REPRESENTATION OF GEOMETRIC VARIATIONS USING MATRIX TRANSFORMS FOR STATISTICAL TOLERANCE ANALYSIS IN ASSEMBLIES

by

OLIVIER L. GILBERT

B. S. Industrial Engineering
École Polytechnique de Montréal
May 1988

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE IN MECHANICAL ENGINEERING
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 1992

© Olivier L. Gilbert, 1992

The author hereby grants to M.I.T. and the Charles Stark Draper Laboratory, Inc. permission to reproduce and to distribute copies of this thesis document in whole or in part.

Signature of Author ________________________________

Department of Mechanical Engineering
May 1, 1992

Certified by ________________________________

Daniel E. Whitney, PhD
Thesis Supervisor

Accepted by ________________________________

An A. Sonin, Chairman
Departmental Graduate Committee

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUN 17 1992
DISCLAIMER

This research is supported by National Science Foundation grants DMC 87-15800 and DDM 90-07062. It represents the opinions of the author only, and is offered for the simulation and exchange of ideas. It does not represent the policies of the National Science Foundation, the Charles Stark Draper Laboratory, or the Massachusetts Institute of Technology.
REPRESENTATION OF GEOMETRIC VARIATIONS USING MATRIX TRANSFORMS
FOR STATISTICAL TOLERANCE ANALYSIS IN ASSEMBLIES

by

Olivier L. Gilbert

Submitted to the Department of Mechanical Engineering
on June 8, 1992 in partial fulfillment of the requirements for the degree of
Masters of Science in Mechanical Engineering at the
Massachusetts Institute of Technology

ABSTRACT

In order to create competitive products, production system requirements must be
considered early in the design process. Since variations are inherent to any manufacturing
and assembly process, geometric tolerances have to be acknowledged and assigned in a fashion
that will not only respect the functional requirements of the product, but that will make the use of
cheap and efficient assembly system possible. Assembly sequence analysis represents a
method for concurrently considering design and assembly, by allowing to reduce the
choice of interesting sequences based on multiple criteria, and by providing feedback on
the critical aspects of the design. In this thesis, tolerance impact evaluation is developed
as a criterion by which assembly sequences can be analyzed.

Tolerance analysis that accounts for variations in both the subassembly and the
assembly system is presented as a method to help evaluate the probability of success of
critical assembly mates in many possible sequences. Representation of tolerances using
homogeneous matrix transforms is particularly well suited for this objective. Using
variations on the six kinematic degrees of freedom associated with a rigid body, feature
(surfaces, axes, etc.) position and orientation variability can be established and
propagated statistically. Through a closed-form error propagation algorithm used in
robotic applications, tolerances so represented can be analyzed rapidly in three
dimensions.

This thesis develops families of homogeneous matrix transforms and associated
errors that can represent most standard tolerances as presented in ANSI Y14.5M-1982,
and peg-hole mate variation. Form tolerances are neither compatible with matrix
transform representation, nor important in evaluation of assembly sequences. This
mathematical representation is therefore suggested as a complementing representation
of tolerances that can co-exist with other definitions which have been developed for
other usage, as for example, quality control activities.

The use of a computer program developed by a previous student has been
facilitated and a case study is presented.

Thesis Supervisor: Dr. Daniel E. Whitney, Lecturer, Mechanical Engineering
ACKNOWLEDGMENTS

First, I thank Dr. Daniel E. Whitney for his supervision. Through a good balance of guidance and freedom, he allowed my making of this thesis to be a worthy experience, both academically and professionally.

I also thank Mr. Rick Metzinger for his continued support in showing me how to deal with computers that sometime seemed to have a mind of their own. I also underline the great help I was able to get from several people associated with the laboratory's "Help Line". Their support in my many computer hardware and communication problems was prompt and friendly; I congratulate them for their professionalism and efficiency.

I extend my appreciation to Dr. Thomas De Fazio, Mr. Richard Gustavson, and Mr. Steve Luby for their help and support on various issues. The discussions they allowed, both professional and personal, have enlightened my reflections and my stay at Draper.

My wife, Marlene Thibodeau, has supported my work in many ways. In addition to her participation in this thesis' presentation, she has cheered me up in those tough moments, shared with me the joys of accomplishment, and helped me believe in me. Merci Marlene.

ACKNOWLEDGMENTS 2
(June 8, 1992)

This thesis was prepared at The Charles Stark Draper Laboratory, Inc., under Contracts 62800 and 82000.

Publication of this thesis does not constitute approval by Draper or the sponsoring agency of the findings or conclusions contained herein. It is published for the exchange and simulation of ideas.

I hereby assign my copyright of this thesis to The Charles Stark Draper Laboratory, Inc., Cambridge, Massachusetts.

Permission is hereby granted by The Charles Stark Draper Laboratory, Inc., to the Massachusetts Institute of Technology to reproduce any or all of this thesis.
À Louis,

and to Anne-Marie
TABLE OF CONTENTS

1 Introduction 13
  1.1 Goal 14
  1.2 Motivation 15
  1.3 Background 16
  1.4 Thesis organization 19

2 Previous Work 23
  Introduction 24
  2.1 Feature-based design for assembly 26
  2.2 Assembly sequence analysis 29
  2.3 Tolerance representation and analysis 38
    2.3.1 Significance of tolerances 39
      • Component specification versus mating variations 40
      • Mating variations versus assembly variations: the assembly system 42
    2.3.2 Tolerance representation 43
    2.3.3 Tolerance analysis 47
      • Possible approaches 48
      • Proposed: 3-D, closed-form algorithm using matrix transforms 53
        - Derivation outline 54
        - Derivation 56
        - Requirements for tolerance representation 66

2.4 Part mating theory 67
  Conclusion 73

3 Overview of geometric dimensioning and tolerancing 75
  Introduction 76
  3.1 Conventional versus geometric 77
  3.2 Generalities 80
    3.2.1 Definitions 80
    3.2.2 Modifiers: maximum material condition 83
    3.2.3 Datums 85
  3.3 Tolerance classification 90
    3.3.1 Related versus individual features 92
    3.3.2 Location 94
    3.3.3 Orientation 97
    3.3.4 Runout 100
    3.3.5 Form and profile 103
6 Case study

Introduction 192

6.1 Evaluation of assembly success 193

6.2 The impeller assembly 196

6.2.1 Sequence with no refixturing 197
  • Propagation of errors 199
  • Consolidation with part mating theory 211

6.2.2 Alternate sequence: using a refixturing 213

6.2.3 Choice of best sequence 217

Conclusion 219

7 Conclusions and recommendations for future work 221

7.1 Recommendations for future work 222
  • Develop more matrix transforms 222
  • Simplify the identification of tolerance chains 223
  • Simplify and automate the input of data in computer program 226
  • Fully incorporate tolerance analysis into ASA 227

7.2 Conclusion 228

References 231
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematization of FDBA prototype</td>
<td>28</td>
</tr>
<tr>
<td>2.2</td>
<td>Assembly sequence generation</td>
<td>31</td>
</tr>
<tr>
<td>2.3</td>
<td>A more complicated assembly graph in the editing module</td>
<td>32</td>
</tr>
<tr>
<td>2.4</td>
<td>Assembly sequences affected by tolerances</td>
<td>34</td>
</tr>
<tr>
<td>2.5</td>
<td>Flow chart of tolerance-based sequence editing</td>
<td>37</td>
</tr>
<tr>
<td>2.6</td>
<td>Phases of mating pegs and holes</td>
<td>68</td>
</tr>
<tr>
<td>2.7</td>
<td>Schematic illustrations of wedging and jamming</td>
<td>70</td>
</tr>
<tr>
<td>2.8</td>
<td>Parameters used in determining allowable lateral and angular errors</td>
<td>71</td>
</tr>
<tr>
<td>2.9</td>
<td>Geometry constraints on allowed lateral and angular errors to permit chamfer crossing and avoidance of wedging</td>
<td>72</td>
</tr>
<tr>
<td>3.1</td>
<td>Conventional dimensioning and tolerancing</td>
<td>78</td>
</tr>
<tr>
<td>3.2</td>
<td>Conventional versus geometric: shape of the tolerance zone</td>
<td>79</td>
</tr>
<tr>
<td>3.3</td>
<td>Maximum material condition (MMC) principle: an example</td>
<td>83</td>
</tr>
<tr>
<td>3.4</td>
<td>Using a surface from a planar feature of size as datum</td>
<td>85</td>
</tr>
<tr>
<td>3.5</td>
<td>Datum versus datum surface or datum feature</td>
<td>86</td>
</tr>
<tr>
<td>3.6</td>
<td>Establishing the datum planes</td>
<td>87</td>
</tr>
<tr>
<td>3.7</td>
<td>Exact constraint assembly specification: similar to a datum reference frame</td>
<td>88</td>
</tr>
<tr>
<td>3.8</td>
<td>Implicit orientation specification using a datum reference frame</td>
<td>89</td>
</tr>
<tr>
<td>3.9</td>
<td>Parallelism: many specifications possible</td>
<td>90</td>
</tr>
<tr>
<td>3.10</td>
<td>Position tolerance associated with a non-cylindrical feature</td>
<td>95</td>
</tr>
<tr>
<td>3.11</td>
<td>Position tolerance used to specify coaxial feature</td>
<td>96</td>
</tr>
<tr>
<td>3.12</td>
<td>Parallelism: used as a refinement of location tolerance</td>
<td>97</td>
</tr>
<tr>
<td>3.13</td>
<td>Orientation: tolerance zone shape depends on datum used</td>
<td>98</td>
</tr>
<tr>
<td>3.14</td>
<td>Angularity tolerance</td>
<td>99</td>
</tr>
<tr>
<td>3.15</td>
<td>Parallelism of a surface</td>
<td>100</td>
</tr>
<tr>
<td>3.16</td>
<td>Total runout</td>
<td>101</td>
</tr>
<tr>
<td>3.17</td>
<td>Runout: can control form (circular, total) and orientation (total)</td>
<td>102</td>
</tr>
<tr>
<td>3.18</td>
<td>Flatness: a form tolerance</td>
<td>104</td>
</tr>
<tr>
<td>3.19</td>
<td>Rough form variations: confined within size or orientation tolerance zones</td>
<td>105</td>
</tr>
<tr>
<td>4.1</td>
<td>Tolerance on a circular feature: cylindrical tolerance zone for its axis</td>
<td>113</td>
</tr>
<tr>
<td>4.2</td>
<td>Representation of tolerance zone using variation ranges on a subset of the 6 degrees of freedom of Cartesia frames</td>
<td>114</td>
</tr>
<tr>
<td>4.3</td>
<td>Tolerance zone in both domains: if they are no linear errors, the maximum angular error are directly proportional to the tolerance specification $2r_{max}$</td>
<td>116</td>
</tr>
<tr>
<td>4.4</td>
<td>If linear errors exist ($r \neq 0$), then angular errors are limited to</td>
<td></td>
</tr>
</tbody>
</table>
a smaller region.

Figure 4.5 If linear errors exist \( (r \neq 0) \), the resulting angle for \( r \) is uniformly distributed in \([0,2\pi]\).

Figure 4.6 Zone boundary in variation space.

Figure 4.7 Density visualization planes.

Figure 4.8 Analytic surfaces resulting from joint probability density of normally distributed angular variations.

Figure 4.9 Radial distance in joint normal density: distributed like the two variables of equal variance.

Figure 4.10 Simulation of real densities: \( \alpha \) and 3 dependent variables: \( r, \theta_x \), and \( \theta_y \).

Figure 4.11 Simulation at specific instances of \( r \) and \( \alpha \).

Figure 4.12 Simulation at specific instances of \( r \) and with \( \alpha \) uniformly distributed in \([0,2\pi]\).

Figure 4.13 Comparison between simulated and analytic densities at specific instances of \( r \).

Figure 4.14 Search for appropriate instances of \( r \) for optimization.

Figure 4.15A Optimization at \( r=0.04 \) \( r_{\text{max}} \).

Figure 4.15B Optimization at \( r=0.36 \) \( r_{\text{max}} \).

Figure 4.15C Optimization at \( r=0.80 \) \( r_{\text{max}} \).

Figure 4.16 Parameters for position using a datum reference frame.

Figure 4.17 Parameters for various tolerance specifications.

Figure 4.18 Parameters for planar size tolerance.

Figure 4.19 Extreme values for variates in size specification.

Figure 4.20 Maximum boundary in parameter space for size specification (2D).

Figure 4.21 Comparison and optimization of 2D size specification: simulated versus analytical probability densities of joint variables.

Figure 4.22 Extreme values for variates in size specification (3D).

Figure 4.23 Maximum boudary in a parameter space for size specification (3D).

Figure 4.24 Approximation of reality-model's diamond by the analytical solution's ellipsoid.

Figure 4.25 Analogy between MMC principle and peg-hole position variation.

Figure 4.26 Parameters for orientation specifications as refinement of size.

Figure 4.27 Total runout tolerance on surface perpendicular to datum axis.

Figure 5.1 Variation zone for peg position in hole.

Figure 5.2 Variates involved in peg in hole problem.

Figure 5.3 Parameters for simulation of peg-hole mating variations.

Figure 5.4 Simulated density curves of peg position and radius of variation zone.

Figure 6.1 Independent normal probability densities for lateral and angular errors.

Figure 6.2 Worst case plane and density curves.
Figure 6.3 Subassembly and assembly system described by a unique matrix
transform sequence. 196
Figure 6.4 Statistical linear and angular errors between frames. 197
Figure 6.5 Sequence graph editing for impeller case study. 198
Figure 6.6 a) Impeller specifications. 200
Figure 6.6 b) Retainer specifications. 200
Figure 6.6 c) Shaft specifications. 201
Figure 6.7 Sequence with no refixturing: the peg-hole mate. 203
Figure 6.8 Perpendicularity of an axis to a surface: a conical tolerance zone. 208
Figure 6.9 Normal density ellipsoids for linear and angular errors of critical
peg-hole mate. 211
Figure 6.10 Results of peg-hole mate for sequence with no refixturing. 213
Figure 6.11 Sequence with refixturing: the peg-hole mate. 216
Figure 6.12 Results of peg-hole mate for sequence with a refixturing. 217

Figure 7.1 Mating and fixturing surface designation. 224

LIST OF TABLES

Table 2.1 Example of features. 26
Table 2.2 Examples of feature attributes. 27
Table 3.1 Tolerance classification. 91
Table 4.1 Errors between analytic and reality-simulated densities at
different values of $r$. 139
Table 4.2 Summary of errors associated with tolerance zones expressed with
homogeneous matrix transforms. 170
Table 5.1 Simulation of tolerance analysis algorithm for peg-hole mate. 186
Chapter 1

Introduction
1.1 Goal

A computer program to propagate statistical tolerances mathematically in closed-form through an assembly or sub-assembly has been written by [Jastrzebski-'90]. The method developed uses homogeneous matrix transforms and associated variabilities to describe the statistical position and orientation of product geometric features (surface, peg, hole, etc.) with respect each other.

Considerable effort is still needed, though, to accurately represent widely used tolerance specifications in the language prescribed by homogeneous matrix transforms. In this context, the designer's intent, or what he or she views as acceptable variations, has to be described with the 6 variables that correspond to the 6 degrees of freedom a feature can have: 3 translation variations along the x, y and z axes, and 3 rotation variations around the same axes.

In this thesis, our goal is to develop families of matrix transforms to represent most tolerance specifications contained in the Y14.5 standard, and to represent variations included in typical mating relationships. An emphasis is put on the specification types pertinent to assembly sequence analysis. Furthermore, since not all specification and variation types can be represented using homogeneous matrix transforms, we want to gain a better understanding of the limitations of this representation.

In the context of automated assembly evaluation, tolerance analysis that includes both the subassembly and the assembly system has to be performed for each assembly
step of each assembly sequence candidate. Such an analysis will determine the relative position and orientation variation of features to mate. The sequences that insure the highest probability of success for critical assembly mates can be kept for further analysis. We start with the assumption that the final assembly has already been evaluated and conforms with the functional requirement and tolerances. The task at hand is to see whether or not the intermediate assembly steps are possible with the existing assembly system, and if so, with what probability of success. It is also our goal to demonstrate, through a case study, that matrix transform representation of variations is specially well suited for evaluating probability of mate success, specially when used in conjunction with the concepts of part mating theory.

1.2 Motivation

Variations are inherent to any manufacturing and assembly process. As a result, products of these processes present geometric variations that have to be accounted for in every step of the development activity and in the product's useful life. Development includes from initial sketches to fully operating production line, including among others, detailed design and assembly process planning. Product development is and activity which rapidity of accomplishment is crucial to competitiveness. Time to market is directly related to development time, and windows of opportunity will be attained if both delays are short. To accomplish this time reduction, development teams apply concurrent engineering concepts, i.e. simultaneous design of products and processes. This approach replaces more traditional methods where design tasks are done sequentially, necessitating many iterations of the "design-prototype-test-trytobuildcheaply" cycle.
Among the development tasks, we find the analysis of assembly sequence. This analysis helps in planning the assembly process by evaluating many possible sequences in terms of feasibility and cost. Needless to say, geometric variations on components will greatly influence the assembly process, specially if that process is automated. Tolerances, therefore, should be assigned in such ways to respect compatibility with cheap, quality-insuring, and efficient assembly processes.

The research in tolerance representation using homogeneous matrix transforms will allow to use a closed-form tolerance analysis algorithm, and therefore permit the evaluation of many assembly sequences in a short period of time. But mostly, it will allow this analysis to be fully compatible with current practices and standards in dimensioning and tolerancing (D&T).

Designers should then be better positioned to assign tolerance specifications to respect assembly processes, and process engineers should be better positioned to choose the assembly sequence that will not only maximize the likelihood of successful mating, but minimize the assembly cost.

1.3 Background

They are many areas of work relevant to this thesis. Tolerance analysis, which finds the tolerance on an assembly dimension using components contributions, and tolerance synthesis, by which a desired tolerance on an assembly dimension is distributed among the components that affect it, are some of them. These two
complementing approaches to tolerances are often used iteratively to help designers assign functional tolerance specifications on components and assemblies. [Bjorke-'89], [Greenwood and Chase-'87], and [Scott and Gabriele-'89], among others, have developed various approaches to perform these tasks.

All manufacturing processes are the source of geometric variabilities on the final product: They have to be taken into account at the design stage by product designers specifying permissible tolerances on important dimensions. [Gossard-'88] presents an enlightening discussion on the problem at hand. The task of assigning tolerances is not an easy one. While the information on which to base the decisions is scarce, the impact of tolerances on product cost and quality is high.

Three factors are taken into account when assigning tolerances. The first one is the respect of functional requirements. In order for a product to perform in a certain manner, a specific dimension might not be allowed to vary outside a given range. Unfortunately, although this is a critical issue, the lack of available knowledge is such that few engineers know the exact relationship between the dimension of the part and its ability to possess the best possible function [Bjorke-'89].

The second factor relates to the manufacturing process used to fabricate the individual parts of a product. As said before, all processes will introduce geometric variability and in general, the smaller the tolerance has to be held, the higher the cost of manufacturing will be. For this reason, the designer must evaluate whether or not the small tolerances he or she has specified are worth the increased cost. This reflection is most applicable at the stage of tolerance synthesis. For this task, knowledge of the
machines capabilities with their respective cost is necessary. This information is often referred to as (manufacturing-related) tolerance-cost models.

The third factor, and the one of highest interest for this thesis, relates to the assembly process per se. Since tolerances also have precise relationships with the analysis of assembly difficulty, sequence, and process capability, it becomes important to develop assembly-related tolerance-cost models as tools to ease decisions related to assembly process. Some have developed assemblability evaluation models ([Okano-'91], [Boothroyd and Dewhurst-'87]) but considerations of variabilities are often limited to the types of fit involved, as presented in ANSI's standard on "Preferred Limits and Fits for Cylindrical Parts" [ANSI B4.1-'79]. The main considerations in these approaches are ease of part feeding, ease of part handling, number of parts per assembly, and ease of contact. Tolerance-cost models directly related to assembly would, among others, ease the task of assigning tolerances.

A clear understanding of the physical relationships between two mating parts will also contribute in establishing guide-lines for variation control and specification. Important work on the theory of mating parts has been done by, among others, [Nevins and Whitney-89], and has resulted in useful knowledge to help evaluate the probability of success of mating parts. The proposed assessment is based on various factors, but most importantly on the lateral and angular position error between two features to mate. The relation between this work and propagation of errors through subassemblies and assembly systems is then clearly defined.

Finally, Assembly Sequence Analysis (ASA) is a critical step in assembly process
planning. Once one has addressed the importance of variations in the choice of a good assembly sequence among many possibilities, one needs a rational methodology to evaluate these possible sequences. ASA is a rational means of considering all feasible assembly sequences in the sequence selection process [Baldwin-'90]. It represents a method for concurrently considering design and assembly, by allowing to reduce the choice of interesting sequences based on multiple criteria, and by providing feedback on the critical aspects of the design. The work of this thesis aims at incorporating tolerances as a criterion used in ASA.

1.4 Thesis organization

Chapter 2 takes a closer look at previous work done in the area of assembly process evaluation. Feature-based design is presented as a way to describe a component and its corresponding assembly by extending the information beyond the classic geometric shape. Assembly sequence analysis (ASA) is also described in more details so that the work of this thesis can be well positioned and justified in a global approach to assembly evaluation. Specific approaches to tolerance analysis, a key task in considering tolerances in ASA, are described along with their requirements in the way tolerances are mathematically represented. In this section, emphasis is put on a closed-loop 3-D algorithm that [Jastrzebski-'91] has implemented as a computer-based tolerance analysis tool. The main results of part mating theory are also presented in order to directly relate, later in the case study of chapter 6, the physics of mating parts with the information obtained from tolerance analysis, and thereby allowing an evaluation of the probability of success of critical automated assembly steps. Finally,
we make an attempt at classifying many variabilities that have to be taken into account when analyzing the assembly process.

In chapter 3, we present an overview the geometric dimensioning and tolerancing (D&T) standard and its various tolerance specifications. Issues such as evolution of the D&T approaches, general concepts, and use of special attributes to the specifications are also presented along with their pertinence and compatibility with matrix transform representation schemes, in light of assembly analysis purpose and requirements, is established.

In chapter 4, we describe the method used to represents these allowed variation, or design intents, in a format appropriate to homogeneous matrix transforms. This permits the 3-D closed-form tolerance analysis algorithm, presented in chapter 2, to be implemented with a reasonable respect of current D&T practices and standards. Furthermore, in order to comply with ANSI's Y14.5 guidelines [ANSI Y14.5M-1982], a widely used D&T standard, we concentrate on transforming the specifications associated with this standard, and we focus our energy on those specifications most influential to the assembly process.

