

Navigating the Tolerance Analysis Maze

Jami J. Shah¹, Gaurav Ameta¹, Zhengshu Shen² and Joseph Davidson¹

¹Arizona State University, Tempe, {j.shah, j.davidson, gameta}@asu.edu

²UGS Corporation, shenzhen@ugs.com

ABSTRACT

This paper classifies and reviews geometric tolerance analysis methods and software. Two of the most popular methods, 1-D Min/Max Charts and Parametric Simulation, are reviewed in detail. The former is fully consistent with international tolerance standards but limited to decoupled analysis of variations in one direction at a time. It is also hard to automate. The latter can handle variations in all directions but is not fully compatible with the standards. The results are highly dependent on the expertise of the analyst. New methods are emerging to overcome these problems. One such method called T-maps is also reviewed. It maps geometric variations from physical space to 3, 4, or 5 dimensional virtual space and computes tolerance accumulations using Minkowski sums. The method so far has shown to be consistent with tolerance standards and can handle all tolerance classes applicable to planar and cylindrical features.

Keywords: GD&T, CATS, Math modeling of tolerances, Computer-aided Tolerance Analysis.

1. INTRODUCTION

Understanding the causes and effects of dimensional and geometric variations is a major concern in the design and manufacture of mechanical products. Designers are essentially concerned with the following geometric dimension and tolerance (GD&T) issues:

- Functionality and/or assemblability
- Tolerance Analysis: consequences of a proposed GD&T scheme
- Tolerance Allocation: determining how to distribute the allowable variation on the dimension of interest amongst all the independent contributors

Unfortunately, today's tools for assisting designers in allocating tolerances and identifying trade-offs during the design process are neither comprehensive nor accurate. The designer must choose between two approaches, both of which are severely deficient. One approach is a manual procedure called Min/Max Tolerance Charting; it is consistent with ASME 14.5/ISO standards [4,18], but limited to 1-D worst case analysis only. The other approach is to use commercial tolerance packages that typically perform both worst case and statistical analysis, but based on point-to-point constraint solving. This makes them incompatible with the current tolerance standards that are based on tolerance zones, not point-to-point variations. Comprehensive 3D analysis of stack-ups involving all types of dimensional and geometric variations is only possible if a mathematical model of such variations exist. But the current international standards are based on ad-hoc conventions collected from years of engineering practice, not on mathematical foundations. The attempt to "retrofit" an "official" math model to the tolerance standard has not gone far enough [5]. Researchers have proposed replacing the standard completely—a proposition unacceptable to industry because the valuable empirical knowledge contained in the current standard will be lost. Emerging methods are attempting to address the challenge to build a math model of geometric variations that is consistent with already existing tolerance standards and capable of supporting comprehensive 3D analysis of stack-up conditions.

2. GD&T STANDARDS OVERVIEW

There are two similar standards in use for specifying tolerances on drawings, ISO 1101 and ANSI/ASME Y14.5M [4,18]. The standards have classified dimensional variations (size) and geometric variations (form, orientation, profile, position, runout) into separate classes (Figure 1). This is because the types of variation that need to be controlled depend on functional and assembly requirements. For example, form needs to be controlled for smooth motion, perpendicularity is important for insertion of long features, and feature size and location must be controlled for proper

assembly. It is not possible to give a review of the entire Y 14.5 standard here, but it is important to note the following points contained in it:

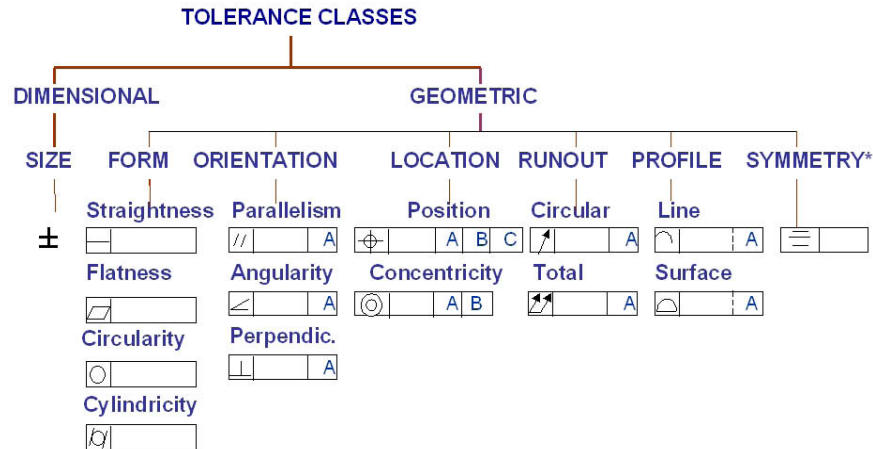


Fig. 1: Tolerance classification in GD&T standards.

- each geometric tolerance class is represented by a region (zone); the shape of the zone depends on the tolerance type and the feature being tolerated; the size depends on the tolerance value, material condition modifiers, and certain rules; the position/orientation of the zone depends on the tolerance type and datums. Fig. 2a shows zones for perpendicularity and flatness; Fig. 2b shows runout zone inside size limits.

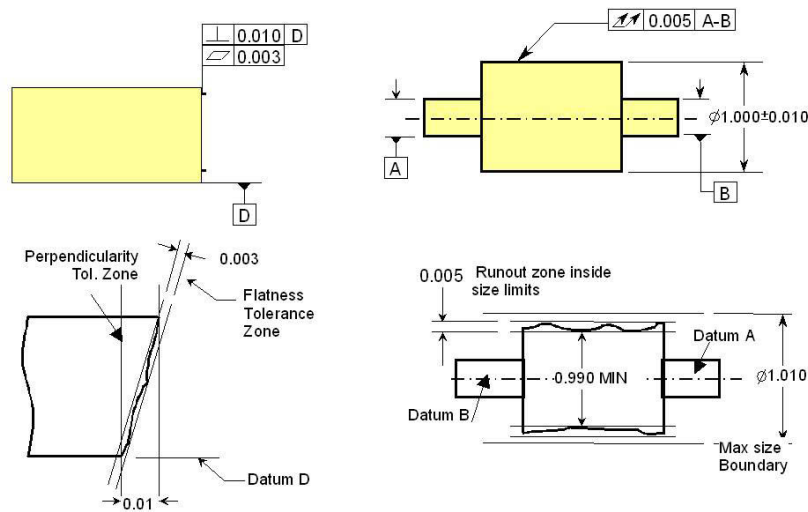


Fig. 2: Tolerance zones.

- datums are references for measurements; they are neither on the part nor on the gage, but simulated by the contact between the two; all D&T relations (except size) are one-directional, i.e. datum-to-target.
- the number of datums used in a tolerance specification depend on the number of degrees of freedom that need to be controlled. When multiple datums are used, the order in which they are specified creates a precedence order used to determine the co-ordinates and directions of control (Fig. 3)
- Rule#1 requires perfect form at max material condition (MMC), i.e. form and orientation tolerance zones must shrink to zero as the feature approaches MMC (max. size for external feature; min. size for internal feature). As a consequence, certain tolerance zones float within other zones; the form tolerance zone has floating position and orientation within a size (see Fig. 4) or orientation zone; an orientation zone floats only in position inside the size zone.

