Chapter

7

Mathematical Definition of Dimensioning and Tolerancing Principles

Mark A. Nasson
Draper Laboratory
Cambridge, Massachusetts

Mr. Nasson is a principal staff engineer at Draper Laboratory and has twenty years of experience in precision metrology, dimensioning and tolerancing, and quality management. Since 1989, he has been a member of various ASME subcommittees pertaining to dimensioning, tolerancing, and metrology, and presently serves as chairman of the ASME Y14.5.1 subcommittee on mathematical definition of dimensioning and tolerancing principles. Mr. Nasson is also a member of the US Technical Advisory Group (TAG) to ISO Technical Committee (TC) 213 on Geometric Product Specification. Mr. Nasson is an ASQ certified quality manager.

7.1 Introduction

This chapter describes a relatively new item on the dimensioning and tolerancing standards scene: mathematically based definitions of geometric tolerances. You will learn how and why such definitions came to be, how to apply them, what they have accomplished for us, and where these definitions may take us in the not-too-distant future.

7.2 Why Mathematical Tolerance Definitions?

After reading this chapter, I hope and trust that you will be asking the reverse question: Why *not* mathematical definitions of tolerances? As you will see, a number of interesting events combined to open the door for their creation. In short, though, mechanical tolerancing is a much more complex discipline than most people realize, and it requires a similar level of treatment as has proven to be necessary for the nominal geometric design discipline (CAD/solid modeling).

Although the seeds for mathematical tolerance definitions were planted well before the early 1980s, a special event of that era indirectly helped trigger a realization of their need. The arrival of the personal computer quite suddenly and dramatically decreased the cost of computing power. As a result, vendors of metrology equipment, predominantly coordinate measuring machines (CMMs) began offering affordably priced measurement systems with integrated personal computers. Also, a number of individuals developed homegrown systems for their companies (as did this author) by pairing an older measuring system that they already owned with a newly purchased personal computer. Just as personal computers have affected us in countless other ways, they also contributed to the resurgence of the coordinate measuring machine.

Another device also contributed to the resurgence of coordinate measuring machines: the touch trigger probe, originally developed in the U.K. by Renishaw. Prior to this invention, conventional coordinate measuring machines used a “hard” probe (a steel sphere) for establishing contact with part features. Not only were hard probes slow to use, but they also were capable of disturbing the part, and even damaging it if the inspector failed to exercise sufficient care. Touch probes improved this state of affairs by enabling the coordinate measuring machine to significantly overtravel after the part feature was triggered upon initial contact. Productivity and accuracy were both improved with touch probes.

The advent of touch probe technology and the availability of relatively inexpensive computing power through new microprocessors enabled quick and sophisticated collection, processing, and display of measurement data. That was the good news of the early 1980s. The bad news? The many instances of software applications developed for metrology equipment did not interpret geometric dimensioning and tolerancing uniformly. Although the personal computer helped us recognize a number of underlying problems with tolerancing and metrology (and hence, for much of manufacturing), other key events helped us further diagnose problems and even chart out plans for resolving them. Writing and using mathematical tolerance definitions were among the suggested corrective actions.

7.2.1 Metrology Crisis (The GIDEP Alert)

In September of 1988, Mr. Richard Walker of Westinghouse Corp. issued a GIDEP Alert[1] against the data reduction software from five unnamed CMM vendors. Himself aware of inconsistency problems with CMM software for some time through painful experience, Mr. Walker sought to bring this serious state of affairs to public light by issuing the GIDEP Alert. Typically, GIDEP issues alerts against specific manufacturer’s product lines or production lots with quality concerns. In this case, the problem was not attributable to just one CMM vendor; this was an industry-wide problem and was not confined to the metrology industry. It was a serious symptom of a larger problem. First, though, let’s deal with the subject of the GIDEP Alert.

Ideally, and not unreasonably, we expect that a measurement process for a given part (say flatness as measured by a CMM) will yield repeatable results. The degree of repeatability depends on many factors such as the number of points sampled, point sampling strategy, stability of the part, and probing force. Each of these factors comes into play on measurements performed on a single, given CMM.

[1] GIDEP (Government-Industry Data Exchange Program, http://www.gidep.corona.navy.mil) is an organization of government and industry participants who share technical information with each other regarding product research, design, development, and production. One function of GIDEP is to issue alerts to its members that pertain to nonconforming parts, processes, etc. In this case, the subjects of the alert were nonconforming software algorithms.

But what about the repeatability of measurements of the same part as performed by CMMs from different manufacturers? Potential contributors to repeatability in this context are the differences in mechanical stability between the CMMs *and the software algorithms used to process the sampled point coordinate data*. It's the latter with which Mr. Walker's GIDEP Alert dealt. Suspicious of inconsistencies between measurement results obtained by different CMMs, Mr. Walker crafted ingeniously simple, but strategically chosen sets of point coordinate data to test the performance of CMM software algorithms for calculating measured values of flatness, parallelism, straightness, and perpendicularity. A data set that could be solved graphically without any algorithms was strategically selected. So not only did Mr. Walker check for consistency between the five CMMs tested, but he also checked for *correctness*.

The results were rather shocking. The worst offending algorithm in one case reported results that were 37% worse than the actual results; in other words, the algorithm indicated that the part feature was worse than it actually was. In another case, the worst offending algorithm reported results that were 50% better than the actual results, indicating that the part feature was better than it actually was. These results led to the realization that many CMM software algorithms were unreliable. Coupling this fact with an increasingly wide awareness that different measurement techniques applied to the same parameter yielded different results, a true metrology crisis was in effect.