In chapter 5, we develop a matrix transform representation of an important case of mate variations: the position variation of a peg in a hole. Unlike individual component variations, controlled by standard tolerance specifications, mate variations have not been analyzed as comprehensively in existing literature. They involve the nature and the behavior of the contact between two or more components, through single or multiple surfaces (or features) on each component.
Finally, in chapter 6, we unite ASA, tolerance analysis based on matrix transforms, and part mating theory through a case study. We look at an impeller assembly that contains a small number of parts so that our approach's validity can be verified with intuition or manual computations.

Chapter 7 proposes research avenues to further advance knowledge in the area of tolerance consideration for assembly process evaluation, and concludes on this thesis.
Chapter 2

Previous Work
Introduction

Much work has been done in tolerancing. Actually, this denomination, tolerancing, is general enough to encompass more than can possibly be discussed in any single paper or thesis. In this section, we will discuss many topics related to this subject in an effort to better understand the subject of this thesis. We will describe the path that lead efforts to be invested in this area, as well as work of similar nature done elsewhere.

In order to do that, we will try to divide tolerancing into sub-topics. First, in section 2.1, we will talk about feature-based design for assembly. In this method, various features of components and assemblies are defined early in the design process, and their description stored in data structures such that the assembly process can be analyzed rapidly. This analysis then gives feedback to the designer by pointing out features that could be modified in order to, among others, simplify the assembly process, ease quality management, and evaluate capabilities of the actual assembly systems to perform the tasks as desired. In this thesis, dimensional tolerances are viewed as features to both components and assemblies, and the goal is to evaluate their impact on the assembly planning process.

In section 2.2, a particular method to analyze assembly processes will be looked at in more details: assembly sequence analysis (ASA). This discussion, in addition to describe work being done in this area, will show how tolerances can affect the choice of assembly sequences. This relation can be of great use in tolerance assignment and synthesis. The designer can use the information provided by sequence evaluation to
decide what tolerances should be assigned to parts which do not affect directly product's functionality, but do affect the assembly process and its related cost. Similarly, when performing tolerance synthesis, guidelines provided by sequence analysis can help in distributing available tolerance of a compounded dimension to individual dimensions which affect it. These two first section will be limited mostly to describing the work done at Charles S. Draper Laboratories in Cambridge, Massachusetts.

The next section will take a closer look at work that investigates how tolerances can be represented and propagated. First, we look at tolerance representation and solid modelling environments in which this representation lies. Efficient computer-based tolerance analysis can be done only if data is defined and stored in a format compatible with the implemented tolerance analysis method. Next, tolerance analysis approaches are examined. We take a closer look at how geometric variations are actually propagated through an assembly to result in uncertainty on a given dimension. In this section, we look at possible analysis methods and describe in more details a statistical closed-form method using matrix transform representation of tolerances, particularly well suited to assembly sequence evaluation.

Finally, in section 2.4, part mating theory and its requirements are presented. This theory describes the conditions that are necessary to prevent jamming and wedging in the mating of two parts. The usefulness of part mating theory, coupled with tolerance analysis, will be explained in an assembly process evaluation point of view.
2.1 Feature-based design for assembly

Any component or assembly will have certain features associated with them. For example, a shaft might be grooved along its length and this groove can be considered a feature. A feature is any geometric or non-geometric attribute of a discrete part whose presence or dimensions are relevant to the product’s or part’s function, manufacture, engineering analysis, use, etc., or whose availability as a primitive or operation facilitates the design process [De Fazio et al.-'90]. The importance of this feature is relative, depending on the analysis being performed on this part or on the assembly it belongs to. Further more, a feature has properties which can be described by attributes that will complement the description of the designer’s intent. Therefore, feature-based design is a way to describe a component and its corresponding assembly by extending the information beyond the classic geometric shape. Tables 2.1 and 2.2 list various features and attributes one can specify to complement nominal geometric information.

Some work has been done in the area of feature recognition [Requicha- '90] which facilitates the automatic identification of geometric features in a B-spline

<table>
<thead>
<tr>
<th>Typical features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hole</td>
<td>Slot</td>
</tr>
<tr>
<td>Pin</td>
<td>Spline</td>
</tr>
<tr>
<td>Flat</td>
<td>Chamfer</td>
</tr>
</tbody>
</table>

Table 2.1 Examples of features (from [Defazio et al.-'90]).
Table 2.2 Examples of feature attributes (from [Defazio et al.-'90]).

A description of a single part's solid model. One important advantage of using feature-based description however, is the possibility to point out the important areas of a part so that one is not overflowed with relatively meaningless information. Automatic recognition needs knowledge of the design's key features that are crucial to the part's function or manufacturability. Furthermore, this recognition limits itself to geometric features and attributes; acknowledging the existence of any important non-geometric attributes, parallel input of this information will have to be implemented sooner or later in the design process.

Among possible features and attributes, there are some that will have a greater influence on the assembly process. Actually, some features are described specifically for assembly planning purposes, like fixtured surfaces or mating relationships. The description and use of these assembly-related features is the basis of feature-based design for assembly. At Draper Laboratories [De Fazio et al.-'90], a prototype for feature-based design for assembly (FDBA) has been implemented on a Sun workstation. The work underlying this implementation intends to show the potential for feature-based modeling in assembly process design. The actual capabilities of this software tool include:
1 - Generation of a feature-level solid model, including assembly structure definition through identification of mating data (hierarchical structure, orientations, mating features, etc.);

2 - Assembly sequence generation which results in a sequence graph containing all possible sequences;

3 - Sequence graph editing to reduce significantly the number of potential sequences to only the ones that respect sound assembly practices (eliminate unstable assembly states, avoid parallel operations or converging assembly lines in the case of bulky assemblies, prevent difficult assembly actions, etc.);

4 - Assembly process planning and assembly system rate and cost evaluation for the remaining sequences based on factors such as weight, size and task type.

One of the most interesting result of such a system is the ability to return to any place in the design process and make revisions, and rapidly see the effects on the remaining phases. It is this capability that allows for design-for-assembly. Figure 2.1 schematizes how the different modules are inter-related.

![Diagram](image)

**Figure 2.1 Schematization of FDBA prototype (adapted from [DeFazio et al. - '90])**

In the present work, focus is on dimensions and tolerances as they relate to assembly sequence evaluation. On the above-described implementation, this would translate to editing the sequence graph (item 3- above) based on tolerances attributes assigned to both functional and non-functional features. Some work has been done on
specifying and representing tolerances as attributes. Most of these effort are oriented at developing solid modeler's capability to convey variational data, to complement nominal geometric description [Gossard et al.-'88], [Turner and Wosny-'88]. Although these advancements are necessary to the development of integrated tolerance information, we want to focus our efforts on the effects of tolerances on the assembly process. Accordingly, we need not to concern ourselves with immediate modeling environment compatibility, but rather try to find a description format for tolerance attributes which is appropriate to sequence analysis. The proposed representation may then be used as a desired output format from tolerance-based solid modelers. Tolerance representation is a key issue that will be discussed further in this chapter, but first, the next section elaborates on the critical role of geometric variations on the choice of assembly sequences.

2.2 Assembly sequence analysis

In the previous section, we described feature-based design and identified geometric tolerances as attributes to geometric features. In this section, we look at Assembly Sequence Analysis (ASA) as a critical step in assembly process planning. In addition, we establish the importance of variations, which are allowed and controlled using tolerance specifications, in the choice of a good assembly sequence among many possibilities. In this section, we concentrate on work being done at the Charles Stark Draper Laboratory in Cambridge, Massachusetts.
ASA is a rational means of considering all feasible assembly sequences in the sequence selection process [Baldwin-'90]. It represents a method for concurrently considering design and assembly, by allowing to reduce the choice of interesting sequences based on multiple criteria, and by providing feedback on the critical aspects of the design. Assembly sequence analysis is a three step process.

First, all mechanically feasible assembly sequences are generated in the form of a sequence graph. For illustrative purposes, we will refer to a simple product, an impeller assembly, which is shown in figure 2.2 a). In order to obtain the assembly graph, a graphical representation of the contacts between parts, called "liaison diagram" [Bourjault-'84, see Baldwin-'90], is used, figure 2.2 b). In this form, each node represents a part, and each arc or "liaison" represents a contact between parts. From this diagram, and a geometric view of the assembly, a set of precedence relations can be established, which enumerates what components have to be assembled before others to respect both geometric and non-geometric constraints. An actual software implementation [Baldwin-'90] helps in creating these precedence relations by using a number of different automated query answering techniques which minimize required user input. Finally, the sequence graph, figure 2.2 c), generated from the precedence relation sequence representation, shows all possible assembly states and transitions between them.

In this representation, the states are described with a keyed square or rectangle that allows a subdivision for each liaisons from the liaison diagram. A particular assembly state is described with darkened subdivisions marking the liaisons that were made up to, and including, that state. Between states, transitions indicate possible
routes in the graph, and can therefore be viewed as assembly operations between states. The combination of all states and transitions represents all mechanically feasible assembly sequence going from separated individual components (the first state) to the fully assembled product (the last, fully darkened, state). A more complex example is shown in figure 2.3.

Figure 2.2  Assembly sequence generation.
Figure 2.3  A more complicated assembly graph in the editing module (from [Baldwin-'90]).
The second step in ASA is the editing of these sequences based on multiple criteria. The goal here is to prune from the graph representation sequences which choice would involve unpractical assembly process. Actual software implementation allows for editing based on assembly line constraints like the requirement of a sequential assembly line and the avoidance of difficult assembly moves. In addition, [Abell-’89] and [Whipple-’90] considered and implemented criteria such as minimizing the number of refixturing and orientation requirements, and eliminating states with unstable components. It is in this editing step of ASA that consideration of tolerances could further reduce the number of valuable sequences.

Finally, the third and last step of assembly sequence analysis as developed in this laboratory, is evaluating the remaining sequences based on economic considerations. The different assembly systems required to implement one or another sequence in the assembly process are evaluated in terms of complexity and related cost. Furthermore, the list of assembly systems can be limited to existing hardware in order to limit expenditures. Here again, valuable feedback can be used to influence both the choice of an appropriate design and the assembly process planning. Most of the tools available in this area are based on work done by [Gustavson-’88].

As said above, it is in the sequence editing process of ASA that the work on tolerances will find a place for implementation. Figure 2.4 shows how geometric variations can be used as a powerful criterion to choose the best assembly sequence.
Figure 2.4  Assembly sequences affected by tolerances (from [Nevins-Whitney-'89]).
An Example

Among the 5 possible sequences for the impeller on figure 2.2, only a few make sense intuitively: although the full ASA process may not be necessary for this simple example, tolerance analysis can point out limitations of a given sequence or useful additions on the design or the fixture’s specifications. Figure 2.4, taken from [Nevins-Whitney-89] shows the problem with exaggerated deviations for clarity.

In B), the desired sequence is shown. This sequence uses the back of the retainer as a jigging surface. The front, being a functional surface, is well tolerated to be perpendicular to the hole so that the shaft will be perpendicular to the retainer. The back is not a functional surface. In C), the sequence in B) is defeated by the fact that the back of the retainer, not well tolerated, tilts the shaft so far that the impeller cannot be assembled by bringing it down over the center of the fixture. In D), a different sequence is used as a remedy: After the shaft and retainer are mated, the subassembly is transferred to a second fixture, shown in E), that grips the shaft, guaranteeing that it will be upright when the impeller is added.

To detect that the sequence in B) may not work given the original tolerances, we need a tool that will help designers discover such problems. Two responses are possible to these problems: search for a sequence that will work or search for the necessary changes in the original tolerances.

In conclusion, we want to stress the fact that part and assembly variations, generally controlled with standard tolerance specifications, have an impact that goes
beyond functional requirements of the finished product. These variabilities will affect
the assembly process in that each and every assembly step's success can be jeopardize by
the tolerance stack-up from the fixtures and the mating subassemblies and parts, to the
full assembly system.

Figure 2.5 is a flow chart that shows a possible implementation of ASA's sequence
editing based on tolerances.
**Figure 2.5 Flow chart of tolerance-based sequence editing.**
2.3 Tolerance representation and analysis

This section takes a closer look at work that investigates how tolerances can be represented and propagated through assemblies.

Many times, the uncertainty is associated with a functional dimension that will affect the product's behavior in a particular way. The linear distance between between two rails is such a dimension. In other cases, the uncertainty is associated with both the position and orientation of a feature in space. For example, if a peg is to be mated with a hole using a robotic manipulator, knowledge of the variability of the position and orientation of the hole relative to the peg will help in predicting the probability of success of the mating operation. Generally, any tolerance analysis method can give the uncertainty associated with either of the cases, but the approach in modeling the particular case might vary slightly. In this section, we look at possible approaches to represent and analyze tolerances and we evaluate their characteristics with regards to assembly sequence evaluation.

We start by a discussion on the significance of tolerances: what they represent and what they are used for. Then we look at means to represent them, mathematically or informally. Tolerance analysis methods are then evaluated and a specific algorithm is proposed as an effective tool for assembly sequence evaluation.
2.3.1 Significance of tolerances

There exist various approaches concerning the interpretation of tolerances which in turn suggest various ways of representing and communicating them. We can list some of the meanings found in the literature:

- Limits to insure specific functionality
- Expected variations from production processes
- Expected variations from the nature of contacts between parts in assemblies
- Designer's intent, possibly including some of the above
- Standardized specifications, possibly related to some of the above

The variety of significance is such that one cannot automatically assume that there is a direct connection between what is specified and what will actually result from the production process. However, it is a widely embraced goal to bring the two as close to each other as possible. There must be a rational and effective link between specifications and resulting products if control of variations is to be given serious consideration.

For this reason, [Scott-'88] proposes to use an idealized interpretation of tolerances based on the actual manufacturing capabilities. However, this approach is not used extensively because it limits the assignment of tolerances to a choice of manufacturing processes, while in fact, other concerns such as functionality might be foremost in the designer's mind. As a result, there exist no standardized notation that designers, manufacturers, and suppliers alike can use to communicate information in this format to deal with variation control. The information needs to be clearly communicated in order to obtain high quality products.
Consequently, we suggest to maintain widely used standards of tolerance specification and control, but we propose to develop tools to help designers evaluate the impact of their specifications on considerations that go beyond functionality. By forcing the evaluation of manufacturing and assembly capabilities with respect to the assigned tolerances, the engineering community will force the standards to change only if the present notation is inadequate; i.e. only if it cannot properly represent limits on variations that can address both functionality and realistic production processes and controls.

Tolerance specifications, therefore, are viewed in this work as representing the designer's intent, i.e. the designer's intention to allow dimensional variations that take into account production processes, while keeping confidence in the product's correct functionality. Appropriate tolerance analysis and synthesis methods, as well as product and process knowledge applied in a rational manner, will insure the match between specifications and results.

**Component specification versus mating variation**

General dimensioning and tolerancing practices, like the ones described in ANSI Y14.5M-1982 or B4.1-1979 standards, can be used to describe allowed variations for components and assemblies. On a component, a given surface may vary in orientation relative to a datum frame on the same component. This could call for a parallelism tolerance specification from Y14.5 "Geometric Dimensioning and Tolerancing" standard for example. On an assembly, the gap between a shaft and a bearing may be specified
with a tolerance grade as described in B4.1 "Preferred Limits and Fits for Cylindrical Parts" standard for mating parts.

The focus of this thesis is on the assembly process. More specifically, we want to analyze the intermediate assembly steps to evaluate their feasibility in face of uncertainty. In this context, we divide product related variations in two groups: 1- Individual component variations and 2- Mating variations.

Individual component variations, although depending greatly on the manufacturing processes used, will be within the boundaries described by the tolerance specification. Mating variations will result from a combination of the component variations and other factors influencing relative part position (gravity, friction, etc.). For example, the actual position of a peg, once mated with a hole, is considered a mating variation and is uniformly distributed within the area defined by the actual clearance, i.e. by the relative diameters of the peg and hole which are considered component variations.

We consider individual component variations to represent a tolerance zone defined by the specifications, or the intent, while mating variations describe the result of mating relationships given the intended component variations. Hence, for tolerance analysis purposes, we disregard any tolerance specifications on dimensions relating two or more features on different parts. This is justified by our desire to analyze intermediate steps of the assembly process, where the possible relative position and orientation of features in the subassembly are more important than dimensions on the final assembly, which we assumed has already been verified for functionality.
Mating variations versus assembly variations: the assembly system

From the assembly process analysis point of view, it is necessary to further clarify the meaning of "mating variations". Mating variations are involved in the relationships between two already mated parts. When considering a peg mated with a hole, "mating variations" describes the conditional position of the peg in the hole, given the intended variation of each component diameter. Hence these variations are useful when performing tolerance analysis to estimate stochastically the position and orientation of a feature in a sub-assembly. In our type of analysis, this feature is typically the one next to be mated with the component manipulated by an automated assembly system.

On the other hand, we consider "assembly variations" to describe all variations that can influence the mating process, i.e. parts to be mated. This includes, in addition to individual component and mating variations, errors which find their source in the complete assembly system. Assuming the use of a simple manipulator-based system, assembly system variations can be classified in three sources: 1- the manipulator and the various factors that influence its precision in trajectory, position and orientation; 2- the fixtures which are created to facilitate the assembly process by positioning parts and sub-assemblies, for both grasping of components to assemble and maintaining parts on which assembly will occur; and 3- the grasping operation itself which can introduce variations that depend on the surface to be grasped, the surface in contact with the feeding fixture if there is one, and on the feeding fixture itself. For comprehensive assembly process evaluation, the method chosen to represent and analyze tolerances in
assemblies should also be capable of representing and propagating errors in the assembly system.

2.3.1 Tolerance representation

Most of the work done in the area of tolerance representation has been motivated by the appearance and use of solid modelers ([Bernstein and Preiss-'89], [Gossard et al.-'88], [Requicha and Chan-'86], [Roy and Liu-'88], [Shah and Miller-'90], [Turner and Wozny-'90]).

Any computer-based tool will require some level of mathematical representation so that information can be stored in the various databases, and used for different applications such as on line quality control and tolerance analysis. Many approaches involve the reformulation of tolerances in a mathematical format more compatible with the parameters used in solid modelers (vertices, lines, polygons, etc.), but the semantics of the Y14.5 standard are not always preserved.

[Guilford and Turner-'91] have suggested that the underlying mathematical approaches to tolerance representation can be considered in three broad headings: parameter space, offset solid, and feasibility space approaches.

In this classification, one of the more popular approaches is termed a parameter space approach ([Hillyard and Braid-'78], [Light and Gossard-'82], [Martino and Gabriele-'89]). This approach views the object shape as being controlled by a set of
parameters. In a CSG-based (Constructive Solid Geometry) or a feature-based system, these are the parameters used in the interactive specification of the object (vertices, lines, etc., or slots, chamfers, etc.). In a dimension-driven boundary based system, the parameters are the part dimensions. Tolerances are viewed as limits on the parameter values. From a mathematical point of view, the tolerance limits define a bounded region of a parameter space. Although this approach can lead to simpler tolerance analysis algorithms, the resulting semantics may be quite different from the ANSI standard, which views a tolerance as a specification of a tolerance zone.

*Offset solids* is another approach that has been proposed mostly by [Requicha-'83]. Here, as in the ANSI standard, tolerance zones in the Cartesian space are defined and act as constraints on the part geometry. However, the zones are positioned at the boundaries of components and do not allow to easily specify a tolerance zone for center-planes and axes, as often is the case in Y14.5.

A good example of a method based on the notion of *feasibility space* is found in [Turner and Wozny-'90]. In this approach, a variational model is derived from a boundary model of the nominal part. The authors have separated the various elements in three types that together specify the allowed variations of the part or assembly: 1. The "design variable" is defined as the specification of highest level. It has limits specified by the design constraints and is an aggregate of the individual part dimensions (ex. clearance between a peg and hole, a total surface area); 2. "Tolerance variables", with limits specified by tolerance constraints of the type ANSI Y14.5, describe simple geometric properties of a single part; And finally, 3- "model variables" are identified to allow a variational model to be defined in the CAD database. Constraints on many of
those variables, usually compatible with tolerance constraints, will ultimately define a design constraint. The vector space of the model variations is called the M-space, and the boundaries in that space represent the various constraints on the model variables. Both conventional and geometrical tolerances can be represented with the M-space, including most notably form tolerances. High order polynomials of model variables are suggested to represent form tolerances.

A mathematical representation of tolerances using homogeneous transformation matrices has been evaluated by [Bernstein-’89]. In the above classification of mathematical approaches, this one can be viewed as a combination of the parameter space and the offset solids approaches. In this presentation, a part boundary is divided into separate topological entities (vertices, edges, faces) which are treated as free rigid bodies and are subject to kinematic analysis. In this context, each entity has 6 associated degrees of freedom (3 translations along x, y, z and 3 rotations about the same axes). These degrees of freedom are restrained by dimensioning constraints and are relaxed by tolerancing constraints. The dimensioning and tolerancing (D&T) design information describes spatial relations and is represented as a constraint network. [Bernstein-’89] has addressed in detail the issue of validity of tolerance schemes using this approach, but has not addressed the semantic integrity with Y14.5. For example, a distance constraint between two planes is represented with a single translation constraint and its associated linear tolerance constraint. In reality, the same distance constraint as expressed by the planar size specification of Y14.5, should be augmented by kinematic angular variations since the planes can vary in orientation while still respecting the D&T constraints. Also, although it is named the "constraint propagation approach", there is no formal
mathematical solution proposed for tolerance analysis within single parts or through assemblies.

Finally, [Shah & Miller-'90] have proposed a product data structure that contains part definition using three types of information: a solid model to define the nominal geometry, a feature model to capture the semantics of the geometry, and variational attributes to specify tolerances. This structure incorporates associations among tolerances, geometric entities, features and feature attributes. Although it seems to be a comprehensive product definition database, useful for integrating and automating Computer Integrated Manufacturing (CIM) applications, variational information often needs more specific mathematical, representational and computational models to provide facilities for design or tolerance analysis, or for process planning. This way of encoding tolerances using objects, object attributes, and associated relations has been extended by [Guilford & Turner-'91] to express in a more rigorous fashion the semantic of Y14.5.

In many of these approaches, the ultimate goal seems to be the actual representation of tolerances. These representations must often be adapted for compatibility with specific analysis applications and methods. Furthermore, one realizes that Y14.5 is not very rigorous and it becomes difficult to find a single formal representation that will convey all pertinent information.

In modeling assemblies for computer-based systems, tolerance representation can be viewed as subordinate to assembly representation, since tolerance constraints often apply to aggregate dimensions, controlling the relative position of two or more features on different parts of the assembly. Assembly modeling has been addressed in
various ways by [Lee & Gossard-'85], [Turner-'90], [Wang & OzSoy-'90] and [Scott-'88], among others. A noticeable problem in this area is the lack of knowledge on the physics of mating parts. Accumulated knowledge and associated standards are very limited; Y14.5 is suitable for individual components but lacks in assembly specification capability.