- tolerances can be applied to both resolved entities (axes, mid-planes), not just to boundary elements (faces, edges);
- in order to allow for trade-offs between feature size and certain types of geometric tolerances, such as position, the standards use material modifiers (MMC, LMC, RFS) to indicate what the geometric tolerance is when the size is at its largest or smallest value. When the feature size deviates from that value, a "bonus tolerance" is added to the geometric tolerance, i.e. trading position variation for size variation; when a modifier is applied to a datum feature of size, the geometric tolerance zones "shift" which is equivalent to a larger zone.

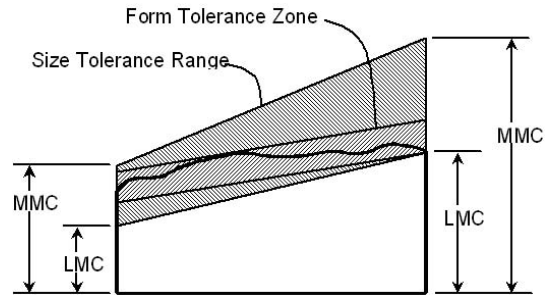
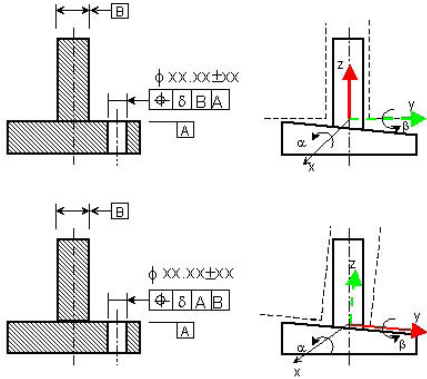


Fig. 3: Effect of datum precedence on measurement direction. Fig. 4: Form zone floating inside size limits.

The standard was not based on any mathematical model but it contains valuable experiential knowledge collected from decades of engineering practice. This lack of a mathematical basis for Y14.5 causes ambiguities in interpretation of the standard, makes it difficult to do tolerance analysis consistent with the standard, and to verify results obtained from analysis software.

3. WHAT IS TOLERANCE ANALYSIS?

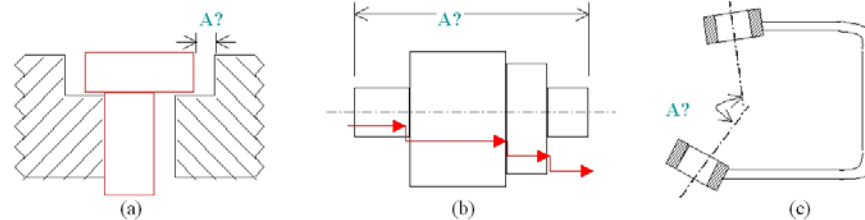


Fig. 5: Examples of analyzed dimension (A) used in tolerance analysis.

The purpose of Tolerance Analysis is to study the accumulation of variations on a geometric attribute of interest (dimension, location, orientation, etc, such as those shown in Fig 5). The need for this arises because the analyzed dimension ($A?$) is not explicitly specified. The most common case is analysis of clearances in assemblies. All dimensions and tolerances that affect A are called the contributors. The stack path, also called a tolerance chain, datum flow chain or dimension loop, is a continuous and shortest series of specified dimensions from one feature of interest to another (start & end of stack). The analysis can involve multiple parts in an assembly, as in Fig 5 (a) or variations on a single part, as in (b) and (c). Worst case analysis is done to determine the maximum, minimum values of A resulting from the limits specified on the contributors. Statistical analysis is used to determine the full frequency distribution of A computed from frequency distributions of the contributors. Whether tolerances are allocated on the basis of worst case or statistical analysis is a matter of economics. Worst case design guarantees 100% interchange-ability and assemblability of parts but is costly. By relaxing the requirement to say 99.97% acceptance, tolerances can be more liberal, thus reducing manufacturing cost. However, one must look at the cost of rejected parts (or selective assembly) to see if the reduced manufacturing cost results in a net gain.

Tolerances are allocated by trial and error. A tentative set of specifications is analyzed to see if the design requirements are met. If not, designers need to determine which tolerance to change and by how much. Two types of information can guide the designer in this search: Sensitivity and Percent Contribution. Sensitivity is the first derivative of A with respect to each contributor and Percent Contribution is an indication of the proportion of the variance of A attributed to each contributor. Obviously, one needs to focus on the biggest contributors to bring A under control and pay careful attention to the contributors with the highest sensitivities.

4. HOW IS TOLERANCE ANALYSIS DONE?

We have found that there are three different communities (designers/draftsmen, engineering analysts, and design researchers) using vastly different tools and techniques. Cultural and education differences between these communities have kept them isolated from one another and unaware of the other's techniques. The draftsmen community uses a manual procedure called Min/Max Tolerance Charts [19,30]; this method is true to Y14.5 but can do only worst-case analysis, and in only one direction at a time. The engineering analysis community uses computer-aided tolerance-software (CATS), such as *Mechanical Advantage* [10], *VSA* [13], *3DCS* [16], *e-TolMate* [9]. These can do both worst case and statistical analyses, but because they are based on the parametric approach discussed in Sec. 7, they have the limitations mentioned previously, such as partial compatibility with Y14.5. Some CATS systems use the same parametric equations as those used in building the CAD model, while others build an abstracted feature model specific to each analysis case. Emerging methods, such as T-maps [12] claim to have overcome the problem of compatibility with standards while providing full 3D worst case and statistical analysis. Table 1 explains the differences between these techniques. The rest of the paper elaborates on each of these analysis methods.

	Dimensionality	Analysis Type	Scope (Tolerance type)	Y14.5 support	Level	Linearization	Automation
Tolerance Charts	1-D	worst case only	geometric & dimensional	full	part & assembly	n/a	manual
Parametric CATS (e.g., <i>MechAdv</i> , <i>IDEAS</i>)	2, 3-D	worst + statistical	dimensional & <u>some</u> geometric	partial	part & assembly	Linearized	Interactive
CATS with abstract feats. (<i>3DCS</i> , <i>VSA</i> , <i>eTolmate</i>)	3-D	worst + statistical	dimensional & <u>some</u> geometric	partial	part & assembly	Non-linear	Interactive
Multi-variate regions (e.g., T-maps)	3-D	worst + statistical	dimensional & <u>all</u> geometric	full	part & assembly	Non-linear	Automated

Tab. 1: Capabilities comparison of Tolerance Analysis Methods.