In true Ralph Nader spirit, Mr. Walker acted on behalf of the customers of metrology equipment vendors. Rather than letting the potential impact on the CMM vendors determine how he handled this discovery, he publicized this information to educate and warn CMM users and the customers of their results. He resisted those that preferred him to keep silent while these problems were solved behind closed doors. Instead, the GIDEP Alert served as a beacon to those who experienced similar problems and had the motivation and technical ability to do something about it. Mr. Walker was criticized by many for his actions—a sure sign that he was on to something.

7.2.2 Specification Crisis

The GIDEP Alert convincingly illustrated the unstable situation with metrology software. However, it is crucial to recognize that the metrology crisis was actually a symptom of the true problem. The inherent ambiguity in the text-based definitions of mechanical tolerances enabled the writing of varied and incorrect computer algorithms for processing inspection data. Though text-based definitions seem to have served engineering well for many years, the robustness and rigor required by computerization has revealed a number of underlying problems. Without the ability to unambiguously specify and assign tolerance controls to mechanical parts, we cannot expect to be able to uniformly verify the adherence of actual parts to those specifications. Thus, one could accurately say that the specification crisis spawned the metrology crisis.

7.2.3 National Science Foundation Tolerancing Workshop

Under a grant from the National Science Foundation, the ASME Board on Research and Development conducted a workshop with invited guests of varied manufacturing backgrounds from a number of domestic and international companies. Held soon after release of the GIDEP Alert, this workshop sought to identify research opportunities in the field of tolerancing of mechanical parts. These research opportunities were determined on the basis of unsolved problems or technological gaps hampering the effectiveness of various engineering disciplines. Among the recommendations generated by the workshop was that mathematically based definitions of mechanical tolerances should be written in order to remove ambiguities and reduce misuse. This recommendation paved the way for the establishment of a body whose sole purpose was to meet that goal.

7.2.4 A New National Standard

In January of 1989 the Y14.5.1 “ad hoc” subcommittee on mathematization of geometric tolerances held its inaugural meeting in Longboat Key, Florida. In approximately fifteen meetings held over five years’ time, Chairman Richard Walker led an inspired group of volunteers to the publication of a new national standard, ASME Y14.5.1M-1994 - Mathematical Definition of Dimensioning and Tolerancing Principles. The continually surprising degree of effort that was necessary to write this document provided constant confirmation that the document was truly needed. Some ambiguities were known before mathematization efforts began, but many other subtle problems were revealed as the subcommittee members took on the challenge of unequivocally specifying what was previously conveyed through written word and figures drawn from specific examples.

7.3 What Are Mathematical Tolerance Definitions?

7.3.1 Parallel, Equivalent, Unambiguous Expression

Mathematical tolerance definitions are a reiteration of the tolerance definitions that appear in textual form in the Y14.5 standard. In many cases, actual mathematical expressions describe geometric constraints on regions of points in space yielding a mathematical/geometrical description of the tolerance zone for each tolerance type. However, tolerance types are only part of the story. The Y14.5.1 standard handles the crucial subject of datum reference frame construction not with mathematical equations, but with mathematical formulations that are expressed textually with supporting tables and logical expressions. In any case, the contents of the Y14.5.1 standard have a direct tracing to an unambiguous mathematical basis. The unfortunate tradeoff is that they are not readily assimilated by human beings, but they are easily converted into programming code.

7.3.2 Metrology Independent

The developers of the Y14.5.1 mathematical standard diligently maintained at arm’s length (or farther!) any influences from current measurement techniques and technology on the mathematical tolerance definitions. There was a frequent tendency to think in terms of inspection procedures when trying to mathematically describe some characteristic of a geometric tolerance, but it was resisted. Measurability was never a criterion that prevailed during the deliberations of the Y14.5.1 subcommittee. The reason was simple: tolerancing is a design function, and it must not be encumbered by metrology, a downstream activity in the product life cycle. Today’s state-of-the-art in measurement technology eventually becomes yesterday’s obsolescence. Desired features and capabilities for dimensioning and tolerancing that enable precise specification of part functionality and producibility should drive technology development in metrology. To have specified mathematical tolerance definitions in terms of industry-accepted measurement techniques would surely have made the definitions more recognizable, but generality would have been sacrificed.

7.4 Detailed Descriptions of Mathematical Tolerance Definitions

7.4.1 Introduction

This section contains introductory material necessary to read and understand mathematical tolerance definitions as they appear in the Y14.5.1 standard. Those readers with a physics and/or mathematics background may bypass the section on vectors that follows. Section 7.4.3 presents some key terms and concepts specific to the Y14.5.1 standard. The remaining sections cover a selection of actual mathematical tolerance definitions. Note that not all aspects of the Y14.5.1 standard are covered here, and that this chapter is designed to provide the reader with enough background to enable him/her to make effective use of the standard.

7.4.2 Vectors

This section contains a brief overview of vectors and the manner in which they are handled in mathematical expressions. Those readers with a physics and/or mathematics background will not find it necessary to read further. The material is included, however, because not all users of geometric dimensioning and tolerancing have had exposure to it, and it is the basis of the definitions that follow.

Vectors are abstract geometric entities that describe direction and magnitude (length). A position vector can describe every point in space, which is simply a line drawn from the origin to the point. Vectors also exist between points in space. The magnitude of a vector is its length as measured from its starting point to its end point. A vector of arbitrary length is typically designated by a letter with an arrow ($\vec{A}$) over it. Graphically, vectors are shown as a line with an arrow at one end; the length of the line represents the vector's magnitude, while the arrow represents its direction. See Fig. 7-1.

