generate the intermediate zones (see Figure 10) necessary to generate the tolerance zone. However, a strong analytical basis is always useful for greater understanding and manipulation of individual tolerances during design iterations or selective assembly. The next section derives the angular stack-up zones mathematically.

Analytical Representation

Homogeneous coordinate system is used to develop the analytical representation. Homogeneous coordinate system can be viewed as the addition of an extra coordinate, a scale factor, to each vector such that the vector has the same meaning if each component including the scale factor, is multiplied by a constant. The reader is referred to reference [9] for more information regarding homogeneous coordinate system.
In this system a point vector

\[ \mathbf{v} = ai + bj + ck \]  

(4)

where \( i, j \) and \( k \) are unit vectors in the \( x, y \) and \( z \) directions, respectively, can be represented in homogeneous coordinate system as a column vector

\[ \mathbf{v} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \\ \mathbf{w} \end{bmatrix} \]  

(5)

where \( a = x/w \), \( b = y/w \) and \( c = z/w \). A plane is represented by a row vector

\[ \mathbf{P} = \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} \end{bmatrix} \]  

(6)

Transformations (rotation, translation, etc.) are represented by 4x4 matrices.

Consider an angular extreme case block, block A. A coordinate frame fixed in A can be used to describe the top edge of the block knowing the homogeneous coordinates of its two end vertices (\( c \) and \( d \), see Figure 11). Let \( \mathbf{cxa}, \mathbf{cy}, \mathbf{cza} \) and \( \mathbf{dxa}, \mathbf{dy}, \mathbf{dza} \) be the \( x, y \) and \( z \) coordinates of points \( c \) and \( d \) in reference frame A.

Figure 11. Angular Extreme Case

Let another similar block, B, be placed on top of A (Figure 12). The location of the top edge \( gh \) of B is known in terms of the coordinate system fixed in B. That is, we know \( gxb, gyb, gzb \) and \( hxb, hyb, hzb \). The two reference frames are linked by a translation and rotation. In order to determine the position of block B in terms of the coordinate system fixed in A every point in B (as described in reference frame B) must be subject to a
transformation, $T_{ba}$, which is equivalent to a translation by $'ad'$ and a rotation by the angle, $q$, between $ab$ and $cd$.

Figure 12. Two angular extreme case blocks stacked up.

The translation by $ad$ ($d_xa^1 + dy_j + dz_k$) can be represented by the transformation matrix

$$T = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

(7)

and the rotation about the x-axis by an angle $\theta$ can be represented by the transformation matrix

$$R = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta & -\sin \theta & 0 \\
0 & \sin \theta & \cos \theta & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

(8)

Now any point in reference frame B can be expressed in terms of reference frame A by combining the two transformations as

$$T_{ba} = TR = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta & -\sin \theta & 0 \\
0 & \sin \theta & \cos \theta & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

(9)

Hence,
Similarly, the vertex 'h' can be determined with respect to reference frame A, and hence the position of the top edge of block B. If there is a third block C on block B we can find a transformation matrix, \( T_{cb} \) (another translation and rotation), to represent any point or plane in C in terms of the coordinate frame fixed in B. The transformation from reference frame C to A is given by

\[
T_{ca} = [T_{cb}][T_{ba}]
\]

(11)

Thus, we can represent the top most edge of any number of blocks stacked on top of one another just by knowing the exact geometry of each block. Also, we can represent the left and right edge of each block with respect to the reference frame fixed in the bottom most block to obtain the angular extreme case zone. This can be done with the second angular extreme case also and the tolerance zone can be obtained as shown in Figure 10.

This analytical approach can be extended to three dimensions easily by transforming planes as follows:

If \( P \) is a plane in reference frame B and \( T_{ba} \) is the transformation matrix that represents transforms from B to A, \( P \) can be expressed in terms of the reference frame A as

\[
Q = P^{-1}T_{ba}
\]

(12)

where \( P \) and \( Q \) are row matrices.

Applications

In addition to generating tolerance zones this approach can be used to make selective assembly of high precision parts and relax tolerance allocation on high precision, small batch production assemblies. For example, consider the case where five similar discs have to be stacked such that the end faces of the assembly have to be parallel to within 2 microns. If the assembly tolerance is allocated equally among all five discs each disc has to be made with a parallelism of 0.4 microns and is not very easy. However, four discs can be manufactured to an accuracy of 2 microns or a little less and using sophisticated measurement techniques the position of the top faces of each
disc can be obtained and using the above analytical approach the position of
the top face of the four disc sub-assembly can be obtained and the fifth piece
can be manufactured accordingly so that the assembly dimensions are as
required.