5. BASIC PRINCIPLES

The simplest problem is that of worst case analysis in one direction considering size variations only, i.e. dimensional only (no geometric tolerances). The analyzed dimension A will simply be an arithmetic sum or difference of the contributing size dimensions d_i . It is easy to see that

$$\text{If } A = \pm d_1 \pm d_2 \pm d_3 \pm \dots \text{ where } d_i = \bar{d}_i \pm \Delta d_i$$

$$\text{then mean value } \bar{A} = \pm \bar{d}_1 \pm \bar{d}_2 \pm \bar{d}_3 \pm \dots = \sum \bar{d}_i \quad (1)$$

$$\text{and } \Delta A = A_{\max} - A_{\min} = 2 * (\Delta d_1 + \Delta d_2 + \Delta d_3 + \dots) = 2 * \sum \Delta d_i \quad (2)$$

where Δd_i are the equal bi-lateral tolerances for contributor i about mean value \bar{d}_i .

For simplified statistical analysis it is typically assumed that the contributors d_i are all normally distributed and their nominal values are set at the mean, and the equal bilateral tolerances are set at $\pm n$ *sigma (n times the standard deviation, typically ± 3). Therefore, the mean and standard deviation of A can be found from:

$$\bar{A} = \bar{d}_1 \pm \bar{d}_2 = \sum (\text{sign}_i) d_i \quad (3)$$

$$\sigma_A^2 = \sigma_1^2 + \sigma_2^2 = \sum \left(\frac{\Delta d_i}{n} \right)^2 \quad (4)$$

Suppose that the acceptable values of A lie between UL and LL (see Fig. 6), then the shaded area under the curve, which can be found from statistical Tables, will give the Acceptance Rate (% parts acceptable).

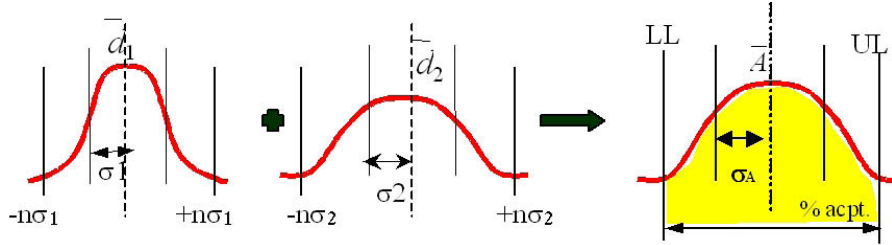


Fig. 6: Simplified statistical analysis.

For 2D or 3D analysis even considering size variations only, A will be a non-linear function f of the dependent variables:

$$A = f(d_1, d_2, d_3, \dots)$$

Two common approaches for non-linear analysis are

- Linearization of function f , leading to a closed form solution for Gaussian distribution
- Monte Carlo simulation, which can handle any distribution but requires thousands of simulations

The linearized approach combines Taylor series expansion (to linearize) with random variable arithmetic, as explained below. Linearization of a function f of multiple variables d_i about a fixed point $d_i = a_i$, yields the variation of A in terms of the sensitivities and variations of the contributors Δd_i .

$$A \approx f(d_1, d_2, d_3, \dots) + \left. \frac{\partial f}{\partial d_1} \right|_{a,b} (d_1 - a_1) + \left. \frac{\partial f}{\partial d_2} \right|_{a,b} (d_2 - a_2) + \left. \frac{\partial f}{\partial d_3} \right|_{a,b} (d_3 - a_3) + \dots$$

$$A - f(d_1, d_2, d_3, \dots) \approx \frac{\partial f}{\partial d_1} \Delta d_1 + \frac{\partial f}{\partial d_2} \Delta d_2 + \dots$$

$$\Delta A = A - f(d_1, d_2, d_3, \dots) \approx \sum_i \frac{\partial f}{\partial d_i} \Delta d_i \quad (5)$$

Consider z which is a function of random variables x and y and deterministic variables k_1, k_2 :

$$z = k_1 x + k_2 y$$

$$\text{then } \bar{z} = k_1 \bar{x} + k_2 \bar{y} \quad \text{and} \quad \sigma_z = \sqrt{k_1^2 \sigma_x^2 + k_2^2 \sigma_y^2} \quad (6)$$

Applying relation (6) to (5), we get the mean value for A :

$$\bar{A} = \frac{\partial f}{\partial d_1} \bar{d}_1 + \frac{\partial f}{\partial d_2} \bar{d}_2 + \dots + \frac{\partial f}{\partial d_n} \bar{d}_n \quad (7)$$

$$\text{and Standard Deviation for } A: \quad \sigma_A = \sqrt{\left(\frac{\partial f}{\partial d_1} \right)^2 \sigma_{d_1}^2 + \left(\frac{\partial f}{\partial d_2} \right)^2 \sigma_{d_2}^2 + \dots + \left(\frac{\partial f}{\partial d_n} \right)^2 \sigma_{d_n}^2} \quad (8)$$

Suppose the tolerance T for each d_i is set at $n \cdot \sigma_i$ (typically we use 6σ), then we can substitute the values of σ_i by T/n in equation (8) to get the standard deviation for A , i.e., σ_A .

For worst case analysis eqn. (7) is also used to calculate the mean but tolerances are from the relation:

$$\Delta A = \frac{\partial f}{\partial d_1} \Delta d_1 + \frac{\partial f}{\partial d_2} \Delta d_2 + \dots + \frac{\partial f}{\partial d_n} \Delta d_n \quad (9)$$

The linearized approach uses arithmetic relations valid for Gaussian distribution. When linearization is not acceptable and/or contributing distributions are non-Gaussian, the only option is to use Monte Carlo simulation. In this method, random numbers are first generated and values of d_1, d_2, d_3 are picked from pre-assigned values in Tables which reflect the probabilities of each variable based on their respective frequency distribution. d_i values outside the tolerance range are discarded. Values of A are calculated from the equation(s) relating A to the independent variables, using values d_i generated at random. After a large number of A values are obtained, simulation is stopped and distribution of A is determined (tabulated or plotted to fit a desired distribution).

The above discussion did not specifically discuss how *geometric* tolerances are taken into account. In the next two sections, we look at how the common tolerance analysis methods, Min/Max Charts and Parametric Simulation, handle both *dimensional* and *geometric* tolerances.

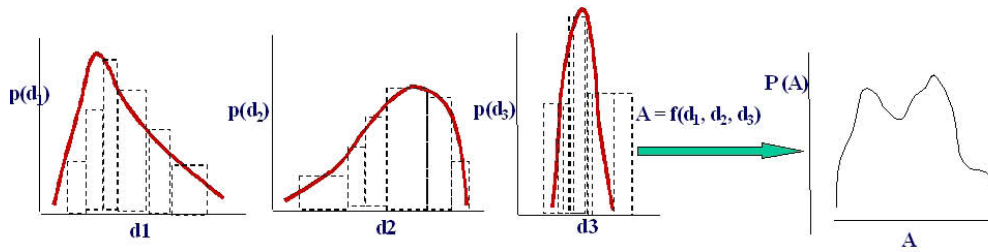
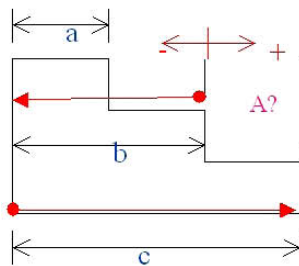


Fig. 7: Monte Carlo Simulation.