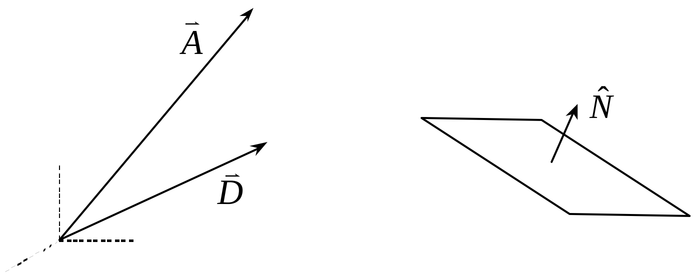

Figure 7-1 Vectors and unit vectors

A special type of vector is the *unit* vector which, not surprisingly, is of unit length. Unit vectors are often used to define or specify the direction of an axis or the direction of a plane's normal; a unit vector is appropriate for such purposes because it is the direction and not the magnitude that is important. A unit vector is typically designated by a letter with a hat, or carat, ($\hat{T}$) over it.

7.4.2.1 Vector Addition and Subtraction

Vectors may be added and subtracted to create other vectors. Two vectors are added by overlapping the starting point of one vector on the end point of the other vector. The resultant vector, or sum vector, is that vector that extends from the starting point of the first vector to the end point of the second vector. See Fig. 7-2.

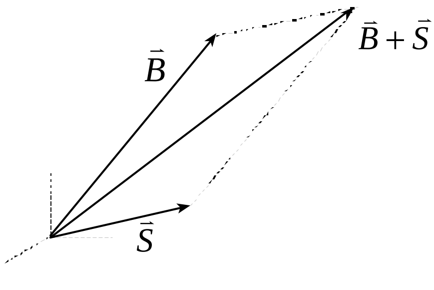

Figure 7-2 Vector addition

Vector subtraction is performed analogously. In Fig. 7-3, the vector $\vec{C}-\vec{R}$ is obtained by adding the negative of vector $\vec{R}$ (which simply points in the opposite direction as $\vec{R}$) to vector $\vec{C}$.

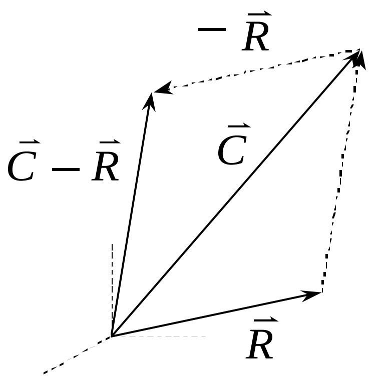

Figure 7-3 Vector subtraction

Vectors may be translated in space without affecting their behavior in mathematical expressions, so long as their length and direction are preserved. For instance, it is common to draw a difference vector as starting at the end point of the "subtrahend" vector ($\vec{R}$ in Fig. 7-3) and ending at the end point of the "minuend" vector ($\vec{C}$ in Fig. 7-3).

7.4.2.2 Vector Dot Products

Vectors may be multiplied in two different ways: by dot product and by cross product. Rules for vector products are different than for products between numbers. Dot products and cross products always involve two vectors. Cross products are discussed in the next section.

The result of a dot product is always a scalar, which is just a fancy term for a number. A dot product is equal to the product of the numerical magnitude of the vectors, which in turn is multiplied by the cosine of the angle between the vectors. The mathematical expression for the dot product between vectors $\vec{A}$ and $\vec{B}$ is $\vec{A} \bullet \vec{B}$. Naturally, for two unit vectors that are 45° apart, their dot product is (1)(1)cos(45) = 0.707. Also, when two vectors have a dot product that equals 0, they must be perpendicular, regardless of their magnitude, because the cosine of 90° is 0. And when two unit vectors have a dot product equal to 1, they must be parallel because the cosine of 0° is 1. Two unit vectors that point in opposite directions yield a dot product of –1 because the cosine of 180° is –1.

When a vector is multiplied with a unit vector via a dot product, the result equals the length of the component of the original vector that is pointing in the direction of the unit vector. The mathematical definitions of geometric tolerances make use of these dot product characteristics.

7.4.2.3 Vector Cross Products

Unlike a vector dot product which yields a number, the result of a vector cross product is always another vector. The mathematical expression for the cross product between vectors $\vec{A}$ and $\vec{B}$ is $\vec{A}\times\vec{B}$, the result of which we will express as $\vec{C}$. By definition, vector $\vec{C}$ is perpendicular to the plane defined by the first two vectors. The magnitude of the vector $\vec{C}$ is equal to the product of the magnitudes of the vectors $\vec{A}$ and $\vec{B}$, which in turn is multiplied by the sine of the angle between $\vec{A}$ and $\vec{B}$. So when two unit vectors are perpendicular, their cross product is another unit vector that is perpendicular to the first two unit vectors;

this because the sine of 90° is 1. And when any two vectors are parallel (or antiparallel), their cross product is a vector of length 0 because the sine of 0°and 180° is 0. The mathematical definitions of geometric tolerances make use of these properties of vector cross products.