2.3.2 Tolerance analysis

Dimensions and tolerances selection is accomplished in a process that includes four general steps [Roy et al-'91]: 1- Identification of functional requirements; 2- Identification of datum reference frames, features, and relation between features that influence both functional and assembly requirements; 3- Development of functional equations, or design functions, in which the product dimensions and tolerances affecting the requirements are the independent variables, and the functional requirements are the dependent variables; 4- Determination of economical solutions to these design functions by means of tolerance analysis and synthesis. In general, the quantification of functional requirements and the formulation of the design functions are complex procedures; the latter is necessary to perform tolerance analysis and synthesis, which in turn helps in the quantification process by means of trial and problem feed-back. This section focuses on the analysis methods, often closely related to the way tolerances are represented.
Possible approaches

There are many factors by which tolerance analysis approaches may be classified. These include "worst-case" or statistical; simulated or analytical; one, two, or three-dimensional space; and combinations of these and other factors.

Worst case

A simple approach is to consider D&T information as nominal values with upper and lower limits. This conservative method, often called worst-case, results in expensive conclusions since it is unlikely that all dimensions take their extreme values at the same time, which would result in excessively tight tolerance specifications.

Statistical

A more realistic approach considers the statistical distributions for each independent variable and their cumulative effect on the functional requirements through the design function. In statistical approaches, the distribution of each individual dimension can be found through measurement for a given manufacturing process, or by using knowledge accumulated from previous experience. However, many factors often prevent this information from being accessible to designers and assumptions on some of the distributions are sometimes made. A common approach is to consider normal distributions for all dimensions. This assumption may be justified for design functions that are dependent on many individual dimensions because of the central limit theorem. This theorem states that the result of the convolution of many distributions (usually
greater than four) approaches the normal distribution, regardless of the shape of the individual dimensions [Bjorke-'89]. The assumption of normal distributions can simplify the error propagation algorithm [Veitschegger and Wu-'86]. On the other hand, a more precise approach may be justified if the information is readily available, and if the analysis applies mostly to smaller assemblies with very few pertinent dimensions. [Bjorke-'89] and [Greenwood and Chase-87] have proposed models taking into account skewness and bias into consideration, but they have limited the analysis to a one-dimensional space.

Simulated

Many methods have been proposed for the actual propagation of errors, or tolerance analysis. One of the simplest technique is the Monte-Carlo simulation. Applicable to many representations of tolerances and solid modeler data structure, this technique uses simulated instances of the individual dimensions (or boundaries in a boundary representation, or parameters in a parameter-space approach, etc) which are taken from the specified distributions. Through the design function or an assembly representation of the product, each instance is accumulated and the corresponding instance of the final dimension is found. Repeating the process many times, the shape of the final dimension's distribution can be inferred. This technique is used in, among others, [Pandit and Starkey-'88], [Scott and Gabriele-'89], and [Turner et al.-'87]. The main drawback of this approach is the computational complexity which increases significantly the time required to do such analysis; "what-if" queries can become burdensome, and analysis of many possible assembly sequences in terms of tolerances becomes highly time-consuming.
**Deterministic and Analytical**

Many deterministic approaches have been proposed and the required time to perform tolerance analysis is often reduced compared to simulation approaches. A widely accepted deterministic approach has been advocated by [Hillyard and Braid-'78], and is based on the variational geometry representation of tolerances, classified under the feasibility space category of the preceding section. Although the authors have developed this approach for analyzing the inconsistency in the specification of D&T in computer-aided mechanical design, the concept is compatible with tolerance analysis. The geometry of an object is dictated by by the specified dimensions which are regarded as constraints between, in their example, vertex geometries. The nominal part geometry is modified by changing the suitable dimension constraints. The relationships between the dimensions of the part and its geometry as defined in the database (vertex coordinates) constitute the required set of constraint equations. With a given, user-defined dimension scheme, the solution of the equations yields the exact vertex geometry, while another system of equations relating the variations in positions of the vertices yields the part tolerances.

[Gossard et al.-'88], [Light and Gossard-'82] also present the use variational geometry, or constraint-based approach, to extend the compatibility with interactive solid modelers, while [Martino and Gabriele-'89] applies it to automatic formulation of design functions. [Fleming-'88] presents a similar concept and investigates the uncertainty in assemblies of rigid parts by building networks of tolerance zones and datums that include mating variations as slop variables. Unfortunately, in constraint-based approaches, the resulting equations and inequalities often constitute a large system.
of non-linear equations which is difficult to solve, thus limiting this technique in terms of practical model size and complexity of the product geometry. Accordingly, the inclusion of the assembly system in the design function, for assembly sequence analysis, would be overly complex, and would require the assembly system to be defined with the same format as parts in the solid modeler.

[Bjorke-'89] elaborates a methodology to find the statistical influence of each individual dimension on what he calls the sum dimension, which is the functional dimension, or the one of interest in the analysis. To do so, he mathematically transforms the statistical parameters to be summed, from the specific directions of the individual dimensions to the direction of the sum dimension. The expression of the sum of all dimension parameters is the tolerance chain, which is equivalent to the design function. The publication presents many useful mathematical formulations of mating variations (gaps, rotational parts, etc). Although formal, the methodology leads to information on variability only in a one dimensional space (what has been identified by the user as the sum dimension), and Y14.5 is not dealt with.

Mathematically programmable

Another frequent method of tolerance analysis is to formulate the constraints in such a way that mathematical programming techniques can be applied. In this context, the objective function is the design function, and the individual dimensions are the decision variables. The design function can be minimized or maximized to find the worst case solutions given the tolerances (constraints) on the dimensions. [Turner and Wozny-'88] discusses such an optimizing approach. [Parkinson-'84] develops a
mathematical programming formulation that preserves the statistical nature of the problem by using the standard deviations of the individual dimensions as the decision variables.

**Commercial packages**

[Turner and Gangoiti-'91] have reviewed many commercial packages and a brief summary of their findings is appropriate here. All packages make some use of geometric models. They all make simplifying assumptions that need to be well understood to prevent false conclusions. Among those are included the non-correlation of design variables, and limited or incomplete interpretations of standard tolerances.

The approaches described in their paper are divided in two types: procedural and declarative. They refer primarily to the way the geometry is modeled but this factor has a significant impact on the tolerance analysis capabilities. For example, in the declarative approach, in which point, line and curve primitives are specified along with dimensions, a 2-D geometric size tolerance will be represented by a single dimension variable, as opposed to a scenario where both size and orientation can be properly represented. In the procedural approach, the user must have a good understanding of the geometry and the modeling primitives used (blocks, cylinders and spheres as in CSG or chamfers, through-holes, etc. in feature-based models) since the variational coverage of the part depends on the choice of primitives and on the sequence in which they were put together to design the part in the first place.
Concerning tolerance analysis per say, many of the packages use a linearized sensitivity analysis. In this method, a sensitivity factor is determined for each model variables and defines the sensitivity of the design function to a variation of that variable. A linear design function is then formed by summing the products of the sensitivity factors multiplied by their corresponding variable’s range or variance, depending if a worst case or a statistical analysis is desired. Hence, the term “variation analysis” may be more appropriate than a more stringent tolerance analysis, the latter being sometimes more complex and burdensome. Monte-Carlo simulation is used in many packages.

On the subject of assembly modeling, all packages treat mating conditions as if they could be satisfied exactly (i.e. nominal parts sharing common edge or surface) which may lead to false conclusions by ignoring, among others, unplanned interferences with other parts of the assembly, or clearances.

**Approach proposed: Closed-form algorithm in 3-D using matrix transforms**

This section presents work done by [Jastrzebski-'91]. The summary of the derivation that led to the closed-form analytical solution to the problem of tolerance propagation in three dimensions is presented from his thesis. His work is largely based on [Veitschegger and Wu-'86], which develops a similar algorithm to propagate errors in kinematic manipulators.
**Derivation outline**

In this work, tri-variable normal probability density distribution of individual tolerances has been assumed. The geometric interpretation of the resulting error distribution of a point position in space is an ellipsoid with the highest density at the nominal position of the point. Due to the angular errors, the axes of the propagated ellipsoid in space might not be parallel to the directions of the base coordinate frame axes. However, all the information necessary to describe the ellipsoid in space is contained in the 3x3 covariance matrix of the resulting tri-variable probability density function. Therefore, after the covariance matrix of the function is determined, the matrix eigenvalues can be computed. The square roots of the obtained three eigenvalues are standard deviations in the principal directions of the ellipsoid. The orientation of the probability density ellipsoid in space is determined by the eigenvectors of the covariance matrix. The eigenvectors are aligned with the principal axes of the ellipsoid.

**Assumptions**

Because of the inherent complexity related to the issues of three dimensional tolerancing, it was necessary to make simplifying assumptions.

1. The propagation of tolerances in 3-D space is treated as propagation of errors during transformations of Cartesian frames. Cartesian frames are attached to a specified point on the surface of a part. There exist 6 degrees of freedom for dimensional transformations of frames or their small errors. The capital letters X, Y, Z signify
linear translations; and the lower case letters \( x, y, z \) represent rotation angles around respective axes.

2. The order of rotation is important and consistency must be maintained through the process of \( N \) transformations. In this work, the following order is assumed and maintained:

\[
T = T_{\text{Trans}} T_{\text{Rot}X} T_{\text{Rot}Y} T_{\text{Rot}Z},
\]

where \( T \) is a general form of a transformation matrix for transformation from one point in Cartesian space to the next. These rotations are taken about the original axes.

3. We assume normal trivariable (Gaussian) distribution of error. Propagation of normal density ellipsoid over \( N \) frames results in the output normal density ellipsoid (possibly of changed size, density and orientation in space).

4. The specified, design tolerance range for any single dimensional quantity (length or angle) is equivalent to \( 6\sigma \) (\( \sigma \) = standard deviation). Therefore, any specified dimension \( a \) with specified tolerance range \( \Delta a \) (\( \Delta a = 3\sigma \)) will be included in the interval \( (a - \Delta a, a + \Delta a) \) with the probability of approximately 99.7\% (in one dimensional space).

5. The probabilistic events are independent. Any single transformation of any frame does not, in any way, affect transformation of the next frame (within their prescribed variances limits).
Derivation

The intermediate objective of this derivation is the development of the tri-variable normal density function for the Mth transformation frame. The ultimate objective is the description of the location, the orientation and the axes' lengths of the two probability density ellipsoids: one for the translation (or linear) errors, and one for the rotational (or angular) errors. Considering for now only the translation error probability, the density function argument is the 3x1 differential translation vector dp. The function describes the distribution of the Mth frame origin position, which is the last frame in the chain of transformations, in the three dimensional space:

\[
    f(dp) = \left(2\pi\right)^{-3/2} |V_p|^{-1/2} \exp\left[-0.5((dp)^\dagger V_p^{-1}(dp))\right]
\]

(2.1)

where \( V_p \) is a 3x3 covariance matrix of the dp vector and the symbol \( \dagger \) means vector transpose. The eigenvalues of the \( V_p \) matrix are variances in the principal directions of the probability density ellipsoid. The directions of the eigenvectors of the matrix determine the ellipsoid orientation in space.

A similar density function can be found for the rotational error probability of the same point (or feature), to which the argument is the 3x1 differential rotation vector \( \delta \), and the associated 3x3 covariance matrix \( V_\delta \).

1st Error transformation between i-1 and ith frame

The \( i \) index denotes a frame number in a chain of \( N \) transformations. A transformation between \( i-1 \) and the ith frame is defined as \( A_i \). Since the order of transformations is
Important, we assume that the translation occurs first, followed by the rotations with respect to X, Y, Z axes respectively. Therefore,

\[ A_i = [A_i]_{\text{Transl}}[A_i]_{\text{Rot}}, \]

where

\[ [A_i]_{\text{Rot}} = [A_i]_{\text{RotX}}[A_i]_{\text{RotY}}[A_i]_{\text{RotZ}}, \]

and the \( A_i \) transform matrices are:

\[
[A_i]_{\text{Transl}} = \begin{bmatrix} 1 & 0 & 0 & X_i \\ 0 & 1 & 0 & Y_i \\ 0 & 0 & 1 & Z_i \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

\[
[A_i]_{\text{RotX}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_{xi}) & -\sin(\theta_{xi}) & 0 \\ 0 & \sin(\theta_{xi}) & \cos(\theta_{xi}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

\[
[A_i]_{\text{RotY}} = \begin{bmatrix} \cos(\theta_{yi}) & 0 & \sin(\theta_{yi}) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta_{yi}) & 0 & \cos(\theta_{yi}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

\[
[A_i]_{\text{RotZ}} = \begin{bmatrix} \cos(\theta_{zi}) & -\sin(\theta_{zi}) & 0 & 0 \\ \sin(\theta_{zi}) & \cos(\theta_{zi}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

where \( X_i, Y_i, Z_i \) are translations in the respective directions; and \( \theta_{xi}, \theta_{yi}, \theta_{zi} \) are rotations around respective axes. The \( i \) index refers to the \( i \)th frame transformation. Therefore, the general form of the \( A_i \) transform can be written in the following simplified homogeneous matrix form:
\[ A_i = \begin{bmatrix} R_i & p_i \\ 0 & 1 \end{bmatrix}, \]  

(2.2)

where \( R_i \) is a 3x3 rotation matrix, and \( p_i \) is a 3x1 translation vector, both taken with respect to the \( i-1 \) frame.

The actual position of the \( N \)th frame is determined by the transform matrix \( T_{N^C} \):

\[ T_{N^C} = T_N + dT_N, \]  

(2.3)

where \( T_N \) is the nominal transform matrix (if no errors were present). \( dT_N \) represents the total differential change transform matrix of the \( N \)th frame position and orientation. The \( dT_N \) matrix can be written in the following form:

\[ dT_N = [\delta T] T_N, \]

where \( \delta T \) can be defined as the differential error matrix of the total transformation chain. For any transformation between \( i \) and \( i-1 \) frame, it is true that:

\[ A_i + dA_i = [I + \delta A_i] A_i. \]

\[ A_1 + \delta A_1 \]

\[ A_1 \]

\[ \delta A_1 + I \]

\( \delta A_i \) is defined as the 4x4 homogeneous error matrix transform with respect to the \( i-1 \) frame. And therefore, the differential change matrix \( dA_i \) (4x4) for the \( N \)th frame is:

\[ dA_i = [\delta A_i] A_i. \]  

(2.4)

Because \( \delta A_i \) transform matrix describes an error between two frames positions and orientations, it is important to find its general form. \( \delta A_i \) can be represented in the following form of a homogeneous matrix:
\[ \delta A_i = \begin{bmatrix} \delta R_i & d_i \\ 0 & 1 \end{bmatrix}, \]

where \( \delta R_i \) is the 3x3 error rotation transform matrix with respect to the \( i-1 \) transformation. Similarly, \( d_i \) is the 3x1 error translation vector. The following flow of the derivation in this section aims at finding the general form of \( d_i \) error translation vector, and the real eigenvector of the \( \delta R_i \) error rotation transform matrix (defined as \( \delta_i \)).

Equation (2.3) can be rewritten as

\[ T_N + dT_N = [A_1 + dA_1][A_2 + dA_2]...[A_N + dA_N], \text{ or } T_N + dT_N = \prod_{i=1}^{N} [A_i + dA_i]. \]

The differential change matrix \( dA_i \) may be obtained by differentiating the matrix \( A_i \) with respect to the three scalar translations \( (X_i, Y_i, Z_i) \), and the three scalar rotations \( (\theta_{x_i}, \theta_{y_i}, \theta_{z_i}) \):

\[ dA_i = \left[ \frac{\partial}{\partial x_i} A_i \right] dX_i + \left[ \frac{\partial}{\partial y_i} A_i \right] dY_i + \left[ \frac{\partial}{\partial z_i} A_i \right] dZ_i + \left[ \frac{\partial}{\partial \theta_{x_i}} A_i \right] d\theta_{x_i} + \left[ \frac{\partial}{\partial \theta_{y_i}} A_i \right] d\theta_{y_i} + \left[ \frac{\partial}{\partial \theta_{z_i}} A_i \right] d\theta_{z_i} \]

Let us rewrite the above equation in a different form:

\[ dA_i = ([D_i]_{\text{Trans}X} dX_i + [D_i]_{\text{Trans}Y} dY_i + [D_i]_{\text{Trans}Z} dZ_i + [D_i]_{\text{Rot}X} d\theta_{x_i} + [D_i]_{\text{Rot}Y} d\theta_{y_i} + [D_i]_{\text{Rot}Z} d\theta_{z_i}) A_i \] (2.5)

The \( [D_i]_{\text{Trans}X} \) 4x4 matrix is defined as:

\[ [D_i]_{\text{Trans}X} = \left[ \frac{\partial}{\partial x_i} A_i \right] A_i^{-1}, \]

and the remaining five 4x4 \( D_i \) matrices are computed analogously.
Comparing (2.4) and (2.5) we can write the $\delta A_i$ transform in the following form:

$$\delta A_i = [D_i]_{\text{Transl}} X dX_i + [D_i]_{\text{Transl}} Y dY_i + [D_i]_{\text{Transl}} Z dZ_i
+ [D_i]_{\text{Rotl}} d\theta_x + [D_i]_{\text{Rotl}} Y d\theta_y + [D_i]_{\text{Rotl}} Z d\theta_z.$$

To find the general form of the error transform matrix $\delta A_i$, first we must derive the inverse of matrix $A_i$, and the partial derivatives of $A_i$ with respect to $X_i$, $Y_i$, $Z_i$, $\theta_x$, $\theta_y$, $\theta_z$. This task requires a rather substantial amount of time and space, and the detailed derivation is not shown here.

The 3x1 $d_i$ vector is the last column of the $\delta A_i$ matrix (excluding the corner element 1). The differential error notations ($d$) can be replaced by the actual (small) errors ($\Delta$) occurring during linear or angular transformations. Therefore, $d_i$ can be written in the following convenient form, with contributing errors isolated:

$$d_i = m_{1i}\Delta X_i + m_{2i}\Delta Y_i + m_{3i}\Delta Z_i + m_{4i}\Delta \theta_x + m_{5i}\Delta \theta_y + m_{6i}\Delta \theta_z.$$

where $m_{1i}$, $m_{2i}$, $m_{3i}$, $m_{4i}$, $m_{5i}$, $m_{6i}$ are all 3x1 vectors contributing to the transformation translation errors occurring between the $i$-1 and the $i$th frames (their expanded forms are shown in the Appendix A). Notice that not only linear errors $\Delta X_i$, $\Delta Y_i$, $\Delta Z_i$, but also the angular errors $\Delta \theta_x$, $\Delta \theta_y$, $\Delta \theta_z$, contribute to the final position error of the $i$th frame.

The 3x1 eigenvector of the $\delta R_i$ matrix can be found and denoted as $\delta_i$. The error rotation vector $\delta_i$ can be written in the following form:

$$\delta_i = m_{7i}\Delta \theta_x + m_{8i}\Delta \theta_y + m_{9i}\Delta \theta_z.$$
The $m_7i, m_8i, m_9i$, are 3x1 vectors contributing to the angular error transformation occurring between the $i-1$ and the $i$th frame. Notice, however, that only the angular errors $(\Delta \theta_x, \Delta \theta_y, \Delta \theta_z)$ affect the $\delta_i$ vector.

2. Total error transformation for a chain of $N$ frames

It can be shown [Veltschegger and Wu-86] that the combined vectors $d$ and $\delta$ associated with the total transformation of $N$ frames, can be computed as follows:

$$d = \sum_{i=1}^{N} \left[ R_{i-1} d_i + p_{i-1} \times [R_{i-1} \delta_i] \right], \quad (2.6)$$

$$\delta = \sum_{i=1}^{N} R_{i-1} \delta_i, \quad (2.7)$$

where $R_{i-1}$ is 3x3 rotation matrix and $p_{i-1}$ is 3x1 translation vector, both obtained from the $A_{i-1}$ transform (equation 2.2). The $d$ and $\delta$ vectors can be represented in the following matrix form:

$$\begin{bmatrix} d \\ \delta \end{bmatrix} = \begin{bmatrix} W_1 \\ 0 \end{bmatrix} \Delta X + \begin{bmatrix} W_2 \\ W_3 \end{bmatrix} \Delta Y + \begin{bmatrix} W_4 \\ 0 \end{bmatrix} \Delta Z + \begin{bmatrix} W_5 \\ W_6 \end{bmatrix} \Delta \theta_x + \begin{bmatrix} W_7 \\ W_8 \end{bmatrix} \Delta \theta_y + \begin{bmatrix} W_9 \\ 0 \end{bmatrix} \Delta \theta_z, \quad (2.8)$$

where

$$\Delta X = [\Delta X_1, \ldots, \Delta X_N]^T \quad \Delta \theta_x = [\Delta \theta_{x1}, \ldots, \Delta \theta_{xn}]^T$$

and $^T$ means vector transpose. The $W$ vectors are all 3x$N$ vectors and (from (2.6), (2.7) and (2.8)) their $n$th columns can be written:

$$W_{1i} = R_{i-1} m_{1i}, \quad W_{2i} = R_{i-1} m_{2i}, \quad W_{3i} = R_{i-1} m_{3i}$$

$$W_{4i} = R_{i-1} m_{4i} + p_{i-1} \times [R_{i-1} m_{7i}], \quad W_{5i} = R_{i-1} m_{7i}$$

Previous Work
\[ W_{6i} = R_i - 1m_{5i} + p_{i-1} \times [R_i - 1m_{8i}], \quad W_{7i} = R_i - 1m_{8i} \]
\[ W_{8i} = R_i - 1m_{6i} + p_{i-1} \times [R_i - 1m_{9i}], \quad W_{9i} = R_i - 1m_{9i}. \]

Because the argument of the probability density function (1) is the 3x1 differential translation vector dp, we should find its general form. The actual (correct) position of the last Nth frame is described by the 3x1 vector pt_N^C:

\[ pt_N^C = pt_N + dp, \]

where the pt_N vector describes the nominal position (no errors present) of the Nth frame. The vector dp can be written as [Veltchevsky and Wu-'86]:

\[ dp = \delta \times pt_N + d, \quad (2 \cdot 9) \]

where \(d\) is obtained from (2.6) and \(\delta\) from (2.7). The above equation shows clearly that the total error of any chain of transformations in space depends both on linear and angular errors (tolerances).