6. MIN/MAX TOLERANCE CHARTS

This is a manual bookkeeping procedure for 1-D stack calculation [19,30]. The analyst typically works with engineering drawings and interprets the Y14.5 symbols. A 1D coordinate system is set up with the origin at the left side of the unknown dimension as shown in Fig. 8 (or lower end for radial stack), with positive direction to the right and negative to the left. The rationale for this convention is that if A comes out to be positive it is a clearance and negative it is an interference (for assemblies); for part level analysis a negative value means that the feature disappears. A stack is a path from the origin to the other side of the analyzed dimension obtained by traversing a series of known dimensions. All tolerances encountered in traversing the stack are accounted for by rules that are specific to each class. The chart contains two main columns in which a value and a sign are entered for each tolerance contained in the stack, based on these rules. For size tolerances, the rule is that if the travel is in the positive direction, the maximum limit of the size is entered in the first column with a positive sign and its minimum in column 2 also with a positive sign. If travel is in the negative direction, the minimum value is put in column 1 and max in column 2, both with negative sign. The arithmetic sum of column 1 gives the max value of A and column 2 gives the min value of A. In the example shown, the stack is $-b+c$, so column 1 sum will be $(-b_{min}+c_{max})$ and column 2 will be $(-b_{max}+c_{min})$.



Part #	Stack Contributor	COLUMN 1		COLUMN 2		Δ
		sign	Maximum max/min	sign	Minimum min/max	
	TOTALS					

Fig. 8: Format of a traditional tolerance chart.

Space does not permit us to discuss rules for all geometric tolerance classes, so we will just give rules for one class, position tolerance, the most complex of all classes. Position tolerances are applied to Features-of-Size, FOS, (hole, boss, tab, slot, pocket) or feature patterns; it requires 2 or 3 datums and is interpreted as a zone within which the center, axis, or center plane of FOS is permitted to vary from its true. The nominal (true position) is specified by basic dimensions, shown by numbers enclosed in boxes in a drawing, as in Fig. 9 ; variations from true position are specified by position tolerance; a material modifier is required on the toleranced entity and any datums that are also FOS. When a position tolerance is specified at MMC (or LMC) and the actual FOS is not at MMC (or LMC), the position tolerance zone is enlarged by the amount of the difference between MMC (or LMC) and the actual size; this is called the bonus tolerance. When a FOS is used as a datum and MMC (or LMC) modifier is used and the actual datum FOS is not at MMC (or LMC), the tolerance zone is shifted by the amount of the difference. The charting rules for position are: The basic dimensions are entered in both columns with the same sign indicating direction of travel half the position tolerance is entered in each column; the left column has +ve and right -ve sign if bonus and shift are applicable, the procedure is as follows. If only the centerline of the feature-of-size is used in the stack: Enter max (radial) bonus in both columns and enter max shift in both columns (LMC radius minus Virtual Condition radius). If MMC & LMC radius of the feature-of-size are used in the stack: Enter zero bonus in the column containing the MMC radius and enter max (radial) bonus in the column containing the LMC radius. If MMC & LMC radius of the datum are used in the stack: Enter min shift in the column containing the MMC radius of the datum and enter max shift in the column containing the LMC radius of the datum. No shift is available for stacks between features in the same pattern

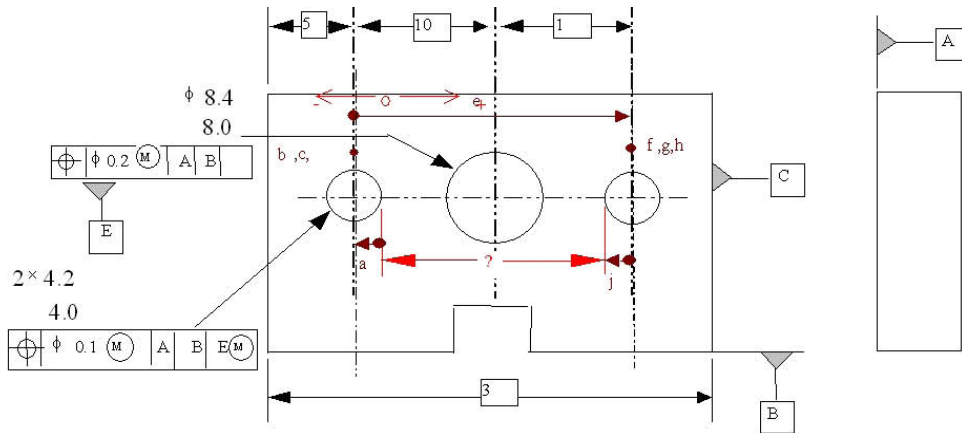


Fig. 9: Stack-up involving size and position tolerance.

6.1 Charts Case Study

Based on the rules for size and position tolerances, the chart is constructed for the analysis of the distance between the closest sides of two small holes (Fig 9). Starting at the left side, travel of radius is in -ve direction, so enter 2 and 2.1 in columns 1,2, both with negative signs. As per rule, half the position tolerance, 0.05 is entered in each column, the left column has +ve and right -ve sign. Since the MMC of the feature-of-size is used in the stack, a zero bonus is entered in the column containing the MMC radius and the max (radial) bonus 0.1 entered in the column containing the LMC radius. Because the holes are a pattern, no shift tolerance applies. The basic dimension of 20 is entered in both columns with a + because the travel is in the positive direction. The second hole is handled in the same way as the first. Adding each column we get the worst case as 16.1 max and 15.5 min. The third column (Δ) is just for checking; it is the difference between columns 1 and 2 for each row. Its sum should be the same as the difference between the max and min columns (0.6 in this case).

Stack Contributor	sign	Maximum max/min	sign	Minimum max/min	Δ
a (radius)	-	2.0	-	2.1	0.1
b (position tol)	+	0.05	-	0.05	0.10
c (bonus tol)	+	0	-	0.1	0.1
d (shift)	+	0	-	0	0
e (basic size)	+	20	+	20	0
f (position tol)	+	0.05	-	0.05	0.1
g (bonus)	+	0	-	0.1	0.1
h (shift)	+	0	-	0	0
j (radius)	-	2.0	-	2.1	0.1
	+	16.1	+	15.5	0.6

Tab. 2: Tolerance chart for part shown in Fig. 9.

Charts can also be constructed for worst case analysis of clearances in assemblies. The analyst mentally “positions” parts in a way that gives the worst cases (min. or max. value of analyzed dimension). So separate charts have to be done for each worst case. Since no algebraic expression for the analyzed dimension in terms of the contributors is generated by this method, no statistical analysis can be done. Also, since some contributors are not aligned with the direction of analysis, they are ignored, which may yield incorrect results.