7.4.3 Actual Value/Measured Value

A subtle but important distinction exists between the actual value and the measured value of a quantity. Soon after beginning its work program, the Y14.5.1 subcommittee quickly recognized the need to clearly draw this distinction. An *actual value* of a measured quantity is the inherently true value. It is the value that would be obtained by a measurement process that is perfect in every way; that is, a measurement process that has no measurement error or uncertainty associated with it, and which makes use of all of the information that is contained in the item being measured (i.e., the infinite number of data points that a surface consists of). In less esoteric terms, it is the value that we always hope to obtain, but never really can. The actual value can never be obtained because every measurement process has some degree of error and uncertainty associated with it, however small. Moreover, discrete measurement techniques operate on a relatively small subset of the infinite number of points of which a surface is comprised. Even though we can never obtain the actual value, it is important to have a concrete definition of it as well as an understanding of the reasons for its elusiveness.

The *measured value* of a quantity is self-explanatory. Quite simply, it is the value generated by a measurement process. A measured value is an estimate of the actual value; it has an uncertainty associated with it. The goal of any measurement process is to obtain a measured value that approximates the actual value within some tolerable level of uncertainty. The uncertainty associated with a measurement process depends on many factors such as the quantity of data sampled, the data sampling strategy, environmental effects, and so on. This uncertainty is never zero, and the degree to which it is minimized amounts to an economic decision based on the time required to conduct the measurement and the expense of the personnel and equipment employed.

It is not uncommon for the distinction between the measured value and the actual value to become blurred, and this may occasionally contribute to miscommunications between design engineers and metrologists. Early on, the Y14.5.1 subcommittee wrestled with these notions and decided that the scope of its work concerned itself solely with actual values and not with measured values. (The issues surrounding measured values were to be taken up by another subcommittee.) That is not to say that mathematical definitions somehow enable us to obtain actual values. Rather, the mathematical definitions presented in the Y14.5.1 standard focus on the geometric controls that the various tolerance types exert on part features. Further, the tolerance types operate not only on actual, tangible part features, but also more importantly on conceptual models of those part features that exist only on drawings or CAD/solid model representations. The genesis of a manufactured product is a representation of the product that is repeatedly modified, typically involving tradeoffs, in response to various constraints upon it. Allowable geometric variation of the product is one constraint, and the intent of the Y14.5.1 subcommittee was to create mathematical definitions of tolerance types that would be applicable to this conceptual design stage of product development. Accordingly, the notion of an actual value is appropriate.

In fact, in writing mathematical definitions it was crucial to maintain this "separation of church and state" as it were. The potential difficulty in obtaining a reliable measured value of a tolerance was of little or no concern during the development of the Y14.5.1 standard. The philosophy is that it is more important to arm a design engineer with flexible tools to uniquely specify a tolerance design rather than to compromise that ability in favor of easing the eventual measurements required to prove conformance of an actual part to those tolerances. It is inappropriate to standardize tolerances around the state-of-the-art in metrology because it is continually changing.

7.4.4 Datums

7.4.4.1 Candidate Datums/Datum Reference Frames

Datums are geometric entities of perfect form that are derived from datum features specified on a drawing. The configuration of one or more datums as specified in a feature control frame results in a datum reference frame. A datum reference frame essentially amounts to a coordinate system that is located and oriented on the datum features of the part, and from which the location and orientation of other part features are controlled.

For two reasons, a given datum feature may yield more than one datum. Most easy to visualize is the situation whereby a primary datum feature of size is referenced at maximum material condition (MMC) and is manufactured at a size between its maximum material size and its least material size. By the rules of Y14.5, the datum may assume any size, location, and orientation between the datum feature and its MMC limit. These potentially numerous datums form a candidate datum set.

Another reason why a set of candidate datums may result from a given datum feature has to do with the fact that actual datum features, like all actual features, necessarily have form error. Form error often undermines the effectiveness of the rules that Y14.5 specifies in section 4.4.1 for associating perfect form datums to imperfect form datum features. These rules are ideally intended to isolate a single datum from a datum feature, but in practice they reduce the size of the candidate datum set, hopefully to a reasonable extent. For instance, consider a nominal flat surface specified as a primary datum, an actual instance of which has form error consisting of small raised areas scattered all over the surface in such a way that a conceptual, perfect form datum feature (a perfectly flat plane) does not engage the actual surface in just one, unique orientation. In fact, there are multiple sets of three raised areas that provide stable engagement. Each results in a potentially valid datum, and they collectively form the candidate datum set.

Thus, we say in general that a datum feature results in a set of candidate datums. Since each datum in a datum reference frame has (or may have) multiple candidate datums, there are potentially a multitude of candidate datum reference frames. What are we to do with all of these candidates? It is reasonable to conclude that one has the freedom to search among the candidate datum reference frame set for a datum reference frame that yields acceptable evaluations of all tolerances. One could also search for a datum reference frame that collectively minimizes (in some unspecified sense) the departure of all of the features controlled with respect to the datum reference frame. Regardless, if a datum reference frame can be found that yields acceptable evaluations of all tolerances, then the part is considered to be acceptable.

7.4.4.2 Degrees of Freedom

The balance of the discussion on datums will focus on degrees of freedom. A datum reference frame can be thought of as a coordinate system that is fixed to datum features on the part according to rules of association and precedence. If we think of a coordinate system as being represented by three mutually perpendicular axes, then the process of establishing a datum reference frame amounts to a series of positioning and orienting operations of these axes relative to datum features on the part. These positioning and orienting operations take place with respect to a fixed “world” coordinate system.

A datum reference frame has three positional degrees of freedom, and three orientational degrees of freedom within the world coordinate system. In other words, the origin of a datum reference frame may be independently located along three world coordinate system axes. Similarly, the three planes formed by the three pairs of datum reference frame axes have angular relationships to the three planes formed by pairs of world coordinate system axes. The establishment of a datum reference frame equates to a systematic reduction of its available degrees of freedom within the world coordinate system. A datum reference frame that has no available degrees of freedom is said to be fully constrained.