By substituting (2.8) into (2.9), the dp vector can be rewritten in the form:

\[ dp = W_1 \Delta X + W_2 \Delta Y + W_3 \Delta Z + W_{10} \Delta \theta_x + W_{11} \Delta \theta_y + W_{12} \Delta \theta_z, \]

where again \(W_{10}, W_{11}, W_{12}\) are all 3xN matrices, and their \(i\)th columns are:

\[ W_{10i} = W_{5i} \times pt_N + W_{4i}, \quad W_{11i} = W_{7i} \times pt_N + W_{6i}, \quad W_{12i} = W_{9i} \times pt_N + W_{8i}. \]

3a Covariance matrices of the triviable normal distribution for the dp and \(\delta\) vectors

\(\Delta X, \Delta Y, \Delta Z, \Delta \theta_x, \Delta \theta_y, \Delta \theta_z\) are Nx1 vectors of designer specified tolerance ranges. For this paper, it has been assumed that \(\Delta X, \Delta Y, \Delta Z, \Delta \theta_x, \Delta \theta_y, \Delta \theta_z\) are six
independent, N-variables with zero means and normal distribution with the following properties:
- \( \mathbf{V}_X = \text{variance of } \Delta X = \text{a NxN diagonal matrix with components (var}X_1, ..., \text{var}X_N) \),
where \( \text{var}X_i = \text{variance of } \Delta X_i \); and similarly for all vectors up to
- \( \mathbf{V}_\theta = \text{variance of } \Delta \theta = \text{a NxN diagonal matrix with components (var}\theta_1, ..., \text{var}\theta_N) \).
where \( \text{var}\theta_i = \text{variance of } \Delta \theta_i \).

The 3x3 covariance matrix \( \mathbf{V}_P \) of the dp differential translation vector can be computed from the following formula:
\[
\mathbf{V}_P = W_1 \mathbf{V}_X W_1^\dagger + W_2 \mathbf{V}_Y W_2^\dagger + W_3 \mathbf{V}_Z W_3^\dagger \\
+ W_{10} \mathbf{V}_\theta W_{10}^\dagger + W_{11} \mathbf{V}_\theta W_{11}^\dagger + W_{12} \mathbf{V}_\theta W_{12}^\dagger
\]

Similarly, the expression for the 3x3 covariance matrix of the total error rotation vector \( \delta \) is:
\[
\mathbf{V}_\delta = W_5 \mathbf{V}_\theta W_5^\dagger + W_7 \mathbf{V}_\theta W_7^\dagger + W_9 \mathbf{V}_\theta W_9^\dagger.
\]

4° Probability density ellipsoid

Knowing the \( \mathbf{V}_P \) covariance matrix, we obtain the ellipsoid of probability density in space, which is fully described by the \( \mathbf{V}_P \) matrix:
\[
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^\dagger \mathbf{V}_P^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = t^2
\]

There is an infinite number of orientations in space that the ellipsoid can assume. We would like to transform it, so that we can use the ellipsoid axes as the coordinate axes.
In other words, we need to diagonalize the matrix $V_p$. The diagonalization can be done by finding the eigenvalues of the matrix. The characteristic equation is:

$$|V_p - \lambda I| = 0,$$

where $\lambda$ means an eigenvalue of $V_p$, and $I$ is a 3x3 identity matrix. The three eigenvalues, obtained from the characteristic equation, are equal to variances in the principal directions of the ellipsoid:

$$\lambda_1 = \sigma_1^2, \lambda_2 = \sigma_2^2, \lambda_3 = \sigma_3^2,$$

where $\sigma_1, \sigma_2, \sigma_3$ are standard deviations in the principal directions. Therefore, after transforming the orientation of the principal axes of the ellipsoid to the position of the base coordinate frame, we obtain the following equation:

$$\frac{y_1^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} + \frac{y_3^2}{\sigma_3^2} = I^2,$$

where $I$ is a constant.

By another transformation,

$$z_i = \frac{y_i}{\sigma_i},$$

we transform the equation of the ellipsoid to a sphere with a radius $r = I$.

$$z_1^2 + z_2^2 + z_3^2 = r^2$$

The probability of finding $z$ inside the hypersphere is [Bryson and Ho-75]:

---

Chapter 2
\[
\int \int \int \frac{1}{(2\pi)^3} \exp \left( -\frac{1}{2} (z_1^2 + z_2^2 + z_3^2) \right) dz_1 dz_2 dz_3,
\]

where the integration is performed over the volume of the hypersphere. Therefore, the probability of finding a point inside the ellipsoid:

\[
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^\dagger V_p^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = l^2
\]
is the same as the probability of finding it within the hypersphere of radius \( r = l \).

\[
\frac{1}{(2\pi)^3} \int_0^l \exp \left( -\frac{1}{2} r^2 \right) f(r) \, dr
\]

or, in the solved form [Bryson and Ho'75]:

\[
\sqrt{\frac{2}{\pi}} \int_0^l \exp \left( -\frac{1}{2} r^2 \right) r^2 \, dr = \text{erf} \left( \frac{l}{\sqrt{2}} \right) - \sqrt{\frac{2}{\pi}} l \exp \left( -\frac{1}{2} l^2 \right)
\]

(2.10)

where \( \text{erf} \) signifies the error function and is defined by the following equation:

\[
\frac{1}{2} \text{erf} \left( \frac{x}{\sqrt{2}} \right) = \int_0^x \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} t^2 \right) \, dt.
\]

In some cases, the three dimensional ellipsoid can degenerate to two dimensional ellipse. This occurs, for example, when the Z is the axis of assembly, and the tolerances are specified in X and Y linear directions only, and the angular tolerances are zero. It is useful to tabularize the values of probabilities for \( l = 1, 2, 3 \); for \( n = 1, 2, 3 \) dimensions. These values are often called the one-, two-, or three-sigma probabilities:
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>$l = 1$</th>
<th>$l = 2$</th>
<th>$l = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$</td>
<td>0.683</td>
<td>0.955</td>
<td>0.997</td>
</tr>
<tr>
<td>$n = 2$</td>
<td>0.394</td>
<td>0.865</td>
<td>0.989</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>0.200</td>
<td>0.739</td>
<td>0.971</td>
</tr>
</tbody>
</table>

Requirements for tolerance representation

As mentioned above, $\Delta X, \Delta Y, \Delta Z, \Delta \theta_x, \Delta \theta_y, \Delta \theta_z$ are the designer specified tolerance ranges. They constitute the only information, along with the nominal values, needed to find the size, position, and orientation of the two final ellipsoid of probability density (one for translational, one for rotational errors).

The nominal values can be incorporated in a homogeneous matrix format as in equation 2.2, while the associated error's usage are best seen through equation 2.5, where the differential notation (d) can be replaced by the actual (small) errors ($\Delta$). Standard tolerance zones and mate variations described in this format will permit the use of this closed-loop algorithm to propagate errors in three dimensions.
2.4 Part mating theory

There is no questions in our mind that the understanding of intricacies of mating parts are necessary to a comprehensive analysis of assembly processes. All the previous sections of this chapter have described work aimed at the advancement of assembly considerations in design decisions. Furthermore, this thesis is developing a representation of tolerances and variations so that an effective closed-loop tolerance analysis algorithm can be used to evaluate many sequences very rapidly. After the propagation of tolerances, this evaluation is focused on the probability of success of critical mating steps through out the assembly sequences: a subject where part mating theory is fundamental.

Important work on the theory of mating parts has been done by, among others, [Nevins and Whitney-89], and has resulted in useful knowledge to help evaluate the probability of success of mating parts. Although this work has addressed many types of mating relationships, the mating of round pegs and holes has been investigated in most details. Of these efforts, we will present important results that are directly related with the information obtained from the tolerance analysis method we are working toward. Many of the following descriptions are taken from [Nevins and Whitney-89].

Phases of mating peg and holes

Figure 2.6 shows the phases of mating peg and holes. The model is two dimensional, although we should remember that assembly is three dimensional in general. The figure defines five typical phases of an assembly: approach, chamfer
crossing, one-point contact, two-point contact, and line contact. Not every assembly contains all of these phases but most do.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.6.png}
\caption{Phases of mating pegs and holes (from [Nevins and Whitney-'89]).}
\end{figure}

In a), we see that parts typically begin mating with some relative lateral and angular error, so the first contact occurs on the chamfers, as shown b). During chamfer crossing, the contact point moves down the chamfer toward the rim of the hole as the parts try to move laterally to remove lateral error. The part is pushed laterally by the force acting on it at the contact point. Once the contact point reaches the rim, it remains there, acting as the "one point" of one-point contact, shown in c). As the peg advances farther into the hole, it finally strikes the opposite side, establishing a second contact point, shown in d). During the two-point contact phase, the parts try to rotate with respect to each other to remove angular errors. The part is turned angularly by the torque created by the forces acting at the two contact points. In some cases, two-point contact may be followed by line contact, in which the parts are exactly parallel and in contact along one wall of the hole.
Conditions for successful assembly

The mechanic of the parts mating are governed by the geometry of the parts, the stiffness of the parts and tooling, the friction between parts as they move past each other during assembly, and the amount of lateral and angular error between the parts as mating begins. The interplay of these factors determines whether assembly will be successful and how large the forces exerted on the parts by the tooling and each other will be.

The success or failure of a peg-hole assembly depends on how the parts behave while passing through two potential danger zones. First, the lateral or angular errors before assembly could be so large that the parts fail to meet within the bounds of the chamfers (or part diameters if they are no chamfers). Second, there are two forms of failure associated with the two-point contact phase; these are called “wedging” and “jamming”. In the present work, we will limit ourselves to an intuitive understanding, but the reader can find firm mathematical formulation behind these ideas in [Nevins and Whitney-89].

Figure 2.7 shows schematic illustrations of both wedging and jamming. In the first case, the contact forces between peg and hole can set up compressive forces inside the peg, effectively trapping it part way in the hole. To avoid wedging, one must keep the angular error between peg and hole at the moment of first two-point contact small enough. The equations describing successful assembly show that there is a relation between avoiding wedging and ensuring that the chamfers meet, i.e. ensuring small
enough initial lateral and angular errors before one point contact. In jamming, the peg cannot advance into the hole because the insertion force vector points too far off the axis of the hole. To avoid jamming, one must support the peg so that the reaction forces set up by the two contact points are able to turn the peg parallel to the hole's axis. A remote compliance center (RCC), often added between the manipulator and the gripper of the assembly system, is such a support which has been developed using knowledge of part mating theory. These supports are also important in chamfer crossing and avoidance of wedging.

Relatively low initial lateral and angular errors and the use an appropriate RCC should increase the probability of successful mating by helping i) chamfers meet, ii) reduce chances of wedging, and iii) reduce chances of jamming. Mathematical derivations in [Nevins and Whitney-89], using parameters defined in figure 2.8, result in a parallelogram-shaped zone in the [lateral error $\varepsilon_0$ - angular error $\theta_0$] space, as shown in figure 2.9. Combinations of errors out of the parallelogram create conditions
unfavorable to success of the mate (because of probable wedging or non-meeting of chamfers). In this figure, \((W + C/2)\) is the sum of chamfer widths on the peg and hole and half the clearance between them, \(\mu\) is the coefficient of friction, and \(c\) is the clearance ratio defined by

\[
c = \frac{D - d}{D}
\]

where \(D\) and \(d\) are respectively the hole and the peg diameters. The slope of the top and bottom edges of the parallelogram is function of the location and stiffness of the compliance center of the support. The compliance center, the fundamental concept used in the RCC, is a point on which lateral forces will cause only lateral motion and about which applying a pure torque causes only rotation. This point can be found mathematically [Nevins and Whitney-89].
Figure 2.9 Geometry constraints on allowed lateral and angular errors to permit chamfer crossing and avoidance of wedging (adapted from [Nevins and Whitney-'89]).

The parallelogram in the \([\varepsilon_0, \theta_0]\) space is used as the prime criterion to evaluate the probability of success of assembly of peg-hole mates. \(\varepsilon_0\) and \(\theta_0\) are found using the proper tolerance analysis algorithm; if the latter gives the errors in three dimensions, i.e. along X, Y, Z, and about X, Y, and Z axes, the six degrees of freedom of a rigid body, one can use the directions of maximum errors. Although a full analysis could be performed in each of the three orthogonal planes, one could start with a worst case analysis performed in a single fictitious "worst case plane". For example, if the maximum linear error is found to be along the Y direction (considered in the Y-Z or X-Y planes) and the maximum angular error is found to be about the Y axis (the angle defined in the X-Z plane), we can consider the worst case by equating these two errors to \(\varepsilon_0\) and \(\theta_0\) which are considered to be in the same fictitious "worst case plane". Figure 2.10 shows the concept.
Conclusion

We have covered a wide array of topics related to the general term "tolerancing". We have also addressed how tolerance consideration can be included in a broader view of product design, including the evaluation of the assembly process.

In section 2.1, we have talked about feature-based design for assembly. In this thesis, dimensional tolerances are viewed as features to both components and assemblies, and the goal is to evaluate their impact on the assembly planning process. In section 2.2, assembly sequence analysis, a particular method to analyze assembly processes, was looked at in more details. We showed by means of an example how tolerances can affect the choice of assembly sequences. Section 2.3 took a closer look at work that has investigated how tolerances can be represented and propagated. We also described in more details a statistical closed-form method using matrix transform representation of tolerances, particularly well suited to assembly sequence evaluation. Finally, in section 2.4, part mating theory and its requirements were presented.
Chapter 3

Overview of Geometric Dimensioning and Tolerancing
Introduction

Dimensioning of parts and assemblies is accompanied by tolerancing. Once one recognizes that the desired size of a cylinder, for example, cannot be obtained automatically to a very high precision, one then indicates the allowable tolerance. Since dimensions and tolerances are so closely related, we often talk about dimensioning and tolerancing (D&T) schemes as methods to describe both constraints (dimensions) and relaxation of constraints (tolerances).

The object of this section is to give an overview of D&T methods, emphasizing the geometric approach as presented in ANSI's Y14.5 standard, through description of their evolution, basic principles and interpretation. Section 3.1 compares what is called the conventional approach to tolerance specification, with the more recent geometric approach. Section 3.2 introduces the reader to various concepts of geometric D&T, and emphasizes the concepts of "maximum material condition" in section 3.2.2 and datum in section 3.2.3, which are key principles in this D&T method. In section 3.3, we take a look at tolerance classification and describe the various specification types (form, profile, size, orientation, location, and runout in sections 3.3.2 to 3.3.7), which are associated with features of two kinds, individual and related as described in section 3.3.1. Throughout these descriptions, we try to evaluate the relation between their object and the matrix transform representation scheme, in light of assembly analysis purpose and requirements.
3.1 Conventional versus geometric dimensioning and tolerancing.

There are two schools of thought on dimensioning and tolerancing: conventional and geometric.

The first type of method, which has been widely used in the past, is a traditional annotation that is intuitive, but limited in scope and precision. It is also often referred to as traditional tolerancing or coordinate dimensioning. It specifies the nominal value of a dimension and adds the maximum and minimum value that will still be accepted, usually by means of a bilateral tolerance (ex.: X.XX +/- 0.01). We note that this type includes 3 major specifications (size, angle, and location) that cover common component variations. Figure 3.1 shows 3 examples of such annotations. The most widely used conventional tolerance is the one of size. Actually, it is typical of most drawings using geometric tolerancing to still include the size type of conventional tolerances, since it has not been necessary to replace it with a new specification in the geometric approach.

Conventional tolerancing, in addition to its limited precision, is different from geometric tolerancing in that there is less freedom in specifying the tolerance zone shape. In geometric tolerancing, this shape will be specified with reference to the nominal position. For example, figure 3.2 shows how the tolerance zone of the positioning of a hole is best represented with a cylinder whose axis is at the nominal position, rather than a square resulting from the bilateral tolerance associated the horizontal and vertical conventional size dimension.
Figure 3.1 Conventional dimensioning and tolerancing.

The other, and more recent method is referred to as geometric D&T or true positioning D&T. This methodology has evolved over many years to ensure that design requirements are interpreted unambiguously. It consists of rules, symbols, definitions and conditions that will serve the above purpose, and is fully described in ANSI's Y14.5M-1982 standards.

Although precision of this method is a significant improvement over other D&T schemes, geometric D&T has been developed predominantly in response to ambiguities when measuring and controlling various features (in quality control activities). Accordingly, developments in computing capability for computer-aided-engineering (CAE, including CAD) are still limited, in the D&T field, by a lack of proper and
Figure 3.2 Conventional versus geometric: shape of the tolerance zone.

complete mathematical representation of variabilities. Furthermore, the mathematical representation requirements will vary greatly depending on the goals of a given engineering analysis.

The main objective of this thesis is to develop a representation of these variabilities which is compatible with sequence analysis, yet not diametrically opposite to present standards. In this work, geometric tolerancing is used as the basis from which our tolerance analysis method, or mostly its tolerance representation requirements, is evaluated in terms of compatibility with current D&T practices. Errors on homogeneous matrix transforms can represent faithfully many tolerance...
zones as defined in Y14.5M standard, but is neither capable nor required to represent all possible specifications necessary to describe all possible physical variations. The goal is to represent variations which, when propagated through a sub-assembly by tolerance analysis, will most influence the position and orientation of a feature to mate. For example, matrix transforms cannot well represent form tolerances since these allow for individual points of a feature to move independently, while matrix transforms must be attached to (a minimum number of) rigid bodies to be useful. This point is discussed in more details in section 3.3.5.

3.2 Generalities

This section goes over definitions, symbols and conditions that characterize geometric D&T. These items are widely used in conjunction with the various tolerance specifications that will be classified and described in the next section. We will review the most important ones and discuss their pertinency in terms of stochastic tolerance analysis and the associated homogeneous matrix transforms. The reader is refered to [Foster-86] or [ANSI Y14.5M-1982] for more details.

3.2.1 Definitions

Here follows definitions of common details of geometric tolerancing, which will ease understanding of the intricacies of representing tolerances with homogeneous matrix transforms.
Basic Dimension:
A theoretically exact value used to describe the exact size, profile, orientation, or location of a feature or datum. It is used as the basis from which permissible variations are established by tolerances in feature control frames. A basic dimension is symbolized by boxing it.

Datum:
A theoretically exact point, axis, or plane derived from the true geometric counterpart of a specified datum feature. A datum is the origin from which the location or geometric characteristics of features of a part are established. A datum feature is an actual (physical) feature of a part used to establish a datum.

Feature control frame:
A rectangular box containing the geometric characteristic symbol and value of the form, orientation, profile, runout, or location tolerance. If necessary, datum references and modifiers applicable to the feature or the datums are also contained in the box.

Feature of size:
A cylindrical or spherical surface, or a set of two plane parallel surfaces, each of which is associated with a size dimension. A size dimension is the distance between two
geometric entities on the same feature. Examples: two opposite points on the circumference of a circle, diameter of a peg; two opposite planes that together constitute a feature, like width of a slot or cube.

**Modifier:**
A term used to describe an attribute to a tolerance type or a datum. Possible attributes include "maximum material condition" (MMC), "least material condition" (LMC), "regardless of feature size" (RFS).

**Maximum (least) material condition, MMC (LMC):**
The condition in which a feature of size contains the maximum (least) amount of material within the stated limits of size. Example: minimal hole diameter and maximum shaft diameter (maximum hole diameter and minimum shaft diameter). See figure 3.3 for an example.

**Regardless of feature size (RFS):**
Term used to indicate that a geometric tolerance or datum reference (on a feature of size) applies at any increment of size of the feature within its size tolerance. RFS is the strictest of the specification attributes as it will not allow any change in the position tolerance zone, regardless of the size of the feature.
**Figure 3.3** Maximum material condition (MMC) principle: an example (from [Foster-’86]).

### 3.2.2 Modifiers

Modifiers are an important part of geometric tolerancing: they add precision to a specification so that it better reflects the designer’s intent. These attributes are all
closely related to one important concept, which is to allow a larger tolerance zone, most commonly for the position of the axis of a circular feature, if the size of the feature to be positioned is such that the possibility of interference with a mating part is diminished. With a MMC modifier for example, if the diameter of a peg to be mated with a hole is smaller than its maximum value, the position of its axis will be allowed to vary more before interference occurs. For a hole, the diameter should be bigger than its minimum value to allow more variation on its axis position. This minimum material condition concept is particularly useful mostly in cases involving mating parts to assure interchangeability or assemblability: the position tolerance zone will be greater than one where size was disregarded and the cost of manufacturing the part will decrease accordingly.

The least material condition (LMC) modifier is a less commonly used attribute that specifies just the opposite of MMC, i.e. the position tolerance zone can increase if the size of the feature departs from its least material condition. For example, a hole which is close to the outside surface might be allowed to vary more in position if its diameter decreases in size as a concern for wall thickness between the hole surface and the outside surface.

The use of MMC and LMC modifiers make a tolerance specification incidental. One needs the knowledge of the actual size of the feature to determine if its position is within tolerance. This precludes direct consideration of the MMC principle in stochastic tolerance analysis, although it can help to establish a conditional variance for the position variation. This is discussed in more details in section 4.2.2
3.2.3 Datums

Datums are the basis for relating features to one another: they are used to help specify related features, as opposed to individual ones. They are specially applicable to the use of homogeneous transforms as tolerance zone descriptors since they correspond to the definition of a reference frame, from which a transform is defined. Actually, even for planar features of size in which two surfaces are located with respect to one another, one planar surface will have to be defined as a datum with respect to which the other surface will be specified. That is, we will add datums, even though the standard size specification does not indicate one. In most cases, this does not cause problems when performing tolerance analysis for sequence evaluation since all parts in the considered assembly will have at least one surface in contact with another part; this surface can be used as a reference from which any other feature can be specified (not necessarily with only one transform). Figure 3.4 shows an example.

Figure 3.4 Using a surface from a planar feature of size as datum.
A distinction is made between a "datum" (theoretically exact, as on the drawing), which represents design requirements and function, and the actual "datum surface" or "datum feature" on the produced part. Figure 3.5 describes this distinction. As it is shown, a datum feature is an actual (physical) feature of a part used to establish a datum.

Finally, it is important to acknowledge the issue of multiple datums. Datum planes are always established by, or are relative to, actual or physical features. The most common datum plane is the type established from a datum surface. Figure 3.5 shows the conventional establishment of a primary datum plane from a primary datum surface. In establishing datum planes in relation to defining or measuring a part, usually three (for parts other than cylindrical, two for cylindrical parts) datum or measuring planes are considered in locating features. These three planes conform to the relationship of the orthogonal X, Y, and Z axes and resulting planes of orientation. We see the pertinency of this concept in using homogeneous matrix transforms.

**Figure 3.5** Datum versus datum surface or datum feature (from [Foster-’86]).
The three planes are referred to as a datum reference frame and are composed of a primary, secondary, and tertiary (third) datum plane as established from the appropriate actual datum surfaces. Unless otherwise specified or controlled, the largest

**Figure 3.6 Establishing the datum planes (from [Foster-'86]).**
or most important surface is usually selected as the primary datum, the next largest or most important as the secondary datum, and the remaining surface as the tertiary datum. Of course, design function requirements should be the first criterion for the establishment of datum priorities, but assembly requirements need not be overlooked.