7. COMMERCIAL CATS SOFTWARE

Current major CATS packages include *UGS/VisVSA*[13], *Tecnomatix/ eTol-Mate* [9], *Cognition /Mechanical Advantage* [10], *3-DCS*[16], and *Sigmatix/CETOL*[29]. Some other CAD systems have an integrated module for tolerance analysis. In the parametric approach, the analyzed dimension is expressed as an algebraic function (an equation, or a set of equations) that relates it to the contributors expressed as geometric parameters. The function is either linearized or directly used for Monte Carlo simulation in nonlinear analysis. Both dimensional tolerances and geometric tolerance classes (most) can be included in the analysis. Results commonly available are the lists of contributors, sensitivities, % contributions (to variance), and tolerance accumulation for worst-case and statistical analysis. CATS systems vary in how they interface with CAD and the type of analysis provided, each of which is discussed below.

7.1 CATS-CAD Interface

Modern CAD systems are history based; they allow solid construction by a combination of 2D constrained sketches and 3D operations like sweeps and lofts. Editing (changes to dimensions) typically requires use of history to rollback to the point of change, re-solving the constraints, and roll-forward to the top [11]. Thus, there is no single 3D constraint network, making it costly to carry out perturbation of variables to calculate sensitivities for linearized analysis, and impractical for repeated simulations for non-linear Monte Carlo calculations. Therefore, integrated CATS modules that use the CAD model directly are not common, and the ones that exist are very limited. The limitation may be of two kinds: either the analysis is limited to 2D cross-sections to allow direct solving of constraint equations (e.g., *UGS/IDEAS*[17]), or only linearized analysis is supported (*MechAdvantage*[10]).

The more mainstream approach adopted is to import the geometry from CAD into CATS and have the user interactively create an abstracted “feature model” for each analyzed dimension. As things stand today, GD&T information is not imported because STEP translators do not yet incorporate the GD&T portion of STEP. So GD&T is lost in importing and the tolerance analyst must redefine all needed tolerances in CATS for the particular analysis that is desired. Abstraction of the geometry and selected dimensions is done in terms of points, lines, planes and distances, angles between them. For example, a cylindrical feature will be represented by “point + vector (axis) + radius + length”, a planar feature will be represented by “point + vector (normal) + width + length”. These abstract objects can be created independently or from actual CAD geometry. Then the dimension of interest to be analyzed is defined in terms of Point Coordinate, Point-to-Point, Point-to-Line, Point-to-Plane, Gap/Flush, Angle, Maximum or Minimum Virtual Clearance. Figure 10 shows an example of geometric entities extracted from the imported CAD model for analyzing the position of the hole (the GD&T frames were added interactively by the user).

Creating the abstract feature model requires considerable skill and experience to get valid results. For a specific dimension of interest, it is not obvious which type of measurement to choose of all the types of measurement options. For example, if one wants to analyze the distance between two parallel planar features, then one can choose point-to-point or point-to-plane measurement. Some packages do not support plane-to-plane measurement, so a way around is to pick 3 points on the target plane (at the farthest points on the plane to maximize accuracy) and take three point-to-plane distances to the other (datum) plane. The results are sensitive to the location of the points on the target plane.

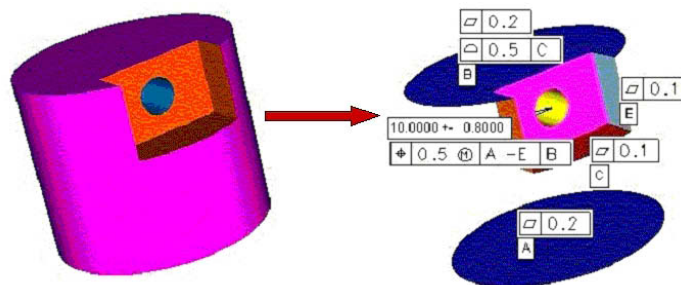


Fig. 10: CAD model used to extract abstract feature model (validity depends on the user).

7.2 Solution Procedure

The first step in any type of analysis is to use the constraint network of the abstracted model to extract the dimension loop (stack of all contributors to the accumulation of interest). The rest of the procedure varies with Linearized or Monte Carlo analysis.

Linearized analysis: This can give both worst case and statistical results for 2D or simplified 3D. The dimension of interest A can be calculated from constraint equations and values of the contributors. If one contributing variable d_i is perturbed at a time, its sensitivity $\partial A / \partial d_i$ can be obtained as the ratio of the change in the analyzed dimension value to the change in the perturbed value. By substituting the mean values of d_i and sensitivities $\partial A / \partial d_i$ in the linearized eqn (7), the mean value of A can be found. By putting in half the tolerance Δd_i (assuming equal bilateral) in eqn (9) the worst case is found, $A = A \pm \Delta A$. Assuming that all contributors are normally distributed, and the tolerance Δd_i for each d_i is set at $n^* \sigma_i$, we can substitute $\sigma_i = T/n$ in equation (8) to obtain the variance for A , which is also assumed to be normally distributed. By superimposing the upper and lower acceptable limits for A on this distribution, as in Fig. 6, the acceptance rate can be found from statistical Tables. The %contribution of each d_i to the variance of A is

$$\% \text{ contribution for } d_i = \frac{(\partial f / \partial d_i)^2 \sigma_i^2}{\sigma_A^2} \times 100\% \quad (10)$$

If Linearized analysis is done in 2D it cannot account for out of plane contributors and several types of tolerance classes, such as runout or surface profile. Neither 2D nor 3D analysis based purely on geometric parameters and not tolerance zones, can account for datum precedence, Rule 1, material conditions, bonus and shift, which are important parts of the standards. Form tolerances are not considered at all in 2D or 3D. Since tolerances in mechanical assemblies are typically two orders of magnitude smaller than dimensions, linearization itself does not cause any significant errors.

Monte Carlo Analysis: CATS packages based on this method allow the contributing dimensions to have any statistical distribution. It also handles both linear and nonlinear response functions, since the values of the response function are computed by simulation. Monte Carlo simulation involves generating random values for each contributor in accordance with its probability distribution, checking to see if the values are within limits of the tolerance on that feature, and then computing the value of A for each simulation using the parametric relations of the abstracted model. This amounts to point-based analysis but the standards are zone based. The main drawback of this method is that to get accurate estimates, it is necessary to generate very large samples (tens of thousands of simulations) and this is computationally intensive. Unlike the linearized analysis, there are no closed form equations for computing worst-case results; these can be estimated by looking at the extreme values generated in all the simulations combined, but there is no guarantee that the very next simulation produces a value outside that range.