Note that it is not always necessary to fully constrain a datum reference frame. Consider a part that only has an orientation tolerance applied to a feature with respect to another datum feature. One can see that it is not necessary or productive to position the datum reference frame in any manner because the orientation of the feature with respect to the datum is not affected by location of the datum nor of the feature.

The rules of datum precedence embodied in Y14.5 can be expressed in terms of degrees of freedom. A primary datum may arrest one or more of the original six degrees of freedom. A secondary datum may arrest one or more additional *available* degrees of freedom; that is, a secondary datum may not arrest or modify any degrees of freedom that the primary datum arrested. A tertiary datum may also arrest any available degrees of freedom, though there may be none after the primary and secondary datums have done their job; in such a case, a tertiary datum is superfluous and can only add confusion.

The Y14.5.1 standard contains several tables that capture the finite number of ways that datum reference frames may be constructed using the geometric entities points, lines, and planes. Included are conditions between the primary, secondary, and tertiary datums for each case.

7.4.5 Form Tolerances

Form tolerances are characterized by the fact that the tolerance zones are not referenced to a datum reference frame. Form tolerances do not control the form of a feature with respect to another feature, nor with respect to a coordinate system established by other features. Form tolerances are often used to refine the inherent form control imparted by a size tolerance, but not always. Therefore, the mathematical definitions presented in this section reflect the *independent* application of form tolerances. The mathematical description of the net effect of simultaneously applied multiple tolerance types to a feature is not covered in this chapter.

Although form tolerances are conceptually simple, too many users of geometric dimensioning and tolerancing seem to attribute erroneous characteristics to them, most notably that the orientation and/or location of the tolerance zone are related to a part feature. As stated in the prior paragraph, form tolerances are independent of part features or datum reference frames. The mathematical definitions that appear below describe in vector form the geometric elements of the tolerance zones associated with form tolerances; these geometric elements are axes, planes, points, and curves in space. The description of these geometric elements must not be misconstrued to mean that they are specified up front as part of the *application* of a form tolerance to a nominal feature; they are not. The geometric elements of form tolerances are dependent only on the characteristics of the toleranced feature itself, and this is information that cannot be known until the feature actually exists and is measured.

7.4.5.1 Circularity

A circularity tolerance controls the form error of a sphere or any other feature that has nominally circular cross sections (there are some exceptions). The cross sections are taken in a plane that is perpendicular to some *spine*, which is a term for a curve in space that has continuous first derivative (or tangent). The circularity tolerance zone for a particular cross-section is an annular area on the cross-section plane, which is centered on the spine. Because circularity is a form tolerance, the tolerance zone is not related to a datum reference frame, nor is the spine specified as part of the tolerance application. Note that the circularity definition described here is consistent with the ANSI/ASME Y14.5M-1994 definition, but is not entirely consistent with the 1982 version of the standard. See the end of this section for a fuller explanation.

The mathematical definition of a circularity tolerance consists of equations that put constraints on a set of points denoted by $\vec{P}$ such that these points are in the circularity tolerance zone, and no others.

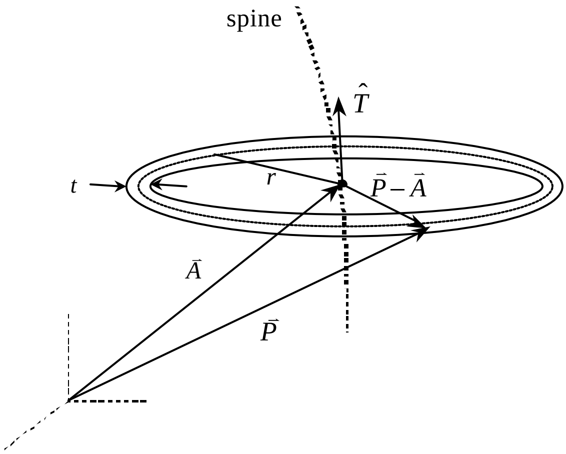

Figure 7-4 Circularity tolerance zone definition

Consider on Fig. 7-4 a point $\vec{A}$ on a spine, and a unit vector $\hat{T}$ which points in the direction of the tangent to the spine at $\vec{A}$.

The set of points $\vec{P}$ on the cross-section that passes through $\vec{A}$ is defined by Eq. (7.1) as follows.

$$\hat{T} \bullet (\vec{P} - \vec{A}) = 0 \tag{7.1}$$

The zero dot product between the vectors $\hat{T}$ and $(\vec{P} - \vec{A})$ indicates that these vectors are perpendicular to one another. Since we know that $\hat{T}$ is perpendicular to the spine at $\vec{A}$, and $\vec{P} - \vec{A}$ is a vector that points from $\vec{A}$ to $\vec{P}$, then the points $\vec{P}$ must be on a plane that contains $\vec{A}$ and that is perpendicular to $\hat{T}$. Thus, we have defined all of the points that are on the cross section. Next, we need to restrict this set of points to be only those in the circularity tolerance zone.

As was stated above, the circularity tolerance zone consists of an annular area, or the area between two concentric circles that are centered on the spine. The difference in radius between these circles is the circularity tolerance t.

$$\left| \left\| \vec{P} - \vec{A} \right\| - r \right| \leq \frac{t}{2} \tag{7.2}$$

Eq. (7.2) says that there is a reference circle at a distance r from the spine, and that the points $\vec{P}$ must be no farther than half of the circularity tolerance from it, either toward or away from the spine. This equation completes the mathematical description of the circularity tolerance zone for a particular cross section.