Figure 3.6 shows how a datum reference frame is established relatively to an actual part. This method is highly similar to the exact constraint scheme often used to implicitly specify part locations in assemblies [Scott-'88], in which all 6 degrees of freedom of a rigid body, no more and no less, are constrained. Figure 3.7 shows such a constraint scheme. Once the datum reference frame is established, it's as if the part can no longer move in relation to this frame and features are specified with respect to it.

![Diagram](image)

**Figure 3.7** Exact constraint assembly specification: similar to a datum reference frame.
In considering datum reference frames within our tolerance analysis approach, one will have to carefully identify the proper sequence of matrix transforms in building the tolerance chain. Figure 3.8 demonstrates how two holes, specified in position with respect to a datum reference frame, implicitly have their cylindrical position tolerance zone parallel to datum B, not surface B. This implicit specification will have to be taken into account when identifying the sequence of transforms for tolerance analysis.

Figure 3.8 Implicit orientation specification using a datum reference frame.
3.3 Tolerance classification

Means to describe allowable variations on components and assemblies have been developed as their need was being observed. The issue of conventional versus geometric tolerancing, discussed in section 3.1, is an example of such evolution. Not only are there many types of specifications, as there are just as many types of features on controlled parts, but there are also multiple specifications that can describe the same feature's allowable variation. For example, the parallelism requirements of two edges of a parallelogram can be appropriately described by two tolerance specifications: obviously, the parallelism tolerance will convene perfectly, but the size tolerance could be used just as well. Figure 3.9 shows such examples. We note that in both cases, the top edge of the square will have to lie within a zone of the same shape. In the case of the size tolerance however, the zone is explicitly constrained in its location with reference to the bottom edge.

**Figure 3.9** Parallelism: many specifications possible.
In face of all these possible specifications, it becomes important to understand them in such a way that one will choose the best one for the problem at hand. Furthermore, there is high potential for confusion when people of different experience use them to communicate often crucial points, mandatory to customer satisfaction. The reader is advised though, that even the best of understandings will still leave way to multiple specification choices for the same intent. Table 3.1 lists major tolerance kinds,

<table>
<thead>
<tr>
<th>KIND OF FEATURES</th>
<th>TYPES</th>
<th>CHARACTERISTIC</th>
<th>SYMBOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Related</td>
<td>Location</td>
<td>Position</td>
<td>⊕</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Concentricity</td>
<td>⊗</td>
</tr>
<tr>
<td></td>
<td>Orientation</td>
<td>Perpendicularity</td>
<td>⊥</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angularity</td>
<td>⊴</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parallelism</td>
<td>//</td>
</tr>
<tr>
<td></td>
<td>Runout</td>
<td>Circular runout</td>
<td>⊥</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total runout</td>
<td>⊥</td>
</tr>
<tr>
<td>Planar size</td>
<td>Distance</td>
<td></td>
<td>←→</td>
</tr>
<tr>
<td>Individual</td>
<td>Form</td>
<td>Flatness</td>
<td>⊥</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Straightness</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Circularity</td>
<td>⊵</td>
</tr>
<tr>
<td></td>
<td>Circular size</td>
<td>Diameter</td>
<td>ø</td>
</tr>
<tr>
<td>Individual or</td>
<td>Profile</td>
<td>Profile of line</td>
<td>⊖</td>
</tr>
<tr>
<td>Related</td>
<td></td>
<td>Profile of surface</td>
<td>⊙</td>
</tr>
</tbody>
</table>

Table 3.1 Tolerance classification.
types and associated symbols, and includes both conventional specifications, and geometric specifications as labelled in ANSI's Y14.5 standard. We should remind the reader here that the geometric classification really applies to characteristics of geometric features which have to be both dimensioned and tolerated; in this work, we refer to a classification of tolerances since our focus is on specifying the variations.

The next few pages describe these typical specifications as explained in an application book on dimensioning and tolerancing which uses ANSI's standard extensively [Foster-'86]. We will describe enough to allow minimum understanding of basic tolerances but mostly, enough to show their compatibility and usefulness, or lack thereof, with stochastic tolerance analysis and homogeneous matrix transforms as tools to assembly sequence evaluation.

3.3.1 Related and individual features

Geometric features can be classified in three kinds: individual, related, and individual or related. This distinction is important in this work because tolerance analysis, and homogeneous matrix transforms, makes mostly use of tolerance on related features. Variations on individual features affect only indirectly, if at all, the position and orientation of the feature to be mated in subassemblies.

An individual feature is a single surface, element, or size feature which relates to a perfect geometric counterpart of itself as the desired form; no datum is used when dimensioning and tolerancing an individual feature. All form tolerances are associated
with individual features. A related feature is a single surface or element feature which relates to a datum, in orientation, runout, or location. An individual or related feature is a single surface or element feature whose perfect geometric profile is described, but may or may not relate to a datum. This is a less commonly used kind of feature that includes only profile type tolerances.

Intuitively, we realize that related features and associated tolerances can be modeled with homogeneous matrix transforms since this representation uses reference frames and relates all transformations to these frames or to each other. Individual features cannot be modeled so straightforwardly with matrix transforms, and features of form will generally prove of little importance in stochastic tolerance analysis for assembly sequence evaluation.

For example, the circularity (form specification) itself of a peg is of little influence in its mating with a hole; on the other hand, the knowledge of the value of the largest diameter (size specification) found on that peg can help in evaluating the peg's position distribution in the hole once the mate is done. Also, the knowledge of this diameter is important in evaluating the probability of mating success. Therefore, individual features of size indirectly affect tolerance analysis using matrix transforms by affecting the stochastic parameters of some variables. This relationship is described in more details in sections 4.2.2 and 5.1.
3.3.2 Location

Location specifications and their associated tolerance zones, along with orientation specifications, control the most influential variations when performing tolerance analysis for assembly sequence analysis. They include position and concentricity. A position tolerance defines a zone within which the axis or center plane of a feature is permitted to vary from true, or theoretically exact, position. A concentricity tolerance defines a cylindrical zone within which the axis of a circular feature must lie; the axis of the tolerance zone is the extension of the axis of the datum feature. We note that concentricity is similar to position except that it applies only to features with an axis (cylinders, cones, spheres, etc.).

The maximum material condition (MMC) principle is most widely used in conjunction with the position tolerance, specially when function or interchangeability of mating part features is involved. The MMC concept however, as seen in section 3.2.2, is based on an instantiation of a physical part, and cannot therefore be considered easily in stochastic tolerance analysis since only independent variations can be represented. Furthermore, we have seen that the MMC principle is based on features of size (most commonly circular), which should not be considered as separate variables in regards to the goal of our tolerance analysis method, but only as variables to help define conditional position.

Position (or more generally, location) tolerances usually involve features of size and relationships to center planes and axes; at least two features, one of which is a size feature, are required before a position tolerancing is valid. The geometric tolerance example of figure 3.2 shows an example of a cylindrical feature positioning, while
**EXAMPLE**

XXX ± XXX

- Φ .005 ± A ± .005

**MEANING**

DATUM A CENTER OR MEDIAN PLANE

WITHIN .005 REGARDLESS OF DATUM SIZE

CENTER OR MEDIAN PLANE OF SLOT, REGARDLESS OF SLOT SIZE, MUST LIE BETWEEN TWO PLANES .005 APART EQUIDISTANT FROM THE MEDIAN PLANE OF THE DATUM, REGARDLESS OF DATUM SIZE.

.005 TOTAL TOL ZONE

.0025

**Figure 3.10 Position tolerance associated with a non-cylindrical feature (from [Foster-'86]).**

Figure 3.10 shows one of non-cylindrical feature positioning which is not using the MMC concept, i.e. specifies "regardless of feature size" (RFS, §). Figure 3.11 shows an example where position tolerancing is used to specify coaxial features. Note here the use of the MMC principle on both the datum and the related feature, which allows for a greater axis position tolerance zone if: 1) the size of the feature is not at it's maximum limit; 2) the datum feature size is not at its maximum limit; or 3) both are not at their maximum size limit. If the MMC principle was not used, i.e. RFS was specified in
Figure 3.11 Position tolerance used to specify coaxial features (from [Foster-'86]).

the position feature control frame, then this resulting position specification would be equivalent to a concentricity specification: i.e., $\varnothing \pm 0.001 \pm A$ is equivalent to $\varnothing \pm 0.001 \pm A$. 

96 

Chapter 3
3.3.3 Orientation

An orientation tolerance controls orientation of features to one another. Expressions of these tolerances refer to perpendicularity, angularity, and parallelism (and in some instances, profile) as related to datum features. Orientation tolerance when applied to a plane surface must be less than the governing size tolerance of the surfaces involved. When applied to a cylindrical size feature, it is a refinement of the locational tolerance which already includes a specification of orientation. Figure 3.12 describes how an implicit orientation (parallelism) tolerance is included in a location

![Diagram](image)

**Figure 3.12 Parallelism: used as a refinement of location tolerance.**
tolerance, and shows the same example with an explicit orientation tolerance as a refinement of requirements. Note how the additional parallelism tolerance really only applies to the view shown, which results in an almost rectangular tolerance zone parallel to the datum plane; this would not be true if the datum was an axis or if a tolerance of perpendicularity to a surface were used, in which cases a cylindrical tolerance zone would result. Figure 3.13 shows that last point.

**Figure 3.13** Orientation: tolerance zone shape depends on datum used.
Figure 3.14 shows an angularity tolerance while figure 3.15 shows a parallelism tolerance applied to a surface. Note again how the tolerance zone specified by the parallelism specification is constrained in position only by the size specification: the first is a mobile subset included in the second.

EXAMPLE

MEANING

Figure 3.14 Angularity tolerance (from Foster-'86).
EXAMPLE

Symbol Meaning

// .002 A
TO DATUM PLANE A
WITHIN .002 WIDE TOL ZONE
THIS SURFACE MUST BE PARALLEL

MEANING

.002 WIDE TOL ZONE PARALLEL. TO DATUM A

THE SURFACE MUST BE WITHIN THE SPECIFIED TOLERANCE OF SIZE AND MUST LIE BETWEEN TWO PLANES .002 APART WHICH ARE PARALLEL TO THE DATUM PLANE.

Figure 3.15 Parallelism of a surface (from [Foster-’86]).

3.3.4 Runout

Runout tolerances states how far an actual surface or feature is permitted to deviate from the desired form and orientation implied by the drawing during full rotation (360°) of the part on a datum axis. It is a method used to control the composite surface effect of one or more features of a part relative to a datum axis. They are two types of runout specifications: circular and total. The first one controls each circular elements individually and the second one controls the complete surface, i.e. all circular elements together. Circular runout controls any composite error effect of circularity,
**EXAMPLE**

\[ \varphi_{XXX} \pm \varphi_{XXX} \]

**MEANING**

**Figure 3.16 Total runout (from [Foster-'86]).**

concentricity (location), and circular cross-sectional profile variation, while total runout controls composite error effect of circularity, concentricity, angularity and parallelism. Figure 3.16 an example of total runout. Although circular runout theoretically does not control orientation, its control of concentricity results in a
cylindrical tolerance zone in which the axis can lie, in any orientation possible. We will therefore consider circular runout to control orientation within the cylindrical location zone; both location and orientation specification are extracted from runout tolerances through the cylindrical tolerance zone. On the other hand, cross-sectional profile or circularity will be disregarded as they have little direct effect on results sought by our analysis.

**EXAMPLE**

![Example Diagram]

**MEANING**

FIXED POSITION OF INDICATOR PARALLEL TO THE AXIS FOR 360° ROTATION

**Figure 3.17** Runout: can control form (circular & total) and orientation (total only) from [Foster-'86].
Size variation has no effect upon the runout tolerance compliance: it applies on a "regardless of feature size" (RFS) basis. Finally, elements or surfaces controlled may be those constructed around a datum axis, or those constructed at right angles to a datum axis. Figure 3.17 shows an example of a runout tolerance applied to a surface perpendicular to the datum axis. This is a (lateral) circular runout and controls the form (wobble) of each circular element, but one can see that a total runout specification on the same surface would also control perpendicularity of the surface.

The collective or composite control of various form, orientation, and location (concentricity) variations of the part provides a more direct representation of part functions, integrates manufacturing operations, and minimizes inspection setup requirements. This makes it a widely used specification that must be considered while performing tolerance analysis. We must keep in mind though, that for our goal of acquiring knowledge about position and orientation of features to mate in the assembly process, location and orientation variations are the ones with most impact.

3.3.5 Form and profile

A form tolerance states how far an actual surface or feature is permitted to vary from the desired form implied by the drawing. Expressions of these tolerances refer to flatness, straightness, circularity and cylindricity, and occasionally profile of surface and profile of line. Figure 3.18 shows an example of flatness tolerance. Note that the flatness tolerance zone is not fixed in position, which is set by the tolerance of size. Flatness, and all the other form tolerances, are associated individual features and are
therefore incidental: they bear little weight in our method of, and goal to, tolerance analysis.

**EXAMPLE**

**MEANING**

**Figure 3.18** Flatness: a form tolerance (from [Foster-'86]).
Furthermore, only explicit form specifications, used as a refinement to other types, are not considered by our approach: If we consider that form is implicitly controlled in both orientation or position specifications, as shows figure 3.17, one can infer that these two specifications account for form variations that are of the same order of magnitude as size or orientation variations.

Profile tolerance is associated with a surface or a line only. The profile of a surface (line) is is the condition permitting uniform amount of profile variation, either unilaterally or bilaterally, on a surface (along a line element of a feature). The surface profile control is the most commonly used; its tolerance often includes both form and orientation, and sometime size or position. For our application, we consider only variations in position and orientation, similarly to a flat surface when tolerated in size with another surface; the only difference is that the profiled surface must lie within a profiled zone while the flat surface must lie within a rectangular zone.

**Figure 3.19** Rough form variations: confined within size or orientation tolerance zones.
Conclusion

In this section, we have tried to describe the most important features of geometric dimensioning and tolerancing (D&T) in light of our goals and methods to implement assembly sequence analysis based on tolerances. Although we essentially described the basic concepts, specifications and attributes used in geometric D&T, we have presented the subject with a perspective that will ease understanding of its interpretation in terms of homogeneous matrix transforms. A proper representation of standard specifications with matrix transforms will allow implementation of a closed-form tolerance analysis algorithm, useful in assembly sequence evaluation.

After discussing the major differences between conventional and geometric D&T, we have seen important concepts of geometric D&T such as the maximum material condition (MMC) principle and the use of datums. We have concluded that the very nature of stochastic analysis prevents the direct consideration of MMC attributes since they are incidental, i.e. related to an actual physical instance of a part. Only the conditional influence of the MMC attribute on the possible position of the feature can be considered. On the other hand, we have seen that datums, and reference datum frames, are inherent to matrix transform representation since transforms are always defined between a reference frame and the considered feature.

On the classification of tolerance types, we have divided them in two main kinds: 1) individual, which includes all form tolerances such as flatness, circularity, etc.; and size tolerance, similar to a conventional specification; and 2) related, which includes location, runout and orientation tolerances. We have concluded that while all
related features can have their associated tolerance zone be taken into account to some degree in matrix transform representation, only size specifications of individual features are important to, and compatible with, our tolerance analysis approach. Size tolerance of circular features is used only to find a conditional variance on the resulting position, while size tolerance on flat features (ex.: two opposite surfaces of a cube) is equivalent to a position specification on one of the feature’s surface using other surface as datum.

In the next chapter, we describe in details the method used to describe these tolerance zones with statistical errors associated with each of the six degrees of freedoms of rigid bodies. These errors are associated with homogeneous matrix transforms that give the nominal position of features with respect to one another.
Chapter 4

Matrix Transform Representation of Tolerance Specifications
Introduction

In this chapter, our goal is to represent allowed variation, or design intents, in a format appropriate to homogeneous matrix transforms, so that the related tolerance analysis technique can be implemented. Furthermore, in order to comply with ANSI's Y14.5 guidelines, a widely used standard, we concentrate on transforming the specifications associated with this standard, and we focus our energy on those specifications most influential in the assembly process.

In the previous chapter, we described the most important geometric specifications and introduced a reflection on the implications of such specifications on stochastic tolerance analysis and matrix transform representation. We have seen that variations of location and orientation were most influential to the assembly process, and that extraction of these allowed errors from all standard specification types was necessary to appropriately evaluate assembly sequences based on tolerances. Accordingly, this chapter will describe methods to represent position and orientation variation zones, whether they are explicitly controlled, as in location or orientation specification types, or implicitly controlled, as in size and runout specification types.

Section 4.1 presents how location specifications can be represented in matrix transforms; position and concentricity are considered. Also included in this section is the location control implicit to the runout specification type. Section 4.2 shows how size specifications can be considered in matrix transform representations; size specifications on a planar feature will have their full description, while size of circular features will only modify location representation through the notion of conditional
variation. The important “maximum material condition” principle, which is most used in conjunction with location tolerances, is a case that well represents the issue of conditional variation. Finally, section 4.3 shows how orientation specifications are represented independently only if they constitute a refinement not already included in other specifications. Also included in this section is the perpendicularity control sometime implicit to total runout specifications. Concerning form specifications, we know that they have a smaller impact on the variation in position and orientation of features to mate in the assembly process. Also, as was explained in the last chapter, they are incidental, which precludes their consideration in a stochastic approach, and they cannot be easily represented with matrix transforms. These reasons justify the absence of form tolerances in this chapter.

4.1 Location

Standard location specifications include position and concentricity. Most of these specifications define a cylindrical tolerance zone in which the axis of a circular feature must lie. In this section, we first provide the general concept used to represent this cylindrical tolerance zone with matrix transforms. Then, we show how standard position, concentricity, and even runout notations are considered as specific applications of this concept. Although modifiers are often used in conjunction with position specifications, they are strongly related to size control, and they are discussed in section 4.2.2. Similarly, we note that position of planar features in space is usually found through an interpretation of the planar feature size specification, which can be found in section 4.2.1.
4.1.1 Axis in a cylindrical tolerance zone

Figure 4.1 shows an example of a tolerance on a circular feature referenced to a datum reference frames. This figure also shows the three-dimensional interpretation of this specification: the cylindrical tolerance zone. In essence, the axis position of the cylindrical tolerance zone is given by the basic dimension, while its diameter is given by the tolerance value in the feature base frame. This variation zone defines where the axis of the actual feature, in this case a hole, can lie; it defines the boundaries outside which no point of the actual feature axis should be, by result of a position or orientation variation, in order to conclude that the part corresponds to the design intent.

To represent this zone with a matrix transform, we attach a Cartesian frame to the middle of the axis of the tolerance zone. This axis also represents the nominal position and orientation of the actual feature axis with respect to the previous (reference) frame. To simulate the possibility for this real axis to be anywhere in the tolerance zone, we associate variabilities to specific degrees of freedom defining its position and orientation; it's as if we consider the axis of the actual feature to be a rigid body with all six degrees of freedom constrained to 0 with respect to the reference frame, but we relax the constraints on specific degrees of freedom. Figure 4.2 demonstrates the concept. Four degrees of freedom have their constraint relaxed: two linear degrees of freedom, X and Y, are transformed in polar coordinates r and α, while the rotation angle around the X axis, θx, and the rotation angle around the Y axis, θy, complete the varying subset. By relaxing the constraints on X, Y, θx, and θy, we allow the axis of the actual feature to vary in position and orientation up to the boundaries of the tolerance zone.
Figure 4.1 Tolerance on a circular feature: cylindrical tolerance zone for its axis.

The key issues in representing the tolerance zone with small variations on these degrees of freedoms include the choice of stochastic distribution and the assignments of limits to each of these variations. The algorithm we propose to use to perform tolerance analysis is based on a certain number of assumptions, two of which are most important to consider in this work: 1) the variates must be independent; 2) the variates must be normally distributed within their range. We will now discuss how these two issues are considered in our work.
Figure 4.2 Representation of tolerance zone using variation ranges on a subset of the 6 degrees of freedom of Cartesian frames.
Dependency of the variations: defining the boundary of the tolerance zone in the parameter space.

For our cylindrical tolerance zone, the variates $X, Y, \theta_x,$ and $\theta_y$ have ranges that cannot be independent if the tolerance zone is to be respected. Assume for a moment that a particular instance of $r$ is 0 (both $X$ and $Y$ are 0). $\theta_x$ and $\theta_y$ can now vary, dependently, within the maximum range $\{0, \theta_{\text{max}}\}$, where $\theta_{\text{max}}$ is an approximated function of the radius and the height of the cylindrical tolerance zone:

$$\theta_{\text{max}} = \frac{2}{L} r_{\text{max}}$$

(4.1)

where $L$ is the height of the tolerance zone and

$$r_{\text{max}} = \frac{\text{tolerance zone diameter}}{2}$$

(4.2)

One will have recognized equation 4.1 to be a small angle approximation to the exact function

$$\sin \theta_{\text{max}} \left( \frac{L}{2 \cos \theta_{\text{max}}} \right) = r_{\text{max}}$$

(4.3)

where $\sin \theta_{\text{max}}$ is replaced by $\theta_{\text{max}}$ and $\cos \theta_{\text{max}}$ is replaced by 1. This approximation is justified by the fact that angles involved in tolerances are very small. To give the reader an order of magnitude, an angle of 30°, which is much greater than what would allow typical position specifications, leads to an approximation error of about 5%. Equation 4.1 then, represents the maximum variation value that either $\theta_x$ or $\theta_y$ can take, while the other angle is kept to 0. The dependence relation between $\theta_x$ and $\theta_y$ is the same as between $X$ and $Y$ which results in a circle of radius $r_{\text{max}}$, equivalent to the
cross section of our original tolerance zone. This is represented by figure 4.3, in which we clearly see that \( r \) is null and \( \alpha \) is nonexistent. For the case where \( r \) is not equal to 0, \( \alpha \) has to be considered.

\[
\theta_{\text{max}} \approx \frac{2}{L} r_{\text{max}}
\]

Figure 4.3 Tolerance zone in both domains: if they are no linear errors, the maximum angular error are directly proportional to the tolerance specification \( 2r_{\text{max}} \).

At a given value of \( r \) (defined by \( X \) and \( Y \)) between 0 and the radius of the cylindrical tolerance zone, \( \theta_x \) and \( \theta_y \) have a range which is dependent on this value of \( r \).