Bonus and shift tolerances can be handled properly. Whenever a Feature-of-Size (FOS) is located with a tolerance of position, the effects of the position tolerance must be included in the stack. If the tolerance and/or the datum are specified with material modifiers, then there exist bonus and/or shift tolerances that must be considered. One possibility is to first generate a random size value within corresponding size limits for each of the FOS involved then calculate the positional tolerance zone for each FOS involved. Next, if the datum specified is modified at MMC or LMC, then a random location of datum feature can be created first inside its "calculated" position zone and then placed at the "true position" from this instance of datum feature. And then a random location for target feature is generated inside this placed zone. Finally, the instance dimension value for the dimension under analysis can be calculated. These steps are repeated for all the rounds of simulations. We have also determined that form tolerances and datum precedence are not supported in leading commercial packages at this time [9,25].

8. EMERGING TECHNOLOGIES

From the foregoing one can see that neither charting methods nor parametric CATS are entirely satisfactory. The key missing ingredient is a fundamental understanding of geometric variations and how they interact and accumulate, i.e., a mathematical model is needed. Many such models have been proposed [8,15,20,22,29] but they lack compatibility with GD&T standards. One promising approach is the modeling of tolerance zones as multi-variate regions [14,21,23,31]. Of these emerging models, T-Maps are the most developed, so we will review them briefly here.

8.1 T-Maps Basics

A T-Map[®] is a hypothetical Euclidean point-space, the size and shape of which reflects all variational possibilities for a toleranced feature (target). It is the range of points resulting from a one-to-one mapping from all the variational possibilities of a feature, within its tolerance-zone, to the Euclidean point-space [12]. These variations are determined by the tolerances that are specified for controlling size, position, orientation etc., of the feature. The T-Map[®] for any combination of tolerances on a feature is constructed from a basis-simplex and described with areal coordinates. If the mapping is done for n-types of variation of a feature, a T-Map[®] will be created of n-dimension; the basis-simplex will also be of n-dimensions. To construct a n-dimensional simplex, n+1 basis points are needed. Consider a cross-section

of a round bar with a size tolerance t on its length, Fig. 11.a. The variational possibilities are represented by region ABCD. The upper and lower extreme positions, AB and CD can be used to map to basis points σ_1, σ_2 and the extreme angle CB to a third basis point σ_3 . Now every line in ABCD can be represented by a corresponding point in the basis triangle $\sigma_1, \sigma_2, \sigma_3$ (Fig 11.b). To represent max rotation in the opposite direction we have a complementary point on the other side (Fig 11.c). To represent all cross-sections, this double triangle can be rotated to get a 3D figure called the T-map for size of the round bar (Fig 11d). Any end plane of the round that satisfies the size tolerance will be represented by σ , uniquely determined by the linear combination

$$\sigma = \lambda_1\sigma_1 + \lambda_2\sigma_2 + \lambda_3\sigma_3 \quad (11)$$

The three λ_i can be normalized by setting $\lambda_1 + \lambda_2 + \lambda_3 = 1$. It is only necessary for affine geometry that the positions for basis-points to be independent.

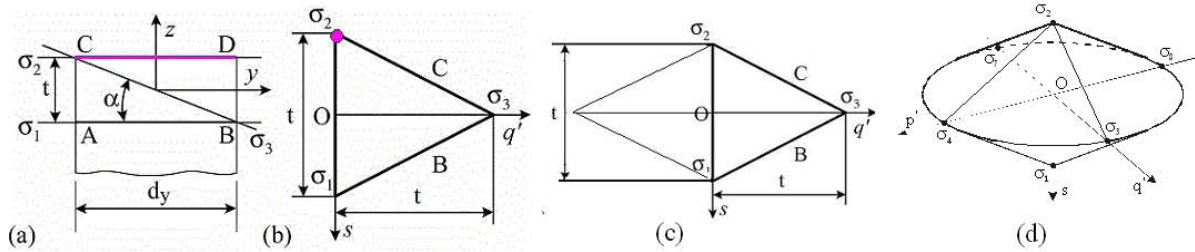


Fig. 11: Generation of T-map for size tolerance of a round bar.

The dimensions of this T-Map® depend on the tolerances and the dimensions of the tolerated plane. Rotational variations of the plane are represented along σ_1 - σ_3 , or σ_1 - σ_3 , direction translational variations of the plane are represented along σ_1 - σ_2 of the T-Map®. Planes σ_i in the tolerance zone correspond to points σ_i in the T-Map®. Thus this T-Map® represents all 3D variations of the tolerated plane in Fig 11(d). If an orientation tolerance (e.g., parallelism) is specified on the end plane, its effect will be restrict the rotational direction, i.e. the T-map will be truncated along σ_3 which gives the map shown in Fig for combined size and orientation.

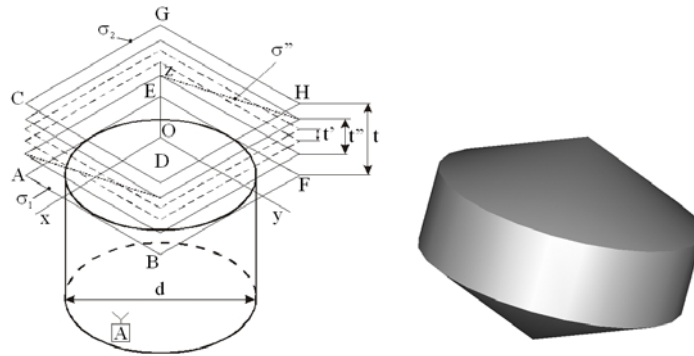


Fig. 12: Modification of size T-map to include orientation tolerance.

T-Maps have been developed for other features, such as for a polygonal bars [21], cylindrical features [7], axes [6], angled faces [1], point-line clusters [29] etc, which represent 3D variations of each. These T-Maps are of different dimensions, depending on the tolerated feature and the tolerance specification. Detailed discussion of different T-Maps is not within the scope of this paper. Next we discuss T-Maps based tolerance analysis.

8.2 T-Maps Based Tolerance Analysis

When two parts are assembled, the errors accumulate. For example, one round bar sitting on top of another round bar (Fig 13.a). This means we must look at all combinations of translation and rotations of the end planes, which is like sweeping one T-map about the other (Fig 13b). Mathematically, this translates to a Minkowski sum of T-maps of all

contributors. In the case of planes, which have 3D T-maps, the Minkowski sum is also a 3D object (Fig 3c). T-maps generated by Minkowski sums of contributing maps are called Accumulation Maps. The limiting values of the analyzed dimensions (target feature) can be used directly to get what is termed the functional map. For 100% inter-changeability (worst case), the accumulation map must fit inside the functional map. Conversely, we can generate the largest possible functional map (circumscribing functional map) from the accumulation map to get the worst case result.

For statistical analysis, the multi-parameter variations are mapped into one-parameter variation of the dimension of interest of the assembly to form the relative frequency distribution for the dimension. The steps are as follows: 1) Get the T-Map for the target feature on the part or assembly, 2) Get the circumscribing functional map for the target feature, 3) Find the intersection of the accumulation map with that surface S of the functional map corresponding to the smallest value of clearance; the resulting value is the first point of the frequency distribution, 4) Obtain each additional point of the distribution as the area obtained by intersecting the Map with surface S that has been translated inward a calculated amount. In general, we need numerical convolution to combine frequency distributions for individual features. Details can be found in [3].