To verify that a measured feature conforms to a circularity tolerance, one must establish that the measured points meet the restrictions imposed by Eqs. (7.1) and (7.2). In geometric terms, one must find a spine that has the circularity tolerance zones that are created according to Eqs. (7.1) and (7.2), containing all of the measured points. The reader will likely find this definition of circularity foreign, so some explanation is in order.

As was stated earlier in this section, the details of circularity that are discussed here correspond to the ANSI/ASME Y14.5M-1994 standard, which contains some changes from the 1982 version. The 1982 version of the standard, as written, required that cross sections be taken perpendicular to a *straight* axis, and that the circularity tolerance zones be centered on that straight axis, thereby effectively limiting the application of circularity to surfaces of revolution. In order to expand the applicability of circularity tolerances to other features that have circular cross sections, such as tail pipes and waveguides, the definition of circularity was modified such that circularity controls form error with respect to a *curved* "axis" (a spine) rather than a straight axis. The 1994 standard preserves the centering of the circularity tolerance zone on the spine.

Unfortunately, the popular interpretation of circularity does not correspond to either the 1982 or the 1994 versions of Y14.5M. Rather, a metrology standard (B89.3.1-1972, Measurement of Out of Roundness) seems to have implicitly provided an alternative definition of circularity by virtue of the measurement techniques that it describes. The main difference between the B89 metrology standard and the Y14.5M tolerance definition standard is that the B89 standard does not require the circularity tolerance zone to be centered on the axis. Instead, various fitting criteria are provided for obtaining the "best" center of the tolerance zone for a given cross section. Without delving into the details of the B89.3.1-1972 standard, suffice it to say that the four criteria are least squares circle (LSC), minimum radial separation (MRS), maximum inscribed circle (MIC), and minimum circumscribed circle (MCC).

There is a rather serious geometrical ramification to allowing the circularity tolerance zone to "float." Consider in Fig.7-5 a three-dimensional figure known as an elliptical cylinder which is created by translating or extruding an ellipse perpendicular to the plane in which it lies. Obviously, such a figure has elliptical cross sections, but it also has perfectly *circular* cross sections if taken perpendicular to a properly titled axis.

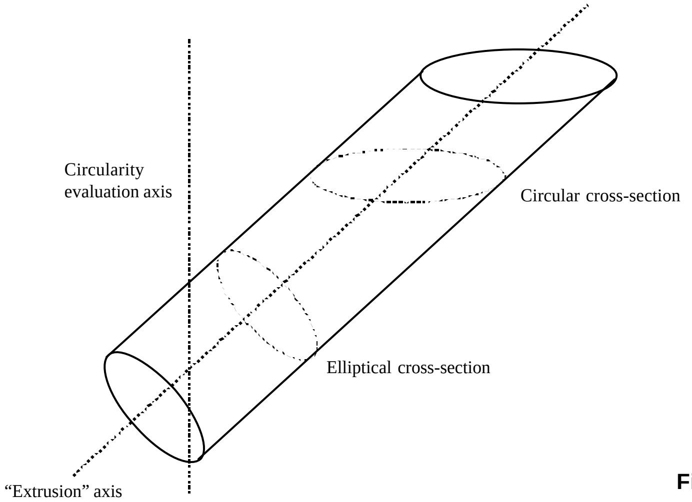

Figure 7-5 Illustration of an elliptical cylinder

Thus, a perfectly formed elliptical cylinder (even one with high eccentricity) would have no circularity error as measured according to the B89.3.1-1972 standard. Of course, any sensible, well-trained metrologist would intuitively select an axis for evaluating circularity that closely matches the axis of symmetry of the feature, and would thus find significant circularity error. However, as tolerancing and metrology progress toward computer-automated approaches (as the design and solid modeling disciplines already have), we must depend less and less on subjective judgment and intuition. It is for this reason that the relevant standards committees have recognized these issues with circularity tolerances and measurements, and they are working toward their resolution.

Creation of a mathematical definition of circularity revealed the inconsistency between the Y14.5M-1982 definition of circularity and common measurement practice as described in B89.3.1-1972, and also revealed subtle but potentially significant problems with the latter. This example illustrates the value that mathematical definitions have brought to the tolerancing and metrology disciplines.

7.4.5.2 Cylindricity

A cylindricity tolerance controls the form error of cylindrically shaped features. The cylindricity tolerance zone consists of a set of points between a pair of coaxial cylinders. The axis of the cylinders has no predefined orientation or location with respect to the toleranced feature, nor with respect to any datum reference frame. Also, the cylinders have no predefined size, although their difference in radii equals the cylindricity tolerance *t*.

We mathematically define a cylindricity tolerance zone as follows. A cylindricity axis is defined by a unit vector $\hat{T}$ and a position vector $\vec{A}$ as illustrated in Fig. 7-6.