Figure 4.4 shows a top view of the cylindrical tolerance zone where \( r \) has a given value and \( \alpha \) is at 90°, the angle of \( r \) around the \( Z \) axis, measured from the \( X \) axis. Considering that: 1) we create a new reference frame in the middle of the axis now vertically
positioned at the end of \( r \); and 2) no point of the real axis can go beyond the original zone of circular cross-sectional shape; we see that any angular variations around the \( X \) axis (\( \theta_x \)) and the \( Y \) axis (\( \theta_y \)) are limited in such a way that none of the two axis extremities should go beyond a zone whose cross-sectional shape is the result of the intersection of two circles of radius \( r_{\text{max}} \), distant of \( 2r \). The middle of this new zone is at the end of the vector \( r \). The shape of this zone in the \( \{X, Y\} \) domain is replicated in the \( \{\theta_x, \theta_y\} \) domain. As the value of \( r \) increases, the ranges of \( \theta_x \) and \( \theta_y \) decrease: the variates \( r, \theta_x \), and \( \theta_y \) are dependent, and this fact will have to be taken into account in our representation approach. Actually, for \( \alpha \) equal to \( 90^\circ \), we have that

\[
\theta_x \leq \frac{2}{L} (r_{\text{max}} - r) \quad \text{and} \quad \theta_y \leq \frac{C}{L}
\]  \hspace{1cm} (4.4)

where

\[
C = 2 \sqrt{r_{\text{max}}^2 - r^2}
\]  \hspace{1cm} (4.5)

is obtained through analytic geometry.
Figure 4.4 If linear errors exist ($r \neq 0$), then angular errors are limited to a smaller region.

The two variates resulting from the polar representation of the two orthogonal X and Y degrees of freedom are $r$ and $\alpha$. We have seen the interaction of $r$, let us now evaluate the effect of $\alpha$. If $\alpha$ varies within $(0, 360^\circ)$, the middle of the "circle intersection-shaped" zone in the $(X, Y)$ domain will vary in position around a circle of radius $r$, since it is coincident with $r$'s extremity. Figure 4.5 shows this variation. The
equivalent variation in the $\{\theta_x, \theta_y\}$ domain results in a circular area of radius $C/L$ as $\alpha$ varies continuously between 0 and $360^\circ$. As $r$ increases, $C$ decreases and the circular

**Small Value of $r$:**

![Diagram showing small value of $r$.]

**Large Value of $r$:**

![Diagram showing large value of $r$.]

*Figure 4.5 If linear errors exist ($r \neq 0$), the resulting angle of $r$ is uniformly distributed in $[0, 2\pi]$.***
area radius diminishes. We note that if the variates $X$ and $Y$ are considered independent and characterized by the same distribution type and stochastic parameters (means, variances, etc.), then $\alpha$ can take any value between 0 and $360^\circ$ with equal probability.

We have just defined how a rigid body, in our case the axis of a cylindrical feature, can be assigned dependent variations on $r$, $\theta_x$, and $\theta_y$, a modified subset of its six degrees of freedom $X$, $Y$, $Z$, $\theta_x$, $\theta_y$, and $\theta_z$, to represent a given tolerance zone. The range of these three variates in the space formed by them will represent the boundary of maximum variation. More specifically, in a parameter space defined by $r$, $\theta_x$, and $\theta_y$, we find that the dependence relations result in an ellipsoidal boundary limit outside of which the position tolerance specification would not be respected. Figure 4.6 shows this boundary and the way it is found. For $r = 0$, the radius of the circular area formed by the relation between $\theta_x$ and $\theta_y$ is at its maximum, with a radius of $C/L$ where $C$ is defined by equation 4.5. As the absolute value of $r$ increases, the corresponding radii of the $\theta_x - \theta_y$ circles in planes perpendicular to the $r = 0$ plane diminish, as we see in figure 4.5. When $r = r_{\text{max}}$, then there is no more possible variation of $\theta_x$ or $\theta_y$, the radius of the circle is null from the fact that $C = 0$. Finally, we find that the rate at which the radius $C/L$ decreases as $r$ increases is not linear but follows an ellipsoidal relation. In other words, in any plane of the $(r, \theta_x, \theta_y)$ domain containing the $r$ axis and rotated about it, the relation between $\theta$, the angle resulting from the combination of $\theta_x$ and $\theta_y$, and $r$ is defined on the boundary curve by

$$\frac{r^2}{r_{\text{max}}^2} + \frac{\theta^2}{\theta_{\text{max}}^2} = 1 \quad (4.6)$$
where $\theta_{\text{max}}$ is equal to $C/L$ for a given value of $r$, and $r_{\text{max}}$ is the radius of the cylindrical tolerance zone. Note that the replacement of $X$ and $Y$ by $r$ is justified in a special case of stochastic distribution of the two variates, which will be shown below.

![Diagram of zone boundary in variation space.](image)

*Figure 4.6 Zone boundary in variation space.*

**Distribution of the variations**

The boundary of the limit variations of the three variates $r$, $\theta_x$, and $\theta_y$ is now defined. The next task is to define how, within this boundary, are the variates distributed. We will first take a look at the distributions in the proposed analytical
approach, where the variates are independent and normally distributed, then we will try to define a model of what really happens within the cylindrical zone. The object of this reflection is to see if this reality can be reasonably approximated by the analytical solution.

1) Proposed tolerance analysis algorithm

As described in section 2.3, we recall that the proposed tolerance analysis based on an analytical solution also results in an ellipsoid when a three-dimensional subset space of the six degrees of freedom is considered. Although in this case, we know that this three-dimensional geometric shape results from the fact that the variates are normally distributed within their respective ranges, in addition to being independent of each other. The probability density function of the variates within this ellipsoid is defined by the 3 variate's joint normal distribution.

\[ r = 0 \quad \text{PLANE} \]

\[ \theta_x, \theta_y, \theta_z \]

*Figure 4.7 Density visualization planes.*
To allow visualization of this probability function, we look at it over two-dimensional planes which are parallel to one of the three orthogonal planes \( r = 0, \ \theta_x = 0, \) or \( \theta_y = 0 \). Figure 4.7 shows the possible orientation of the visualization planes.

Since any two variates are joined by the same probability function, we need only to look at one orientation of these planes: we choose planes parallel to the plane \( r = 0 \). In these planes, the joint density probability function of \( \theta_x \) and \( \theta_y \) is

\[
f(\theta_x, \theta_y) = \frac{1}{2 \pi \sigma_{\theta_x} \sigma_{\theta_y}} \exp \left( -\frac{1}{2} \left( \frac{\theta_x^2}{\sigma_{\theta_x}^2} + \frac{\theta_y^2}{\sigma_{\theta_y}^2} \right) \right)
\]

and is shown in figure 4.8 for three values of \( r \), \( r = 0 \), \( r = 0.33 \ r_{\text{max}} \) and \( r = 0.8 \ r_{\text{max}} \).

We note that equation 4.7 is valid only in cases where the means of the variates are equal to 0.

These density functions are the result of independent, normally distributed joint variates; what can we say about the dependent distributions of the variations on the four degrees of freedom for our cylindrical tolerance zone?

ii) Real distributions in the cylindrical tolerance zone

In this section, we will define a model of the real distributions of the four variates \( X, Y, \theta_x, \) and \( \theta_y \), and we will show how this model can be described using only 3 variates by representing \( X \) and \( Y \) by \( r \) uniformly distributed in all directions.
Figure 4.8 Analytic surfaces resulting from joint probability density of normally distributed angular variations.

In the toleranced case of figure 4.1, we assume that the hole normally varies in position symmetrically and circularly around the nominal value specified by the basic dimensions. That is, the probability density function above the X-Y plane (perpendicular to the hole axis) is equivalent to the joint density function of X and Y

\[ f(X,Y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left(-\frac{1}{2} \left( \frac{X^2}{\sigma_x^2} + \frac{Y^2}{\sigma_y^2} \right) \right) \]  

(4.8)
where \( \sigma_x = \sigma_y \). Equation 4.8 is equivalent to equation 4.7 where \( \theta_x \) and \( \theta_y \) and their variances are replaced respectively by \( X \) and \( Y \) and their variance. This assumption implies that in any given direction in the \( X-Y \) plane, the probability of the pair \( (X,Y) \) is normally distributed. Figure 4.9 demonstrates the approach. We will denote by \( r \) the resulting distance from the nominal position, a value which is normally distributed, with 0 means, along any direction in the \( X-Y \) plane.

![Diagram](image)

**Figure 4.9** Radial distance in joint normal density: distributed like the two variables of equal variance.

The assumption of a normal distribution has been the object of much work in the past and is most often justified by the central limit theorem. This theorem states the result of the convolution of many (usually more than two) distributions with similar
ranges approaches the normal distribution shape. Ultimately, the exact distribution of each feature dimension depends on the manufacturing process and the corresponding quality control strategy used, and we will not consider this level of detail in the present work.

We have demonstrated how \( r \), a normally distributed variate in all directions, can represent the joint density of the relaxed \( X \) and \( Y \) degrees of freedoms that characterize the stochastic position of the actual hole, or rather, of its axis. We will now evaluate how \( \theta_x \) and \( \theta_y \), the combined orthogonal angles the hole axis can take, are distributed within their circular tolerance zone as described earlier by figure 4.5.

We have seen in figure 4.4 that, for a given value of \( r \), any angular variations around the \( X \) axis (\( \theta_x \)) and the \( Y \) axis (\( \theta_y \)) are limited in such a way that none of the two axis extremities should go beyond a zone which cross-sectional shape is the result of the intersection of two circles of radius \( r_{\text{max}} \), distant of \( 2r \). We now make the assumption that \( \theta_x \) and \( \theta_y \) are dependently distributed within that region, each of them being almost normally distributed in the range allowed by the actual value of the other. In other words, at \( \alpha = 90^\circ \) for example, if \( \theta_x \) is such that the top extremity of the hole axis is at \( r + (r_{\text{max}} - r)/2 \), then \( \theta_y \) will be almost normally distributed in its range defined by the linear space left for the top extremity of the hole axis to move as a result of the rotation around the \( Y \) axis. This linear space is in turn defined by the boundary of the zone at a specific value of \( r \). This example is shown in figure 4.10. We say almost normally distributed because the probability that the axis be within the boundary specified by the tolerance has to be 1, or else the part would be out of specification and
would not respect the design intent. A true normal distribution does not have such hard limits.

![Diagram showing the relationship between \( r \), \( \theta_x \), and \( \theta_y \).]

*Figure 4.10 Simulation of real densities: \( \alpha \) and 3 dependent variables: \( r \), \( \theta_x \), and \( \theta_y \).*

To visualize what this last assumption about the distributions of \( \theta_x \) and \( \theta_y \) results in, we have written a program that simulates a large number of instances of these two variates for a given value of \( r \). Figure 4.11 shows the results of this Monte-Carlo simulation for two cases of \( r \). These surfaces are read as histograms of the instances that were found to be in a certain (square) range of the two variates: the number of instances in this range is represented by a height at the (discrete) middle point of the range, and all points are joined using a linear interpolation to create the final mesh or surface. If not smoothed, the surfaces are irregular. All subsequent surfaces were smoothed by averaging all points within the boundary with their eight adjacent points.
Additional Monte-Carlo simulations have been performed to include the effect of $\alpha$, i.e. to obtain a complete model of what we assume to be close to reality for specific values of $r$. The resulting probability density functions, shown in figure 4.12, constitute a plausible model for the distributions of $\theta_x$ and $\theta_y$ in the circular boundaries shown in figure 4.5. It is this model of reality that we have to compare to the analytical solution resulting from the error propagation algorithm we propose to implement to perform tolerance analysis.

Figure 4.12 Simulation at specific instances of $r$, and with $\alpha$ uniformly distributed in $[0,2\pi]$. 
Comparison between analytical solution and model of reality obtained by simulation: optimization

Figure 4.13 shows a qualitative comparison between the surfaces obtained from the analytical solution and the model of reality obtained by Monte-Carlo simulation. We recall that the analytical solution is based on the independency of the variates while the model of reality involves obvious dependencies. Furthermore, it is important to remember that the volume under the analytical surfaces is infinite since it is described by an exponential function (normal probability) while the real model has hard limits at the boundary of the ranges to respect the tolerance specification. To obtain this clear limit when simulating the real model, we simulate each variates with a normal distribution but we reject any instances outside the 6-σ range. This would effectively result in a 98.9 % probability of finding all instances under the surface but the latter is normalized to obtain 100 %. Never the less, the two set of surfaces, the one created by the analytical solution and the one resulting from the Monte-Carlo simulation of reality, are very similar.

A preliminary evaluation of our analytical solution states that it is highly compatible with a plausible model of reality in the case of a cylindrical tolerance zone. The analytical algorithm not only results in an equivalent ellipsoidal 6-σ, 97.1 % boundary shape in the {r (or X and Y), θx, θy} domain, it also results in variate distributions of similar shapes.
Figure 4.13 Comparison between simulated and analytic densities at specific instances of r.

Quantitative evaluation

When considering the probability density functions for the parameter pair $\theta_x - \theta_y$, which are the same for the two other possible pairs, we can evaluate quantitatively
the error between the two models. To do so, we can look at the Chi-square ($\chi^2$) error
between the two discretized surfaces

$$\chi^2 = \sum_{l=0}^{n} \sum_{j=0}^{m} ( p_{AU}(\theta_x, \theta_y) - p_{RU}(\theta_x, \theta_y) )^2$$  \hspace{1cm} (4.9)

where $m$ and $n$ are the number of equally distant discretized points in the ranges of $\theta_x$
and $\theta_y$ respectively, and $p_{AU}(\theta_x, \theta_y)$ and $p_{RU}(\theta_x, \theta_y)$ are the probability values at
these points of respectively the analytical and the reality-based model surfaces. One
realizes that the actual $\chi^2$ value will change with $m$ and $n$. This criterion, widely used in
numerical error minimization algorithms such as the method of least square
approximation, gives an absolute number which is then used to compare many possible
solutions to the same problem. It is more a relative appreciation of those solutions
rather than an exact computation of the error; the difference between two curved
surfaces is an evaluation of how different they are, and this can be described by many
factors (continuity, 1st, 2nd and 3rd curvatures in any possible directions, etc.). In
this work, we will concern ourselves more with finding a means to compare solutions
than to completely and exactly define their results. We therefore use the $\chi^2$ error
computed in all cases with the same values of $m$ and $n$, as well as the same ranges of
variates.

In the case of the cylindrical tolerance zone defined by an ellipsoid in the \{r, \theta_x, 
\theta_y\} parameter space, we notice that the error between the analytical and reality-based
probability surfaces increases as $r$ decreases. For $r = 0$, the $\chi^2$ error is very small at
4.7 while for $r = 0.8 r_{\text{max}}$, the $\chi^2$ error is 3197, about 650 times larger. In the
process of looking for the best analytical solution to represent our model of reality, one is forced to ask if there is a way to reduce this error, which is the object of the next section.

Optimization of analytical parameters

In this section, we define a method to optimize the analytical solution whose results are presented in figure 4.13. On the one side, it is our desire to limit ourselves to the closed-loop and rapid algorithm we propose to use to propagate the tolerances, which implies use of the six degrees of freedom associated to a rigid body, and their independency from each other; on the other side, this algorithm still gives us a relative flexibility in defining the ranges in which these degrees can vary. By changing them in a specific way, there is a possibility that the analytical solution be even closer to our model of reality.

Typically, these ranges are based on information obtained from the geometry of the tolerance zone: its radius will define the ranges for \( r \), and for \( \theta_x \) and \( \theta_y \) through a small angle approximation (note therefore that the maximum values of the three variates are linearly proportional), and the height \( L \) of the zone. In an effort to reduce the error of the solution, one might implement an optimization factor that would multiply one or more of these parameters. However, in the search for such factors, we are limited to the guidance provided by the error found in a two-dimensional space: the analytical and simulated surfaces are generated for specific values of \( r \), and we thereby loose a degree of freedom. The error between the surfaces should be evaluated at more
than one of these \( r \) values, so that an overall and more general error can be averaged as the sum of weighted errors at each considered values of \( r \). The weighting factors are similar to the probabilities of \( r \) for each case.

One should keep in mind however, that these computations will be limited in precision since we will search the optimizing factor using a limited number of discrete values of \( r \). In this work, we will limit ourselves to three values of \( r \) wisely chosen on its possible range. For each case, we find the factor, applied to both \( \theta_x \) and \( \theta_y \), that will minimize the \( \chi^2 \) error between the analytical and the reality-based model. We wrote a program to perform such optimization on a numerical basis. Then, we find an overall factor that minimizes the overall error on the whole range of \( r \). To do so, similarly to the overall error computation, we sum the weighted optimum factors found for each considered values of \( r \).

We now turn to our method for choosing the three values of \( r \). The weighting factors are proportional to the probabilities of \( r \) being equal to each considered value. We add that the computation of these probabilities involves an approximation. \( r \) is a continuous variable with a normal probability density curve, hence the probability of it being equal to a given value is, in reality, the probability of it being in a small range centered on that value. Since we have decided to use only three values of \( r \), we will consider the weighting factors as the probability of finding \( r \) in three non-overlapping ranges covering the 3-\( \sigma \) length. We now need to find a wise choice for the separation of these three ranges. Once found, we will simulate our reality-based model at the three middles of the considered ranges.
Since the normal curve is non-linear, we first find a linear approximation of it in three segments, and we limit the extremities of these segments to be on the original curve for simplicity. This is shown in figure 4.14. Then, we assume the ranges to be limited by these four extremities and we find the middle of each range. These three values of $r$ are the ones used for the three Monte-Carlo simulations from which the optimized factor search is performed. In our case, we find $r$ equal to 0.04 $r_{\text{max}}$, 0.36 $r_{\text{max}}$, and 0.8 $r_{\text{max}}$ respectively for ranges 1, 2 and 3 shown on figure 4.14. The specific optimization factors for these values of $r$ are respectively 1.00, 0.84 and 0.67.

**Figure 4.14** Search for appropriate instances of $r$ for optimization.

Finally, we compute the probability of each $r$ being found in their respective range using a table of unit standard normal distribution, and we multiply this probability to the optimum factor found in each case. The result is a weighted, overall
optimization factor that can multiply the ranges of $\theta_x$ and $\theta_y$ in the analytical solution. In our case, the probabilities are 0.27, 0.68 and 0.05 for each of the ranges, and the final overall factor $f_{opt}$ is computed with

$$f_{opt} = \text{prob}(r \in R_1) \times f_{opt_1} + \text{prob}(r \in R_2) \times f_{opt_2} + \text{prob}(r \in R_3) \times f_{opt_3}$$

$$f_{opt} = 0.27 \times 1.00 + 0.68 \times 0.84 + 0.05 \times 0.67 = 0.92$$

where $\text{prob}(r \in R_1)$ is the probability of $r$ being element of range 1 and $f_{opt_1}$ is the specific optimizing factor found for the value of $r$ at the middle of range 1, shown in figure 4.14. Now, each simulation at the specific values of $r$ can be redone applying the averaged optimizing factor $f_{opt}$ to $\theta_x$ and $\theta_y$, and the new error recomputed. The resulting density functions and error evaluation are shown in figure 4.15 for the case of a cylindrical tolerance zone. The related quantitative data is summarized in table 5.1. The $\chi^2$ error is computed between the analytical and the reality-based surfaces with and without specific and averaged optimizing factors. We find that the use of this optimization method can bring a significant reduction of the overall error, a figure of 37% in our example.

We note that since similar relations exist between $r$ and $\theta_x$, and $r$ and $\theta_y$, we can assume that a similar optimizing factor can be found and applied to the range of $r$, given by $r_{max}$ in the analytical solution. Note that this will in turn modify the distribution that was used to find the values of $r$ at which we have computed the specific optimizing factors. However, the resulting error is small compared to the approximations already
<table>
<thead>
<tr>
<th>No Optimizing Factor</th>
<th>Analytical</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>max = 5.13</td>
<td>max = 0.34</td>
</tr>
<tr>
<td>Specific Optimizing Factor (1.00)</td>
<td>max = 5.13</td>
<td>max = 0.34</td>
</tr>
<tr>
<td>Average Optimizing Factor (0.92)</td>
<td>max = 6.05</td>
<td>max = 1.18</td>
</tr>
</tbody>
</table>

*Figure 4.15 A  Optimization at $r = 0.04 r_{max}$.*
\( r = 0.36 r_{\text{max}} \)

**SIMULATED**

\[ \text{max} = 8.33 \]

<table>
<thead>
<tr>
<th></th>
<th><strong>ANALYTICAL</strong></th>
<th></th>
<th><strong>ERROR</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>NO OPTIMIZING FACTOR</td>
<td>max = 5.94</td>
<td></td>
<td>max = 2.39</td>
</tr>
<tr>
<td>SPECIFIC OPTIMIZING FACTOR (0.84)</td>
<td>max = 8.39</td>
<td></td>
<td>max = 0.17</td>
</tr>
<tr>
<td>AVERAGE OPTIMIZING FACTOR (0.92)</td>
<td>max = 7.01</td>
<td></td>
<td>max = 1.32</td>
</tr>
</tbody>
</table>

*Figure 4.15 B  Optimization at \( r = 0.36 r_{\text{max}} \).*
Figure 4.15 C  Optimization at $r=0.80 \ r_{\text{max}}$. 
performed in the process. In this work, we optimize the ranges of \( r \), \( \theta_x \) and \( \theta_y \) equally with the same averaged optimizing factor. Its use in the above example shows very well that the proposed analytical solution's efficiency can be increased by a simple modification of its parameter ranges.

We have seen how a cylindrical tolerance zone, to which the axis of a circular feature is limited, can be represented with ranges of variations on four degrees of freedom associated with the axis: \( X \) and \( Y \) through \( r \), \( \theta_x \), and \( \theta_y \). This shape of zone is by far the most common of the position specification type. We now turn to its actual matrix representation, which can also be extracted from a concentricity or a runout tolerance specification.