Further if we have twenty discs manufactured to one micron accuracy
and we want to make four assemblies as above, it is easier to get the geometry
of each disc and use the above analytical approach along with some kind of
optimization technique to try different combinations of discs such that all
four assemblies meet the assembly criteria of 2 micron parallelism.

Conclusion

The existing tolerance analysis methods do not cover the instances of
angular accumulation and hence are insufficient to ascertain universal
interchangeability of parts or sub-assemblies. The new perspective based on
tolerance zones encompasses instances of both linear and angular
accumulation, and hence is clearly better suited for analyzing designs
involving universal interchangeability. Further, this method has the
advantage that it can be easily extended to three dimensions.
References


Appendix

The analytical representation developed is applied to our two dimensional example. Figure. 13 shows two similar angular extreme case blocks, A and B, with the coordinates of each point shown in their respective reference frames. The angle of the top edge of each block with respect to the bottom edge is 0.01 radians. The point a can be represented in homogeneous coordinate system as

\[
a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\]

Similarly all other points in A and B can be represented in their respective coordinate system. When the two blocks are placed on top of another the position of points e,f,g and h can be obtained by subjecting them to a transformation, \( T_{ba} \), equivalent to a translation by \( ad \) and a rotation by 0.01 radians about the x axis, fixed in reference frame A. The translation is given by

\[
T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

and the rotation is given by
\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(0.01) & -\sin(0.01) & 0 \\
0 & \sin(0.01) & \cos(0.01) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

The transformation is given by
\[
T_{ba} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(0.01) & -\sin(0.01) & 0 \\
0 & \sin(0.01) & \cos(0.01) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
\[
T_{ba} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.995 & -0.01 & 0 \\
0 & 0.01 & 0.995 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

Now, the position of \( g \) when block B is placed on top of block A is given by its coordinates in reference frame A and we can obtain it by applying the transformation \( T_{ba} \) to the coordinates of \( g \) in the reference frame B.

\[
\begin{bmatrix}
g_{xa} \\
g_{ya} \\
g_{za} \\
1
\end{bmatrix} = T_{ba} \begin{bmatrix}
g_{xb} \\
g_{yb} \\
g_{zb} \\
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
g_{xa} \\
g_{ya} \\
g_{za} \\
1
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.995 & -0.01 & 0 \\
0 & 0.01 & 0.995 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
0 \\
1.1 \\
2.1045 \\
1
\end{bmatrix} = \begin{bmatrix}
0 \\
0.9840 \\
2.1045 \\
1.0000
\end{bmatrix}
\]

Similarly all other points can also be obtained in order to determine the position of block B. The position of the third block can be obtained in terms of reference frame B with the transformation matrix \( T_{cb} \) and in terms of reference frame A by transformation \( T_{ca} \) which is given by the product of \( T_{cb} \) and \( T_{ba} \). Since the blocks are identical in our case we can write

\[
T_{ca} = T_{ba}T_{ba}
\]

Extending this argument the position of the sixth block F can be obtained by

\[
T_{fa} = (T_{ba})^5
\]

\[
T_{fa} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.9743 & -0.0490 & -0.0990 \\
0 & 0.0490 & 0.9743 & 4.9493 \\
0 & 0 & 0 & 1.0000
\end{bmatrix}
\]
If $pq$ is the top edge of block F and the homogeneous coordinates of $p$ and $q$ in reference frame F are

\[
p = \begin{bmatrix} 0 \\ 0 \\ 1.1 \end{bmatrix} \text{ and } q = \begin{bmatrix} 0 \\ 1 \\ 1.1 \end{bmatrix}
\]

respectively then the coordinates of the top edge of the assembly of the six blocks is given by the two points

\[
p = \begin{bmatrix} -0.1480 \\ 5.9235 \\ 1.0000 \end{bmatrix} \text{ and } q = \begin{bmatrix} 0.8214 \\ 6.0699 \\ 1.0000 \end{bmatrix}
\]

Thus we can obtain the extreme angular position of the assembly. The angular zones required to obtain the final tolerance zone of the assembly (Figure. 10) can be constructed by calculating the homogeneous coordinates of key points in the zone and defining a polygon. Having done this actual tolerance zone which encompasses all possible assemblies made from "good" parts can be obtained by simple union and difference operations between polygonal zones. The result can be easily displayed on a computer screen.