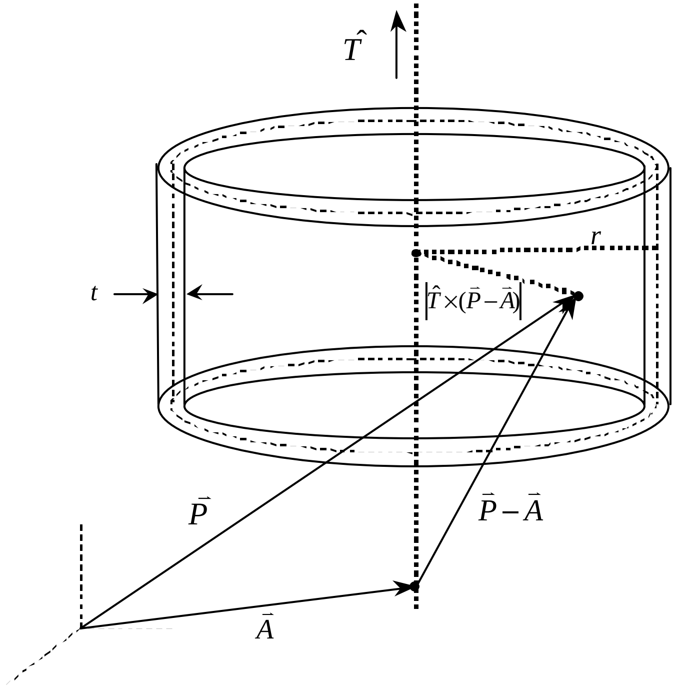

Figure 7-6 Cylindricity tolerance definition

If we consider the unit vector $\hat{T}$, which points parallel to the cylindricity axis, to be anchored at the end of the vector $\vec{A}$, one can see from Fig. 7-6 that the distance from the cylindricity axis to point $\vec{P}$ is obtained by multiplying the length of the unit vector $\hat{T}$ (equal to one by definition) by the length of the vector $\vec{P}-\vec{A}$, and by the sine of the angle between $\hat{T}$ and $\vec{P}-\vec{A}$. The mathematical operations just described are those of the vector cross product. Thus, the distance from the axis to a point $\vec{P}$ is expressed mathematically as $\left|\hat{T}\times(\vec{P}-\vec{A})\right|$. To generate a cylindricity tolerance zone, the points $\vec{P}$ must be restricted to be between two coaxial cylinders whose radii differ by the cylindricity tolerance t.

Eq. (7.3) constrains the points $\vec{P}$ such that their distance from the surface of an imaginary cylinder of radius *r* is less than half of the cylindricity tolerance.

$$\left|\left|\hat{T}\times(\vec{P}-\vec{A})\right|-r\right|\leq\frac{t}{2} \tag{7.3}$$

If, when assessing a feature for conformance to a cylindricity tolerance, we can find an axis whose direction and location in space are defined by $\hat{T}$ and $\vec{A}$, and a radius r such that all of the points of the actual feature consist of a subset of these points $\vec{P}$, then the feature meets the cylindricity tolerance.

7.4.5.3 Flatness

A flatness tolerance zone controls the form error of a nominally flat feature. Quite simply, the toleranced surface is required to be contained between two parallel planes that are separated by the flatness tolerance. See Fig. 7-7.

To express a flatness tolerance mathematically, we define a reference plane by an arbitrary locating point $\vec{A}$ on the plane and a unit direction $\hat{T}$ that points in a direction normal to the plane. The quantity

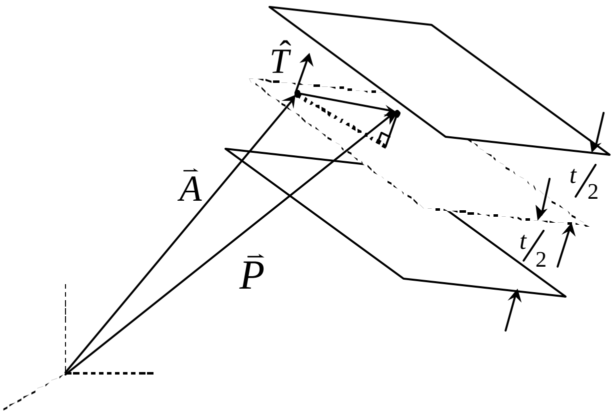

Figure 7-7 Flatness tolerance definition

$\vec{P} - \vec{A}$ is the vector distance from the reference plane's locating point to any other point $\vec{P}$. Of more interest though is the component of that distance in the direction normal to the reference plane. This is obtained by taking the dot product of $\vec{P} - \vec{A}$ and $\hat{T}$.

$$\left|\hat{T} \bullet (\vec{P} - \vec{A})\right| \leq \frac{t}{2} \tag{7.4}$$

Eq. (7.4) requires that the points $\vec{P}$ be within a distance equal to half of the flatness tolerance from the reference plane.

In mathematical terms, to determine conformance of a measured feature to a flatness tolerance, we must find a reference plane from which the distances to the farthest measured point to each side of the reference plane are less than half of the flatness tolerance.

Note that Eq. (7.4) is not as general as it could be. The true requirement for flatness is that the *sum* of the normal distances of the most extreme points of the feature to each side of the reference plane be no more than the flatness tolerance. Stated differently, although Eq. (7.4) is not incorrect, there is no requirement that the reference plane equally straddle the most extreme points to either side. In fact, many coordinate measuring machine software algorithms for flatness will calculate a least squares plane through the measured data points and assess the distances to the most extreme points to each side of this plane. In general, the least squares plane will not equally straddle the extreme points, but it may serve as an adequate reference plane nevertheless.

7.5 Where Do We Go from Here?

Release of the Y14.5.1 standard in 1994 addressed one of the major recommendations that emanated from the NSF Tolerancing Workshop. However, the work of the Y14.5.1 subcommittee is not complete. The Y14.5.1 standard represents an important first step in increasing the formalism of geometric tolerancing, but many other things must happen before we can claim to have resolved the metrology crisis. The good news is that things are happening. Research efforts related to tolerancing and metrology have accelerated over the time frame since the GIDEP Alert, and we are moving forward.