<table>
<thead>
<tr>
<th>INSTANCE OF ( r ) % OF ( r_{\text{max}} )</th>
<th>4</th>
<th>36</th>
<th>80</th>
<th>OVERALL AVERAGE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- WITH NO OPTIMIZING FACTOR</td>
<td>4.8</td>
<td>198.7</td>
<td>3197.8</td>
<td>269.3</td>
</tr>
<tr>
<td>2- WITH SPECIFIC OPTIMIZING FACTOR (FACTOR VALUE)</td>
<td>4.8 (1.00)</td>
<td>1.1 (0.84)</td>
<td>106.0 (0.67)</td>
<td>7.3</td>
</tr>
<tr>
<td>3- WITH AVERAGE OPTIMIZING FACTOR (0.92)</td>
<td>46.8</td>
<td>67.0</td>
<td>2217.0</td>
<td>169.0</td>
</tr>
<tr>
<td>REDUCTION OF ERROR FROM 1- TO 3- (IN %)</td>
<td>-1000</td>
<td>34</td>
<td>69</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 4.1 Errors between analytic and reality-simulated densities at different values of \( r \).
4.1.2 Position

Figure 4.1 shows a common case of circular feature positioning. To represent the cylindrical tolerance zone with an homogeneous matrix transform and its associated errors, one will first have to attach a Cartesian frame to the axis of the tolerance zone, which is the nominal position of the actual feature axis, given by the basic dimensions. Assume we start from the intersection of the datum planes A, B, and C. A first transformation is necessary to attach the Cartesian frame on the middle of the zone axis: this transformation only translates the frame $F_0$ to the frame $F_1$, as shows figure 4.16. For the six degrees of freedom, we have therefore these nominal values:

$$X = D_1 \quad \theta_x = 0$$

$$F_{0,1} : \quad Y = -D_2 \quad \theta_y = 0$$

$$Z = -D_3 \quad \theta_z = 0$$

which translate into this homogeneous matrix transform:

$$F_{0,1} = \begin{bmatrix}
1 & 0 & 0 & D_1 \\
0 & 1 & 0 & -D_2 \\
0 & 0 & 1 & -D_3 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

and the associated errors in +/- format for the six degrees of freedom

$$\Delta X = 0 \quad \Delta \theta_x = 0$$

$$\Delta Y = 0 \quad \Delta \theta_y = 0$$

$$\Delta Z = 0 \quad \Delta \theta_z = 0$$
It is in the transformation from the frame \( F_1 \) to the frame \( F_2 \) that the work described in the previous section comes into play. The nominal values are

\[
\begin{align*}
X &= 0 & \theta_x &= 0 \\
F_{1,2} : \quad Y &= 0 & \theta_y &= 0 \\
Z &= 0 & \theta_z &= 0
\end{align*}
\] (4.10)

which translate into the homogeneous matrix transform

\[
F_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\] (4.11)

and the associated errors in +/- format

\[
\begin{align*}
\Delta X &= f_{\text{opt}} \times T_p / 2 & \Delta \theta_x &= f_{\text{opt}} \times T_p / L \\
E_{1,2} : \quad \Delta Y &= f_{\text{opt}} \times T_p / 2 & \Delta \theta_y &= f_{\text{opt}} \times T_p / L \\
\Delta Z &= 0 & \Delta \theta_z &= 0
\end{align*}
\] (4.12)

where \( T_p \) is the diameter of the cylindrical tolerance zone, \( L \) its height and \( f_{\text{opt}} \) the averaged optimizing factor found in section 4.1.1.

Figure 4.17 a) shows another common case of cylindrical tolerance zone resulting from a position specification on a circular feature. This time, the axis of the tolerance zone is referenced from the axis of another circular feature, which is used as a datum. Again, starting from the datum axis with frame \( F_0 \), a nominal transformation with no associated errors is used to translate to the frame \( F_1 \), positioned at the middle of
**Figure 4.16** Parameters for position using a datum reference frame.

exactly the same as in the above case. The nominal values are

\[ \begin{align*}
X &= 0 & \theta_x &= 0 \\
F_{1,2} : & \quad Y = 0 & \theta_y &= 0 \\
& \quad Z = 0 & \theta_z &= 0
\end{align*} \]

which translates into the homogeneous matrix transform

\[ F_{1,2} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \]

and the associated errors in +/- format

\[ \begin{align*}
\Delta X &= f_{\text{opt}} \times T_p / 2 & \Delta \theta_x &= f_{\text{opt}} \times T_p / L \\
E_{1,2} : & \quad \Delta Y = f_{\text{opt}} \times T_p / 2 & \Delta \theta_y &= f_{\text{opt}} \times T_p / L \\
& \quad \Delta Z = 0 & \Delta \theta_z &= 0
\end{align*} \]
4.1.3 Concentricity

As explained in chapter 3, the case of figure 4.17 a) could be tolerated using a concentricity specification, as shown in figure 4.17 b). The transformation between the axis of the cylindrical tolerance zone and the axis of the actual feature is the same as if it were a position specification except that $T_p$ is replaced by $T_c$. The nominal transformation is the identity matrix and the associated errors are

\[ \Delta X = f_{opt} \times T_c / 2 \quad \Delta \theta_x = f_{opt} \times T_c / L \]
\[ \Delta Y = f_{opt} \times T_c / 2 \quad \Delta \theta_y = f_{opt} \times T_c / L \]
\[ \Delta Z = 0 \quad \Delta \theta_z = 0 \]  \hspace{1cm} (4.13)

where $T_c$, similar to $T_p$, is the diameter of the cylindrical tolerance zone, $L$ its height and $f_{opt}$ the average optimizing factor found in section 4.1.1. We note that the only difference between the position and the concentricity specifications is that the former allows the use of modifiers that relax the constraints on the tolerance zone. This particularity is discussed in section 4.2.2.

4.1.4 Runout

We have seen in chapter 3 that a position control can be extracted from a runout specification. An example of such specification, resulting also in a cylindrical tolerance zone, is shown in figure 4.17 c) Again, the appropriate transformation to describe this zone is the identity matrix with the associated errors.
\[
\begin{align*}
\Delta X &= f_{opt} x T_R / 2 \\
\Delta Y &= f_{opt} x T_R / 2 \\
\Delta Z &= 0 \\
\Delta \theta_x &= f_{opt} x T_R / L \\
\Delta \theta_y &= f_{opt} x T_R / L \\
\Delta \theta_z &= 0
\end{align*}
\] (4.14)

where $T_R$ is the diameter of the cylindrical tolerance zone defined by the runout notation, $L$ is the length of the zone and $f_{opt}$ the average optimizing factor found in section 4.1.1. We note that $T_R$ is the measuring indicator's maximum movement as the part is rotated about the datum axis as show in figure 3.16. Such a limit would allow the position of the actual feature's axis, if considered as the only existing variation, to be anywhere in a cylindrical zone of diameter $T_R$, similarly to $T_c$ and $T_p$.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.17.png}
\caption{Parameters for various tolerance specifications.}
\end{figure}
4.2 Size

As seen in chapter 3, the size tolerance specification is used to control the size of a planar feature, as the size of a cube in a particular direction, or of a circular feature, as the diameter of a cylinder or a hole.

When components are assembled, the position and orientation of a feature of concern somewhere in a sub-assembly will be influenced mainly by planar feature types, rather than circular ones. If two cubes are stacked together for example, the position of the top surface of the top cube, with respect to the workstation table, will depend on the size of each cube in the vertical direction. The analysis of such variations can be separated in two and three-dimensional cases, the former dealing with the position of an edge, the latter dealing with the position of a surface by extending the two-dimensional case.

Although size variation of a circular feature will have less impact on the position and orientation of various features in a sub-assembly, it can still influence in two specific cases: one of which, regarding the use of modifiers, is discussed in this chapter, while the other one, regarding mating of parts, is the main object of chapter 5.

We start by discussing planar features of size in section 4.21, circular features follow in section 4.22.
4.2.1 Planar features

When planar features of slabs are stacked, one of the surfaces can be considered as a datum surface, while the other one will vary in position with respect to the resulting datum plane. Figure 4.18 shows how one can view the zone allowed by the size specification with a transformation between two Cartesian frames, one attached to the middle of the tolerance zone, the other one to the middle of the actual surface, considered as a planar and rigid body. Figure 4.18 shows the surface at nominal position, which makes the two Cartesian frames coincident.

Figure 4.18 Parameters for planar size tolerance.
The task at hand is to model the real variations in position and orientation of this surface within the tolerance zone with a homogeneous matrix transform and associated errors. We have to define what degrees of freedom can vary, and what are the ranges and distribution shapes of these variations. To do so, it is possible to think the problem through in two dimensions first, and then extend the reflection in the more realistic three-dimensional space. This is a similar approach to the problem of positioning an axis in a cylindrical tolerance zone where, in the parameter space, we were considering cases at fixed values of \( r \). In the size tolerance zone, we can easily limit ourselves to two variables in the more intuitive \( \{ X, Y, Z \} \) space, and then evaluate the impact on a specific parameter space.

Two dimensions: edges

Consider for a moment the front view of the assembly in figure 4.18. This two-dimensional view of the cube clearly constrains the position of the top edge with respect to the datum line created by the datum edge. In the three-dimensional space created by the \( X \) axis, coming out of the page, the \( Y \) axis horizontally positioned, and the \( Z \) axis vertically positioned, the edge has only two degrees of freedom: a translation along the \( Z \) axis and a rotation \( \theta_x \) about the \( X \) axis. For now, since \( \theta_y \) and \( \theta_z \) are fixed, we simply note the possible rotation as \( \theta \). We now want to consider the ranges and the distributions of these two variates \( Z \) and \( \theta \). The procedure to do so is very similar to the one used in the case of the cylindrical tolerance zone. We first find the limiting boundary in the parameter space \( \{ Z, \theta \} \), we then evaluate the probability density of these variables by performing a Monte-Carlo simulation according to a model of reality, and we finally
compare this model to the analytical density resulting from the proposed tolerance analysis algorithm.

Boundary

Taken independently, the range of $Z$ is $+/- T_b$, with $2T_b$ as the height of the tolerance zone, while the range of $\theta$ is $+/- 2T_b/L$, where $L$ is the length of the tolerance zone. This last expression is obtained from a small angle approximation. We also note that the variation on $L$ is considered to have a negligible impact on the value of the range of $\theta$. However, these ranges cannot be considered separately, since if both variates were at their maximum value, a portion of the edge would be out of the tolerance zone, as shown in figure 4.19. For this reason, we define a relation between $Z$ and $\theta$ which results in a parallelogram (a diamond-like shape) in the two-dimensional ($Z$, $\theta$) parameter space, as shown in figure 4.20. This parallelogram is therefore a boundary, resulting from the combined ranges, out of which the parameters (degrees of freedom) of the edge movement would result in a non-respect of the size tolerance specification.

![Parallelogram Diagram](image)

Figure 4.19  Extreme values for variates in size specification.
Figure 4.20 Maximum boundary in parameter space for size specification (2D).

**Probability Density**

Another concern is the actual distribution of the parameters within this boundary. As in the previous section, we use the normal distribution and consider it realistic for both \( Z \) and \( \theta \), as long as they are kept dependent in order to respect the maximum boundary. As a model of the real joint probability density of the two variables, we have performed Monte-Carlo simulations that randomly and normally generated each of the variables, the second one generated being limited to the range left by the the first one, to create the actual dependency. Any instance generated outside the 6-\( \sigma \) limits were disregarded to respect the tolerance zone. Each instance of value pair of generated values were then added appropriately to create a histogram of the instances above the plane created by the parameter space \( \{ Z, \theta \} \). The resulting surface was finally normalized to insure a total cumulative distribution (volume under the surface) of 1. The results are shown in figure 4.21, along with the analytical, joint probability
density surface resulting from the proposed tolerance analysis software, in which the variables are considered independent and normally distributed.

**Figure 4.21** Comparison and optimization of 2D size specification: simulated versus analytical probability densities of joint variables.
To compare the analytic solution with our model of reality, we look at the chi-square error between the two surfaces. This value can be used to find an optimizing factor that can be multiplied to the range values used in the analytical solution. We have found this factor $f_{opt}$ to be 0.95, which resulted in a 20% decrease in the error between the two surfaces. We now extend this reflection to the three-dimensional case, the object of the next section.

**Three dimensions: surfaces**

Considering again figure 4.18, we now look at all possible movement of the tolerated surface in the $\{X, Y, Z\}$ domain. We see that a translation along the $Z$ axis, a rotation $\theta_x$ about the $X$ axis, and a rotation $\theta_y$ about the $Y$ axis are the only required degrees of freedom. We have seen in the two-dimensional case the maximum ranges for $Z$ and $\theta_x$, and we now add the possible range $\pm 2T_u/L$ for $\theta_y$. We want to consider the combined ranges and the distributions of these three variates $Z$, $\theta_x$, and $\theta_y$.

These ranges cannot be considered separately, since if all variates were at their maximum value, a portion of the surface would be outside of the tolerance zone, as shown in figure 4.22. For this reason, we define the relations between $Z$, $\theta_x$, and $\theta_y$, resulting in a diamond in the three-dimensional $\{Z, \theta_x, \theta_y\}$ parameter space, as shown in figure 4.23. This diamond is therefore a boundary, resulting from the combined ranges, out of which the parameters (degrees of freedom) of the surface movement would result in a non-respect of the size tolerance specification.
Figure 4.22 Extreme values for variates in size specification (3D).

We recall that the result of the analytical solution for three independent, normally distributed variates is an ellipsoid in a three-dimensional parameter space; and that the probability density function of the variates within this ellipsoid is defined by the three variate's joint normal distribution. We note also that a visual representation of such an ellipsoid is usually the 6-σ surface inside which 97.1% of the probability is found. In essence, this ellipsoid is comparable to the diamond of the reality-based model; the object of this work is to approximate the diamond with the ellipsoid, of which we can modify the parameter values to reduce the approximation error.

The extremities of the diamond cross the Z, θx, and θy axes at the same points the unmodified ellipsoid crosses these axes. By modifying the length of the principal axes of the ellipsoid, i.e. by specifying different ranges to the three variates, one can better approximate the diamond. Figure 4.24 demonstrates the concept. Intuition might suggest that if one reduces the ellipsoid's size so that difference in volume of the two
shapes in minimized, then one will have found the optimum analytic solution. That would be true if the variates were uniformly distributed within both regions, but it is not the case, as shows figure 4.21 relating to the two-dimensional case in the precedent section.
As in the problem of finding the distribution of the variates for a cylindrical
tolerance zone, we fix one of the variates and we evaluate the difference between the two
probability density surfaces, the joint normal analytic and the reality-based
simulation, above the plane formed by the other two variates. This is equivalent to the
two-dimensional case where \( \theta_y \) is fixed. We have found the optimizing factor \( f_{opt} \) for
the two-dimensional case and since the relations between all pairs of variates are
similar to the one between \( Z \) and \( \theta_x \), we consider that this factor can be applied to all
variates equally in the three-dimensional problem.

\[ \]

Figure 4.24 Approximation of reality-model's diamond by the analytical
solution's ellipsoid.
Hence, in figure 4.18, starting from the Cartesian frame attached to the middle of the datum plane, the transformation leading to the frame attached to the middle of the tolerance zone contains only a translation along $Z$ of the nominal size dimension, without any associated error

$$F_{0,1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & D \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$  

And, from the Cartesian frame attached to the middle of the rectangular tolerance zone of dimensions $L_x$ by $L_y$ by $2T_z$, its lengths along respectively the $X, Y$ and $Z$ axes in the $\{X, Y, Z\}$ domain, the nominal transformation values to the middle of the actual surface are

$$X = 0 \quad \theta_x = 0$$
$$F_{1,2} : \quad Y = 0 \quad \theta_y = 0$$
$$Z = 0 \quad \theta_z = 0$$

which translates into the homogeneous matrix transform

$$F_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the associated errors in +/- format for the six degrees of freedom.
\[
\begin{align*}
\Delta X &= 0 & \Delta \theta x &= f_{opt} \times 2T_e / L_y \\
E_{1,2} : & \quad \Delta Y = 0 & \Delta \theta y &= f_{opt} \times 2T_e / L_x \\
\Delta Z &= f_{opt} \times T_e & \Delta \theta z &= 0
\end{align*}
\] (4.15)

where \( f_{opt} \) is equal to 0.95 for this size tolerance case. We note that here, \( T_e \) is already in the +/- format, also called bilateral, unlike the tolerance notations for position, concentricity and runout specifications using respectively \( T_p, T_c, \) and \( T_R \) for the full size of the cylindrical tolerance zone (diameter).

4.2.2 Circular feature: the effect of the MMC principle

A "maximum material condition" (MMC) modifier is often specified in the position specification associated with a circular feature. In such a case, the diameter of the position tolerance zone can change depending on the actual size of the feature. The reader is referred to chapter 3 for more details on this concept. Since the position tolerance zone diameter itself becomes conditional to the actual size of the feature, it is appropriate to introduce the concept of conditional probability. Of course, when performing statistical tolerance analysis, the actual size of the feature is not known. Yet, since it is distributed statistically, the parameters of its distribution can be considered in the computation to find a conditional probability on the position; in other words, one could look for the probability of a peg position error being smaller than a specified value, given the probability of the size of the peg being smaller from its nominal specification.
In using the proposed tolerance analysis algorithm, one is concerned with computing the variance on the position of the circular feature conditional to its size, since the conditional probability itself will automatically result from the use of the conditional variance. To compute it, we use the analogy with the case of a peg mating a hole which is discussed in more details in chapter 5.

In the case of a peg mating a hole, we realize that the actual peg’s position will distributed within the space available in the hole. In other words, the axis of the peg will be positioned somewhere in a cylindrical variation zone which diameter is equal to the clearance between the peg and the hole. This clearance is in turn directly related to the relative size of the two mating features. This sequence of dependency makes the position of the peg a function of, among others, its diameter size: it is also the case when one wants to describe the tolerance zone on the position of a peg specified with a modifier.

To represent this analogy, we compare the clearance between two feature with the size of a single feature. For example, figure 4.17 d) shows a cylindrical feature specified in position with respect to another cylindrical feature. This example is the same as the one in figure 4.17 a) except that a modifier is added to the specification. This additional notation makes it necessary to compute a variance on the position which will be conditional to the statistical size of the feature:

\[ \sigma_{cp}^2 = g(E(S), \sigma_e^2) \]  \hspace{1cm} (4.16)
where $\sigma_{ep}^2$ is the feature's axis conditional position variance, $E(S)$ is the feature size (diameter) expectation, and $\sigma_{s}^2$ is the feature size variance.

Figure 4.25 demonstrates the analogy between the clearance in a peg mating a hole and a peg size variation when describing a position tolerance zone. When the size of the peg is at its maximum value, its axis can only be positioned in a tolerance zone of diameter $T_p$. On the other hand, when the size is at its minimum value, the axis can be in a greater tolerance zone, of diameter $(T_p + R_s)$ where $R_s$ is the allowable range of

**MMC (MAXIMUM MATERIAL CONDITION):**

**LMC (LEAST MATERIAL CONDITION):**

*Figure 4.25 Analogy between MMC principle and peg-hole position variation.*
peg size variation \(2T_p\). In our analogy, we consider \(T_p\) and \((T_p + R_s)\) to be equivalent to respectively the minimum and maximum values the clearance can take in the case of a peg mating a hole. We can now define the stochastic parameters of the equivalent clearance \(C\), the expectation \(E(C)\) and the variance \(\sigma_c^2\):

\[
E(C) = T_p + \text{size}_{\text{max}} - E(\text{size})
\]

\[
\sigma_c^2 = \sigma_s^2 + \sigma_p^2
\]  

(4.17)  

(4.18)

where \(T_p\) is the position tolerance usually given as the diameter of the position tolerance zone, \(\sigma_s^2\) is the size variance, and \(\sigma_p^2\) is the unconditional position variance at maximum size. Considering both distributions normal, we compute the variances using

\[
\sigma_s^2 = \left(\frac{R_s}{6}\right)^2, \quad \sigma_p^2 = \left(\frac{T_p}{6}\right)^2
\]

(4.19)

where \(R_s\) is the range of variation on the size of the peg and \(T_p\) is the position tolerance specification (diameter of the tolerance zone). It can be shown (see chapter 5) that:

\[
\sigma_p^2 = \frac{1}{9} \left(\sigma_s^2 + [E(X_R)]^2\right),
\]

(4.20)

we find that \(\sigma_p^2\) is effectively a conditional variance that will note here with \(\sigma_{cp}^2\), and that the parameters for \(X_R\), the variable defined as half the clearance, are found using the equivalent clearance parameters as computed above:

\[
E(X_R) = \frac{E(C)}{2} = \frac{T_p + \text{size}_{\text{max}} - E(\text{size})}{2}
\]

(4.21)
\[ \sigma_{x_r}^2 = (1\frac{1}{2})^2 \sigma_e^2 = \frac{\sigma_e^2 + \sigma_p^2}{4}. \] (4.22)

Equation 4.21 and 4.22 into 4.20 yields

\[ \sigma_{cp}^2 = \frac{1}{36} \left( \sigma_e^2 + \sigma_p^2 + \left[ T_p + \text{size}_{max} - E(\text{size}) \right]^2 \right) \] (4.23)

Equation 4.23 results in a conditional variance which is greater than the variance resulting from the unconditional position tolerance alone, \( \sigma_p^2 \). For example, with a size variation range \( R_s \) which is about twice the position tolerance \( T_p \), \( \sigma_{cp}^2 \) is about twice as big as \( \sigma_p^2 \); for \( R_s \) about half \( T_p \), \( \sigma_{cp}^2 \) is about 10% larger than \( \sigma_p^2 \); finally, if the MMC modifier is disregarded, i.e. if \( R_s = \sigma_e^2 = 0 \) and \( E(\text{size}) = \text{size}_{max} \), then one finds \( \sigma_{cp}^2 \) to be about 3% larger than \( \sigma_p^2 \). We conclude that the use of equation 4.23 is justified since it is compatible with the MMC principle which allows for more position variance if the peg size departs from its maximum value.

Although the above discussion uses the case of position and size specifications for a peg, a similar reflection can be made with specifications on a hole. In that case, a MMC modifier would allow the position tolerance zone to be greater if the hole is at its maximum size. But since the position tolerance is specified at MMC, i.e. when the hole size is at its minimum, we also find that the conditional variance is greater than the variance resulting from the unconditional position tolerance alone. Therefore, equation 4.23 is also applicable to define the conditional variance of the position of the hole axis.
Effect of MMC on the matrix transform of a circular feature position specification

When a MMC modifier is specified, we use the result given by equation 4.23 to modify the errors associated with the matrix transform representing the position of a circular feature. In defining the new errors, the position tolerance $T_p$ of equations 4.12 is replaced by a conditional position tolerance $T_{cp}$, which we define with

$$T_{cp} = 5 \sigma_{cp} \quad (4.24)$$

given the assumption of normal distributions. The homogeneous matrix transform will remain the identity matrix as in equation 4.11, but the associated errors in +/- format become

$$\Delta X = f_{opt} \times T_{cp} / 2 \quad \Delta \theta_x = f_{opt} \times T_{cp} / L$$

$$\Delta Y = f_{opt} \times T_{cp} / 2 \quad \Delta \theta_y = f_{opt} \times T_{cp} / L$$

$$\Delta Z = 0 \quad \Delta \theta_z = 0 \quad (4.25)$$

where $T_{cp}$ is computed with equations 4.24 and 4.23, and the optimizing factor for a cylindrical tolerance zone $f_{opt}$ is equal to 0.92, as computed in section 4.1.1.
4.3 Orientation

Orientation specifications include parallelism, perpendicularity and angularity specifications. The matrix transform representation of these controls are very similar; they all involve constraints on two rotational degrees of freedom, a subset of the degrees of freedom controlled by the planar size tolerance of section 4.1.1. Figure 4.26 shows a component with all three orientation specifications as refinements of size controls, the usual application of such specifications.