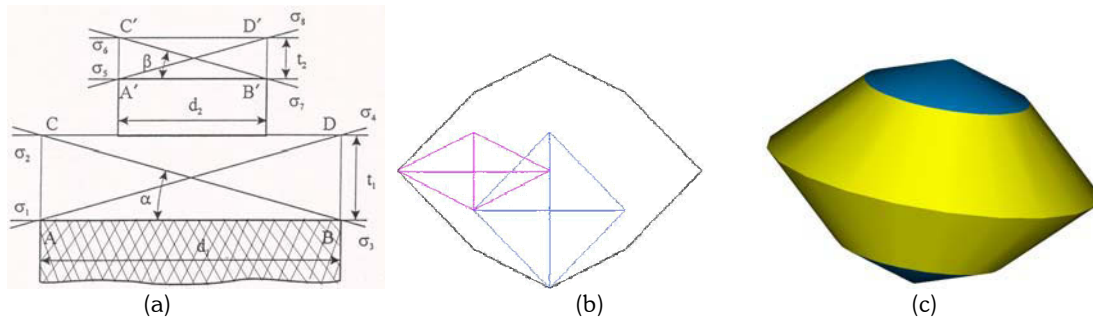


Fig. 13: Accumulation T-maps as Minkowski sums.

8.3. T-maps Case study

Figure 14 shows two parts that need to be assembled (tab-slot) with A as the common datum (mating). The size and position tolerances are inter-related by the modifier \textcircled{M} . Using T-maps, it is possible to do statistical analysis with any material condition: MMC or RFS in this case. The relative frequency distribution of clearance between the engaged tab and slot for both are compared to simulation based analysis in Fig. 15 and the acceptance rates given in Table 3.. All part variations are assumed to be normally distributed within the tolerance zone.

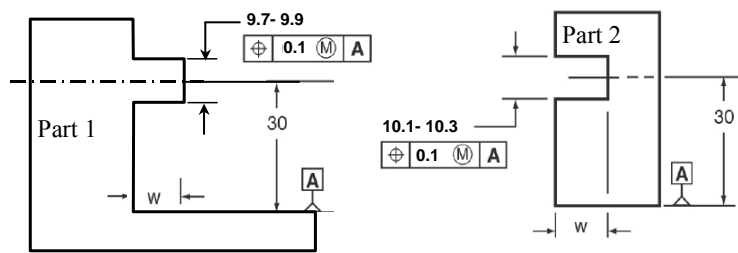


Fig. 14: Assembly case study.

10. DISCUSSION

The min/max tolerance charting method is consistent with the Y14.5 standard; it can take into account both dimensional and geometric tolerances, and is applicable at both part and assembly level. Because it is relatively easy to understand and use, tolerance charting has been widely used in the draftsmen community. However, it is a manual procedure that is hard to automate, the user must remember all the rules for all the classes while constructing the chart, making the process tedious and prone to errors. Charts analyze accumulation in one direction at a time, ignoring possible contributions from other directions, which often leads to inaccurate results, and only worst-case analysis can be done. Shen *et al.* [24,26] has made some progress towards automating min/max charts. Tolerance charting,

especially automated charting method can offer a quick tolerance check and relatively accurate results when the contributions in the other directions are at the negligible level. This method is especially useful for a quick check of an initial tolerance allocation. When the part features are offset from the common analysis direction or axis, or angle dimensions are involved, charting method cannot be used.

Parametric tolerance analysis integrated with CAD models may be limited to 2D and usually uses linearized approach. Model parameters and constraints used in creating a CAD model do not correspond directly to the tolerances that may be specified on the drawings. This approach is unable to accurately represent datum precedence, material modifier condition, form tolerance and other features of standards. Only Gaussian distribution is supported.

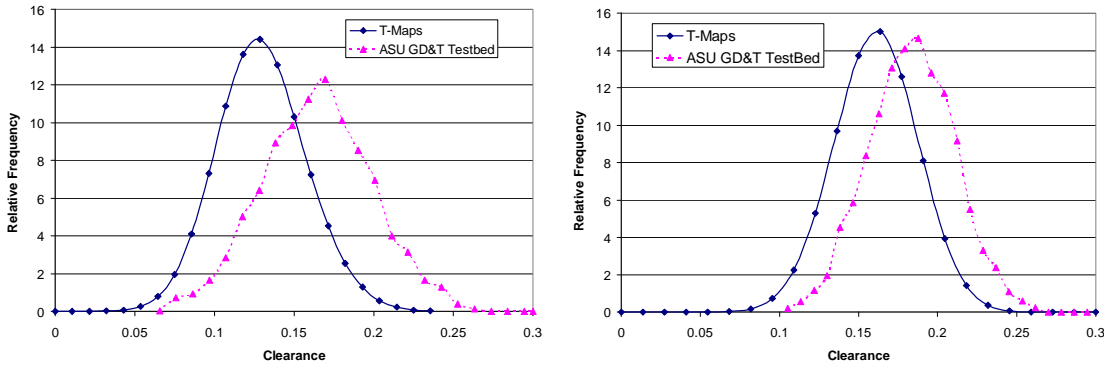


Fig. 15: Statistical distribution of tab-slot clearance for (a) MMC (b) RFS.

	Material Condition	Acceptable Clearance	% Acceptance
T-Maps	MMC	0-0.15	82.3%
	RFS	0-0.15	42.4%
ASU GD&T Testbed	MMC	0-0.15	49.5%
	RFS	0-0.15	18.5%

Tab. 3: Percent acceptance rates with RFS and MMC conditions.

Mainstream CATS use abstracted feature-parameter model for Monte Carlo simulation. The major problems with these packages are that the results depend on the expertise of the user, not just the GD&T specifications, often requiring companies to hire consultants. Monte Carlo simulation does not produce a closed form solution; the solution changes with the number of simulations performed; one can never be sure of the worst case results. Although many Y14.5 rules can be incorporated, such as material conditions, many cannot, such as datum precedence. Commercial CATS are basically point based analysis, which does not conform to true 3D tolerance zones and geometric variations.

T-Maps are based on traditional principles of geometry, such as areal coordinates and screws and they do not depend on perturbations of just the points that are used to create a solid model. Compared with other methods, it has the following strengths: (1) it can model all the 3D variations of a feature, such as size, form, orientation and position, consistently with ASME Y14.5 Standard; (2) it models completely the interactions of these variations; (3) it incorporates Rule #1, floating zones, bonus/shift tolerances, datum precedence, and material modifiers; (4) it can easily demonstrate accumulation of various part tolerances in an assembly; and (5) it provides multiple stack up equations and metric measures to aid a designer in selecting optimal tolerances. The analysis model is independent of the user's choices, i.e. the results from this method will be the same for the same problem no matter how the user creates the analysis model. The method is not yet fully developed and the math concepts may be difficult for most designers to understand. Visualization of higher-dimensional T-Maps is not straightforward. The challenge to create a user-interface that makes a good working platform for tolerance analysis, yet hides most of the geometric intricacies of the method.