7.5.1 ASME Standards Committees

Though five years have passed since the release of the Y14.5.1 standard, it is difficult to discern the impact that it has had on the practitioners of geometric tolerancing. However, the impact that it has had on the standards development scene is easier to measure. Advances in standards work are greatly facilitated when standards developers have a minimal dependence on subjective interpretations of the standardized materials. Indeed, it is the specific duty and responsibility of standards developers to define their subject matter in objectively interpretable terms; otherwise standardization is not achieved. The Y14.5.1 standard, and the philosophy that it embodies, provides a means for ensuring a lack of ambiguity in standardized definitions of tolerances.

Despite the alphanumeric subcommittee designation (Y14.5.1), which suggests that it sit below the Y14.5 subcommittee, the Y14.5.1 subcommittee has the same reporting relationship to the Y14 main committee, as does the Y14.5 subcommittee. The new Y14.5.1 effort was truly a parallel effort to that of Y14.5 (though certainly not entirely independent). Its value has been sufficiently demonstrated within the subcommittees to the extent that the leaders of each group are establishing a much closer degree of collaboration. The result will undoubtedly be better standards, better tools for specifying allowable part variation, less disagreement between suppliers and customers regarding acceptability of parts, and better and cheaper products.

7.5.2 International Standards Efforts

The impact of the Y14.5.1 standard extends to the international standards scene as well. Over the past few years, the International Organization for Standardization (ISO) has been engaged in a bold effort to integrate international standards development across the disciplines from design through inspection. As a participating member body to this effort, the United States has made its share of contributions. Among these contributions are mathematical definitions of form tolerances. These definitions are closely derived from the Y14.5.1 versions, but customized to reflect the particular detailed differences, where they exist, between the Y14.5 definitions and the ISO definitions. As other ISO standards are developed or revised, additional mathematical tolerance definitions will be part of the package.

7.5.3 CAE Software Developers

Aside from standards developers, computer aided engineering (CAE) software developers should be the key group of users of mathematical tolerance definitions. Recalling the lack of uniformity and correctness in CMM software as brought to light by the GIDEP Alert, it should not be difficult to see the need for programmers of CAE systems (including design, tolerancing, and metrology) to know the detailed aspects of the tolerance types and code their software accordingly. In some cases, this can be achieved by coding the mathematical expressions from the Y14.5.1 standard directly into their software.

We are not yet aware of the actual extent of usage of the mathematical tolerance definitions from the Y14.5.1 standard among CAE software developers. Where vendors of such software claim compliance to US dimensioning and tolerancing standards, customers should rightly expect that the vendor owns a copy of the Y14.5.1 standard and has ensured that its algorithms are consistent with its requirements. It might be reasonable to assume that this is not the case across the board, and it would be a worthy endeavor to determine the extent of any such lack of compliance. As of this writing, ten years have passed since the GIDEP Alert, and perhaps the time is right to see whether the situation has improved with metrology software.

7.6 Acknowledgments

The groundbreaking Y14.5.1 standard was the result of a collective effort by a team of talented and unique individuals with diverse but related backgrounds. This author was but one contributor to the effort, and I would like to sincerely thank the other contributors for their wit, wisdom, and camaraderie; I learned quite a lot from them through this process. Rather than list them here, I refer the reader to page v of the standard for their names and their sponsoring organizations. At the top of that list is Mr. Richard Walker who demonstrated notable dedication and leadership through several years of intense development.

Unlike many other countries, standards of these types in the United States are voluntarily specified and observed by customers and suppliers rather than mandated by government. Moreover, the standards are developed primarily with private funding by companies that have an interest in the field and have personnel with the proper expertise. These companies enable committee members to contribute to standards development by providing them with travel expenses for meetings and other tools and resources needed for such work.

7.7 References

1. British Standards Institution. 1989. *BS 7172, British Standard Guide to Assessment of Position, Size and Departure from Nominal Form of Geometric Features*. United Kingdom. British Standards Institution.
2. Hocken, R.J., J. Raja, and U. Babu. Sampling Issues in Coordinate Metrology. *Manufacturing Review* 6(4): 282-294.
3. James/James. 1976. *Mathematical Dictionary - 4th Edition.* New York, New York: Van Nostrand.
4. Srinivasan V., H.B. Voelcker, eds. 1993. *Proceedings of the 1993 International Forum on Dimensional Tolerancing and Metrology*, CRTD-27. New York, New York: The American Society of Mechanical Engineers.
5. The American Society of Mechanical Engineers. 1972. *ANSI B89.3.1 - Measurement of Out-of-Roundness*. New York, New York: The American Society of Mechanical Engineers.
6. The American Society of Mechanical Engineers. 1994. *ASME Y14.5 - Dimensioning and Tolerancing*. New York, New York: The American Society of Mechanical Engineers.
7. The American Society of Mechanical Engineers. 1994. *ASME Y14.5.1 - Mathematical Definition of Dimensioning and Tolerancing Principles*. New York, New York: The American Society of Mechanical Engineers.
8. Tipnis V. 1990. *Research Needs and Technological Opportunities in Mechanical Tolerancing,* CRTD-15. New York, New York: The American Society of Mechanical Engineers.
9. Walker, R.K. 1988. *CMM Form Tolerance Algorithm Testing, GIDEP Alert, #X1-A-88-01A.*
10. Walker R.K., V. Srinivasan. 1994. Creation and Evolution of the ASME Y14.5.1M Standard. *Manufacturing Review* 7(1): 16-23.