11. ACKNOWLEDGEMENTS

This study was supported in part by National Science Foundation Grants DMI-9821008 and DMI-0245422. The opinions expressed in the paper are those of the authors and not endorsed by the NSF. The authors wish to thank NSF for the financial support.

12. REFERENCES

- [1] Ameta, G.; Davidson, J. K.; Shah, J. J.: The Effects of Different Specifications on the Tolerance-Maps for an Angled Face, 30th Design Automation Conf., Salt Lake City, UT, Sept. 28-Oct. 2, 2004.
- [2] Ameta, G.; Davidson, J. K.; Shah, J. J.: Tolerance-Maps Applied to a Point-Line Cluster of Features, CD-ROM Proceedings, ASME Design Engr. Technical Conf.s. (31st Design Automation Conf.), Long Beach, CA, Sept. 25-28, 2005, Paper #IDETC2005/DAC-85115, 11pp.
- [3] Ameta, G.; Davidson, J. K.; Shah, J. J.: Using Tolerance-Maps to Generate Frequency Distributions of Clearance for Pin-Hole Assemblies, ASME Design Automation Conf, Philadelphia, Sept. 2006.
- [4] ASME Standard, 1994, Dimensioning and Tolerancing, ASME Y14.5M-1994, American Society of Mechanical Engineers, New York
- [5] ASME Y14.5.1M-1994, Mathematical Definition of Dimensioning and Tolerancing Principles, An ASME National Standard, ASME, New York, N.Y.
- [6] Bhide, S.; Ameta, G.; Davidson, J. K.; Shah, J. J.: Tolerance-Maps Applied to the Straightness and Orientation of an Axis, Proceedings of 9th CIRP 2005 Seminar on Computer Aided Tolerancing.
- [7] Bhide, S.; Davidson, J. K.; Shah, J. J.: A new mathematical model for geometric tolerances as applied to axes, CD-ROM Proceedings, ASME Design Engr. Technical Conf.s. (29th Design Automation Conf.), Chicago, IL, Sept. 2-6, 2003, Paper #DETC2003/DAC-48736, 9.
- [8] Chase, K.; Gao, J.; Magleby, S.; Sorensen, C.: Including geometric feature variations in tolerance analysis of mechanical assemblies. IIE Transactions, 28, 1996, 795-807.
- [9] Chiesi, F.; Governi, L.: Software Review-Tolerance analysis with eTol-Mate, ASME Transactions on the J. of Computing & Information Science in Eng., special issue on GD&T, 3(1), 2003, 100-105.
- [10] Cognition, Inc., Mechanical Advantage User Solutions Training Manual, Cognition, Inc., Boston, MA.
- [11] Corney, J. ; Lim, T. : 3D modeling with ACIS , Saxe-Coburg, 2nd edition, 2001.
- [12] Davidson, J. K.; Mujezinović, A.; Shah, J. J.: A new mathematical model for geometric tolerances as applied to round faces, ASME Transactions, JMD, 124, 2002, 609-621.
- [13] EAI Corp., 2000, Vis VSA Manual, Ames, IA.
- [14] Giordano, M.; Pairel, E.; Samper, S.: Mathematical Representation of Tolerance Zones,' In Global Consistency of Tolerances, Proc., 6th CIRP Int'l Seminar on CAT, Univ. of Twente, Enschede, Netherlands, March 22-24, F. van Houten and H. Kals, eds., 1999, 177-186.
- [15] Hillyard, R.; Braid, I.: Analysis of Dimensions and Tolerances in Computer Aided Mechanical Design, Computer-Aided Design, 10(3), 1978, 161-166.
- [16] <http://www.3dcs.com/>
- [17] <http://www.ugs.com/products/nx/ideas/>
- [18] ISO, International Organization for Standardization ISO 1101: Geometrical tolerancing - Tolerancing of form, orientation, location and run-out - Generalities, definitions, symbols, indications on drawings, 1983.
- [19] Krulikowski, A.: Tolerance Stacks - A Self-Study Course, Effective Training Inc., Westland, MI, USA, 2002.
- [20] Light, R.; Gossard, D.: Modification of Geometric Models Through Variational Geometry, Computer-Aided Design, 14(4), 1982, 209-214.
- [21] Mujezinović, A.; Davidson, J. K.; Shah, J. J.: 2004, A New Mathematical Model for Geometric Tolerances as Applied to Polygonal Faces, ASME Transactions, JMD, 126(3), 2004, 504-518.
- [22] Requicha, A. A. G.: Representation of Tolerances in Solid Modeling: Issues and Alternative Approaches. In Solid Modeling by computers: from theory to applications, eds. Boyse, J. W. and Pickett, M. S., Plenum Press, New York, 1984, 3-22.
- [23] Roy, U.; Li, B.: Representation and Interpretation of Geometric Tolerances for Polyhedral Objects: II. Size, Orientation and Position Tolerances, Computer-Aided Design, 31(4), 1999, 273-285.
- [24] Shen, Z.; Shah, J. J.; Davidson, J. K.: Automation of linear tolerance charts and extension to statistical tolerance analysis, CD-ROM Proceedings, 23rd ASME Computers & Information in Engr. Conf., Chicago, IL, Sept. 2-6, 2003, Paper #DETC2003/CIE-48179.
- [25] Shen, Z.: Software Review-Tolerance analysis with EDS/VisVSA, ASME Transactions on the Journal of Computing and Information Science in Engineering, special issue on GD&T, 3(1), 2003, 95-99.
- [26] Shen, Z.; Shah, J. J.; Davidson, J. K.: Simulation-Based Tolerance and Assemblability Analyses of Assemblies with Multiple Pin-Hole Mating Conditions, ASME CIE conf, Long Beach, California, Sep. 2005.
- [27] Sigmetrix: <http://www.sigmetrix.com>.

- [28] Teissandier, D.; Delos, V.; Couetard, Y.: Operations on Polytopes: Application to Tolerance Analysis, In Global Consistency of Tolerances, Proc., 6th CIRP Int'l Seminar on CAT, Univ. of Twente, Enschede, Netherlands, March 22–24, F. vanHouten and H. Kals, eds, 1999, 425–434.
- [29] Turner, J. U.; Wozny, M. J.: The M-space theory of tolerances, In B. Ravani, ed., Proc. of 16th ASME Design Automation Conf., ASME Press, 1990, 217–225.
- [30] Wade, O. R.: 1967, Tolerance Control in Design and Manufacturing, Industrial Press, New York, 1967.
- [31] Whitney, D. E.; Gilbert, O. L.; Jastrzebski, M.: Representation of Geometric Variations Using Matrix Transforms for Statistical Tolerance Analysis in Assemblies, Res. in Eng. Design, 6, 1994, 191–210.