![Diagram of orientation specifications]

Figure 4.26 Parameters for orientation specifications as refinement of size.
4.3.1 Parallelism

As we saw in section 4.2.1, parallelism control of two planar surfaces is already achieved by a specification of size. In such a case, the height of the tolerance zone $2T_s$ dictates the ranges in which $\theta_x$ and $\theta_y$ are allowed to vary. If a parallelism specification is added to a size control, it is because although the size tolerance might allow a $2T_s$ variation range along Z, the corresponding angular variations would probably affect product functionality or ease of assembly. When such a concern is justified, the angular errors associated with the identity matrix transform, equations 4.15, are modified to further constrain $\theta_x$ and $\theta_y$ by using the additional parallelism tolerance $T_{pa}$ (the total height of the orientation zone) instead of the larger size tolerance zone $2T_s$. The errors associated with the identity matrix transform to go from frame $F_1$, attached to the middle of the tolerance zone, and $F_2$, attached to the actual surface, become:

$$\Delta X = 0 \quad \Delta \theta_x = f_{opt} x T_{pa} / L_y$$
$$\Delta Y = 0 \quad \Delta \theta_y = f_{opt} x T_{pa} / L_x$$
$$\Delta Z = f_{opt} x T_s \quad \Delta \theta_z = 0$$

(4.26)

where $L_x$ and $L_y$ are the lengths of the tolerance zones respectively in the X and Y directions (same for size and parallelism zones), and the optimizing factor $f_{opt}$ is equal to 0.95 as computed in section 4.1.2.

4.3.2 Perpendicularly

The same reflection applies to the perpendicularity specification. The perpendicularity tolerance $T_{pa}$ will be dictating the angular variation ranges $\Delta \theta_x$ and
\( \Delta \theta_y \) while the size tolerance range dictates the linear variation range \( \Delta Z \). We establish a convention on the orientation of the Cartesian frame attached to the middle of the tolerance zone to simplify and standardize matrix transform representations: in all orientation specifications, we will assume the linear variation to be always along the \( Z \) axis. When, for any reason, there is no size specification dictating that variation, the perpendicularity tolerance \( T_{pe} \) is used and the error on the range becomes \( \Delta Z = f_{opt} \times T_{pe}/2 \). This last remark also applies to the other orientation specifications.

Therefore, from the Cartesian frame \( F_0 \) attached to the middle of the datum surface, the nominal transformation values to the frame \( F_1 \) attached to the middle of the tolerance zone are

\[
F_{0,1} : \begin{align*}
X &= 0 \\
Y &= -D_3/2 \\
Z &= D_1/2 \\
\theta_x &= \pi/2 \\
\theta_y &= 0 \\
\theta_z &= 0
\end{align*}
\]

where \( D_1 \) and \( D_3 \) are the size dimensions respectively along the \( Z \) and \( Y \) directions of the datum Cartesian frame, as shown in figure 4.26. These values translate into this homogeneous matrix transform:

\[
F_{0,1} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(\pi/2) & -\sin(\pi/2) & -D_3/2 \\
0 & \sin(\pi/2) & \cos(\pi/2) & D_1/2 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & -1 & -D_3/2 \\
0 & 1 & 0 & D_1/2 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

(4.27)

And, from the Cartesian frame \( F_1 \) attached to the middle of the rectangular tolerance zone of dimensions \( L_x \) by \( L_y \) by \( 2T_z \), its lengths along respectively the new \( X, Y \) and \( Z \) axes, the nominal transformation values to the middle of the actual surface are

\[
\begin{align*}
X &= 0 & \theta_x &= 0 \\
F_{1,2} : \quad Y &= 0 & \theta_y &= 0 \\
Z &= 0 & \theta_z &= 0
\end{align*}
\]

which translates into the homogeneous matrix transform

\[
F_{1,2} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

and the associated errors in +/- format for the six degrees of freedom

\[
\begin{align*}
\Delta X &= 0 & \Delta \theta_x &= f_{opt} \times T_{ps} / L_y \\
E_{1,2} : \quad \Delta Y &= 0 & \Delta \theta_y &= f_{opt} \times T_{ps} / L_x \\
\Delta Z &= f_{opt} \times T_z & \Delta \theta_z &= 0
\end{align*}
\]

(4.28)

where \( f_{opt} \) is equal to 0.95. We recall that \( T_z \) is the linear +/- variation along the \( Z \) axis of frame \( F_2 \), while \( T_{ps}/2 \) is used for the angular +/- variations about the \( X \) and \( Y \) axes. In the perpendicularity specification of figure 4.26, one finds that \( L_x \) is equal to \( D_4 \) and \( L_y \) is equal to \( D_1 \).
4.3.3 Angularity

Again, the same method is used to represent the slanted tolerance zone.

From the Cartesian frame \( F_0 \), attached to the middle of the datum surface, the nominal transformation values to the frame \( F_1 \), attached to the middle of the tolerance zone are

\[
X = 0 \quad \theta_x = -\alpha \\
F_{0,1} : \quad Y = D_2 / 2 \quad \theta_y = 0 \\
Z = D_1 / 2 \quad \theta_z = 0
\]

where \( D_1 \) and \( D_2 \) are the size dimensions respectively along the \( Z \) and \( Y \) directions of the datum Cartesian frame, and \( \alpha \) is the nominal angle dimension, as shown in figure 4.26.

These values translate into this homogeneous matrix transform:

\[
F_{0,1} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(-\alpha) - \sin(-\alpha) & D_2 / 2 & 0 \\
0 & \sin(-\alpha) & \cos(-\alpha) & D_1 / 2 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(\alpha) & \sin(\alpha) & D_3 / 2 \\
0 & -\sin(\alpha) & \cos(\alpha) & D_1 / 2 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

(4.29)

And, from the Cartesian frame \( F_1 \) attached to the middle of the rectangular tolerance zone of dimensions \( L_x \) by \( L_y \) by \( 2T_x \cos(\alpha) \), its lengths along respectively the \( X, Y \) and \( Z \) axes of frame \( F_1 \), the nominal transformation values to the middle of the actual surface are
\[ X = 0 \quad \theta_x = 0 \]
\[ F_{1,2} : \quad Y = 0 \quad \theta_y = 0 \]
\[ Z = 0 \quad \theta_z = 0 \]

which translates into the homogeneous matrix transform

\[
F_{1,2} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

and the associated errors in +/- format for the six degrees of freedom

\[
\Delta X = 0 \quad \Delta \theta_x = f_{opt} \times T_a / L_y \\
\Delta Y = 0 \quad \Delta \theta_y = f_{opt} \times T_a / L_x \\
\Delta Z = f_{opt} \times T_b \cos(\alpha) \quad \Delta \theta_z = 0
\]

where \( f_{opt} \) is equal to 0.95. We recall that \( T_a \) is the specified linear +/- variation along the \( Z \) axis of datum frame \( F_0 \), while \( T_a / 2 \) is used for the angular +/- variations about the \( X \) and \( Y \) axes of frame \( F_1 \). In the angularity specification of figure 4.26, one finds that \( L_x \) is equal to \( D_4 \) and \( L_y \) is equal to \( (D_3 - D_2) / \cos(\alpha) \).

4.3.4 Perpendicularity included in runout specification

Figure 4.27 shows a surface controlled with a total runout specification. This implies that the surface must be perpendicular to the datum axis defined by the actual datum (circular) feature. As the part is rotated 360° about the datum axis, an Indicator
Figure 4.27 Total runout tolerance on surface perpendicular to datum axis (adapted from [Foster-86]).

positioned parallelly the the datum, measuring the surface lateral position at any distance from the axis, must not show a Full Indicator Movement (FIM) greater than the runout tolerance zone $T_R$.

Once one has positioned a Cartesian frame $F_1$ on the middle of the tolerance zone with the $Z$ axis along the datum axis, the nominal transformation values to the Cartesian frame $F_2$ attached to the actual surface are all 0. This is equivalent to the identity matrix transform. As for the corresponding errors in +/- format for the six degrees of freedom:

$$
\begin{align*}
\Delta X &= 0 \\
\Delta Y &= 0 \\
\Delta Z &= f_{opt} \times T_s
\end{align*}
\quad \begin{align*}
\Delta \theta_x &= f_{opt} \times T_R / L_y \\
\Delta \theta_y &= f_{opt} \times T_R / L_x \\
\Delta \theta_z &= 0
\end{align*}

(4.31)

where $f_{opt}$ is equal to 0.95. $T_s$ is the size tolerance giving the linear +/- variation along the $Z$ axis of frame $F_1$, while $T_R/2$ is used for the angular +/- variations about the $X$ and $Y$ axes of frame $F_1$. In the runout specification of figure 4.27, one finds that $L_x$ and $L_y$ are both equal to the diameter of the controlled surface.
We note that a circular runout specification theoretically does not control the perpendicularly of a surface, since it controls only a single circular element of the surface. On the other hand, if a FIM measurement at the outermost radius of the surface is out of circular runout tolerance, there is a strong possibility that measurements at smaller radii would also be unacceptable, also affecting orientation: in light of this possible interpretation, one should investigate the real intent conveyed by this tolerance.

Conclusion

In this chapter, we have seen how standard tolerance zones associated with features of individual components could be represented with homogeneous matrix transforms. We have described the methods to represent position and orientation variation zones, whether they are explicitly controlled, as in location or orientation specification types, or implicitly controlled, as in size and runout types.

A given tolerance zone is always represented starting from a Cartesian frame (that we have called \( F_1 \)) attached to the middle of the tolerance zone. From this frame, an identity transform matrix is specified to describe the transformation to the Cartesian frame (that we have called \( F_2 \)) attached to the middle of the actual feature (axis or surface). The nominal position and orientation of the actual feature is therefore coincident with the middle of the tolerance zone. To represent the allowed variations, we identify a subset of the six possible degrees of freedom, and we assign a range of variation for each of the variates in the subset. The range of variation is specified in the +/- format around the nominal value, which is null.
Table 4.2: Summary of errors associated with homogeneous matrix transforms.

<table>
<thead>
<tr>
<th>TYPES</th>
<th>CHARACTERISTIC</th>
<th>CONTROL FRAME</th>
<th>GEOMETRY</th>
<th>ERRORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Position MMC</td>
<td>ρ T A</td>
<td></td>
<td>same as below but replace T by T&lt;sub&gt;op&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>Position</td>
<td>ρ T A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Concentricity</td>
<td>ρ T A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Runout</td>
<td>Circular runout</td>
<td>T&lt;sub&gt;a&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total runout (on surface // to datum axis)</td>
<td>T&lt;sub&gt;a&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total runout (on surface ⊥ to datum axis)</td>
<td>T&lt;sub&gt;a&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planar size</td>
<td>Distance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orientation</td>
<td>Parallelism</td>
<td>H T&lt;sub&gt;p&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perpendicular</td>
<td>L H T&lt;sub&gt;k&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Angularity</td>
<td>L T&lt;sub&gt;e&lt;/sub&gt; A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
T_{op}^2 = \alpha^2 + \sigma^2 \times (T \times S_{\text{max}} + E(S))^2
\]

\[
\Delta X = T_{opt} / T \quad \Delta Y = T_{opt} / T \quad \Delta Z = 0
\]

\[
\Delta Z = 0
\]

\[
T = T_{op} = T_{k}
\]

\[
L = \text{length of cylindrical tolerance zone}
\]

\[
f_{\text{ext}} = 0.92
\]

\[
\Delta X = 0 \quad \Delta Y = 0 \quad \Delta Z = 0
\]

\[
\Delta X = 0 \quad \Delta Y = 0 \quad \Delta Z = 0
\]

\[
T = T_{op} = T_{k}
\]

\[
L_x \text{ and } L_y \text{ are length and width of tolerance zone}
\]

\[
f_{\text{ext}} = 0.95
\]

\[
\Delta X = 0 \quad \Delta Y = 0 \quad \Delta Z = 0
\]

\[
T = T_{op} = T_{k}
\]

\[
L_x \text{ and } L_y \text{ are length and width of tolerance zone}
\]

\[
f_{\text{ext}} = 0.92
\]
Identifying the range of variation for each variate is critical: the real joint density of the (usually three) dependent variates has to be approximated by a joint normal density of (usually three) independent variates. The considerations of normal distribution and independency of variates are required by the use of a closed loop tolerance analysis algorithm, of which they are important assumptions. We have done that for all pertinent specification types.

The critical data resulting from these transformations are the errors associated with the identity matrix transforms. These errors are specified in terms of the parameters used in the standard notation of Y14.5M-1982. Table 4.2 summarizes the work we have done. Location errors associated with position of circular features, concentricity, and runout specifications are given respectively by equations 4.12, 4.13, and 4.14. Errors associated with the size of planar features are given by equation 4.15, while the impact of the MMC modifier is described by equation 4.25. Parallelism, perpendicularity, and angularity errors are given respectively by equations 4.26, 4.28, and 4.30. Finally, if a runout specification is used to control a surface perpendicular to the datum axis, one can extract a tolerance of perpendicularity with the errors given by equation 4.31.

In the next chapter, we describe the transformations involved with a frequent case of mating variation, a peg in a hole.
Chapter 5

Matrix Transform Representation of Mating Variations: Peg in a Hole
Introduction

Mating variations are another source of variations when considering tolerance analysis in assemblies. Unlike individual component variations controlled by standard tolerance specifications, mating variations have not been analyzed as comprehensively in existing literature. They involve the nature and the behavior of the contact between two or more components, through single or multiple surfaces or features on each component.

In this chapter, we analyze only one important case: the representation of the position variation of a peg which is in a hole.

5.1 Problem statement

The relative sizes of the mating diameters will establish an important parameter influencing the nature of the contact between the peg and the hole. If the size of the peg is greater than the size of the hole, we are dealing with an interference fit. If the inverse is found, the we are dealing with a clearance fit. If the variations on the sizes are such that the fit can be sometime interference and sometime clearance, then we talk about a transition fit. It is only in clearance and transition fits that there can be a question on the stochastic position of the peg within the hole. In an interference fit, the axis of the peg will always be coincident with the axis of the hole, assuming both features reasonably circular. In this section, we will concentrate on the case of a clearance fit, in which the mating variation has the greatest impact on position and orientation of other
features in the assembly, since the related clearance usually has a greater means than in transition fits.

Since circular features are usually referred to by there axis, we establish a stochastic relationship between the axis of the hole and the axis of the peg. The work in this section is aimed at obtaining a homogeneous matrix transform and the associated stochastic errors to describe realistically the relative position and orientation of the peg axis in reference to the hole axis.

5.2 Definition of the variation zone

If the sizes of both the peg and the hole were fixed, the portion of the peg axis within the hole (of length L) would be confined to a fixed cylindrical variation zone, as shown by figure 5.1. However, the assumption of exact sizes is not realistic in existing manufacturing environments; the diameters (sizes) of the components have to be considered as statistical variables. As a result, the clearance itself is a statistical variable, and the radius of the variation zone $X_R$ shown in figure 5.1 also becomes a stochastic variable, whose distribution parameters are based on the parameters of the component diameter distributions. Finally, at the end of this dependence chain, the position of the peg's axis within this variable zone has its own distribution and associated parameters, which are function of the zone parameters. We therefore use the concept of conditional probability.
Figure 5.1 Variation zone for peg position in hole.

Assuming the use of distributions involving means and variances as parameters, we want to compute the conditional variance of the position of the peg as a function of the variance of the tolerance zone's diameter or radius:

\[ \sigma_{cp}^2 = g(E(C), \sigma_c^2) \]

or

\[ \sigma_{cp}^2 = g(E(X_R), \sigma_{X_R}^2) \]  

where \( \sigma_{cp}^2 \) is the conditional variance of the position of the peg axis, and \( E(C) \) and \( \sigma_c^2 \) are respectively the expectation and the variance of the clearance of between the peg and the hole (or \( E(X_R) \) and \( \sigma_{X_R}^2 \)). Equations 5.1 result in a parameter that can then be used by the proposed tolerance analysis algorithm to compute the probability of the peg's axis being within a defined zone, given the probability of the peg and hole sizes being
within defined ranges. [Bjørke-'89] has worked on this problem and we present the principal results below.

5.3 Derivation of the conditional variance

Figure 5.2 shows the variates used in this problem. $X_a$ and $X_b$ are the diameters of respectively the peg and the hole, while $X_R$ is the manufacturing-dependent radius of the position variation zone for the peg axis. $X_R$ is equal to half the clearance between the peg and the hole

$$X_R = \frac{X_b - X_a}{2}.$$  \hspace{1cm} (5.2)

The expectation $E(X_R)$ and variance $\sigma_{X_R}^2$ of $X_R$ can be found easily with

$$E(X_R) = \frac{E(X_b) - E(X_a)}{2} \hspace{1cm} \text{and} \hspace{1cm} \sigma_{X_R}^2 = \frac{\sigma_{X_b}^2 + \sigma_{X_a}^2}{4}.$$  \hspace{1cm} (5.3)

We now seek the expression giving $\sigma_{cp}^2$ as a function of these stochastic parameters. The vector, whose length is the distance of the peg axis from the cylindrical variation zone axis, has its direction uniformly distributed in $[0, 2\pi]$. If we find the parameters of its position when along the $X$ direction, they will be the same in the $Y$ direction, and these two directions are compatible with our need to input, in the proposed tolerance analysis algorithm, variation ranges only in terms of the 6 degrees of freedom $X, Y, Z, \theta_x, \theta_y, \text{and} \theta_z$. Our goal then is to determine the statistical
parameters along the X axis, that is, the unconditional marginal density \( g(X) \) with respect to the trivariate density \( f(X,Y,X_R) \). We can denote the bivariate conditional probability density of \( X \) and \( Y \ given X_R \) by \( g_2 (X, Y \mid X_R) \) which respects

\[
g_2 (X, Y \mid X_R) = \frac{f(X,Y,X_R)}{f_1(X_R)}
\]

(5.4)

where \( f_1(X_R) \) is the marginal probability density of \( X_R \) with respect to the trivariate density \( f(X,Y,X_R) \). Then, the univariate conditional marginal density \( g_1 (X \mid X_R) \) with respect to the trivariate density \( f(X,Y,X_R) \) is

\[
g_1 (X \mid X_R) = \int_{-\infty}^{\infty} g_2 (X, Y \mid X_R) dY
\]

(5.5)

And the unconditional univariate marginal density will be found from
\[ g(X) = \int_{-\infty}^{\infty} g_1(X|X_R)f_1(X_R)dX_R \]  

(5.6)

where the function \( g(\) is different from the one in equation 5.1. The first step is to set \( g_2(X, Y|X_R) \), the bivariate conditional probability density for \( X \) and \( Y \) given \( X_R \). We have many choices in choosing this probability density shape, but we will limit ourselves to assumptions of uniform or normal distributions. We start with the case of uniform distribution.

5.3.1 Uniform distribution

Assuming a uniform distribution for the position of the peg axis in the clearance-created variation zone, \( g_2(X, Y|X_R) \), the bivariate conditional probability density for \( X \) and \( Y \) given \( X_R \) becomes

\[ g_2(X, Y|X_R) = \frac{1}{\pi X_R^2} \quad \text{if} \quad X^2 + Y^2 \leq X_R^2 \quad \text{and} \]

\[ g_2(X, Y|X_R) = 0 \quad \text{if} \quad X^2 + Y^2 > X_R^2 \]  

(5.7)

Substituting equation 5.7 into equation 5.5, we obtain the univariate conditional density of \( X \) given \( X_R \).
\[ g_1(X \mid X_R) = \int_{-\sqrt{X_R^2 - x}}^{\sqrt{X_R^2 - x}} g_2(X, Y \mid X_R) \, dy = \int_{-\sqrt{X_R^2 - x}}^{\sqrt{X_R^2 - x}} \frac{1}{\pi X_R^2} \, dy \]  \quad (5.8)

Equation 5.8 results in

\[ g_1(X \mid X_R) = \frac{2}{\pi X_R^2} \sqrt{1 - \left(\frac{X}{X_R}\right)^2} \quad \text{if} \quad -X_R \leq X \leq X_R \quad \text{and} \]

\[ g_1(X \mid X_R) = 0 \quad \text{if} \quad |X| > |X_R| \]  \quad (5.9)

And by substituting equation 5.9 into equation 5.6, we get the unconditional univariate marginal density of \( X \)

\[ g(X) = \int_0^{X_R} \left( \frac{2}{\pi X_R^2} \sqrt{1 - \left[ \frac{X}{X_R} \right]^2} \right) f_1(X_R) \, dX_R . \]  \quad (5.10)

This relatively complex integral is solved in [Bjørke-89] by splitting the \( X \) variable in two variables \( X = X_R X_U \) with associated probability density \( f_1(X_R) \) and \( h(X_U) \). This allows to compare equation 6.10 to the convolution integral

\[ g(X) = \int_0^{X_R} \frac{1}{|X_R|} h\left(\frac{X}{X_R}\right) f_1(X_R) \, dX_R \]  \quad (5.11)
and infer the probability density of $X_u$, $h(X_u)$, such that equation 5.11 can be
identical to equation 5.10. Solving the integral of equation 5.10 then becomes
unnecessary: the knowledge of the probability density $h(X_u)$ allows to compute the
statistical parameters $E(X_u)$ and $\sigma_{X_u}^2$ associated with $X_u$, respectively 0 and 1/4.
Finally, knowledge of the statistical parameters of both $X_R$ and $X_u$ in the equation
$X = X_R X_u$ allows to find the statistical parameters of $X$

$$E(X) = 0 \quad \text{and} \quad \sigma_X^2 = \frac{1}{4} (\sigma_{X_R}^2 + (E(X_R))^2 )$$ (5.12)

Equations 5.3 give the parameters $E(X_R)$ and $\sigma_{X_R}^2$ of $X_R$, the variable
representing the radius of the cylindrical tolerance zone resulting from the clearance.
This variance, computed in the $X$ direction of the position of the peg axis, is equivalent in
all directions, including in $Y$. As a result, we now have the conditional variance of the
position of the peg axis $\sigma_{cp}^2$, assumed to be uniformly distributed, as a function of the
parameters of the variable $X_R$, directly proportional to the clearance $C$. The function
$g( \cdot )$ of equation 5.1 is now defined for one of the two possible cases. We now look at the
case where the position of the peg axis is assumed to be normally distributed.

5.3.2 Normal distribution

Through a similar process, we find function $g( \cdot )$ of equation 5.1 such that

$$\sigma_X^2 = \frac{1}{9} (\sigma_{X_R}^2 + (E(X_R))^2 )$$ (5.13)
where $\sigma_{x^2}$ is equivalent to $\sigma_{c_p^2}$. Equation 5.13 is valid in the case where we assume the position of the peg axis to be normally distributed within the cylindrical variation zone defined by the radius $X_R$, a statistical variable with distribution parameters resulting from the manufacturing process. We note that the variance resulting from the normal assumption is almost half the variance resulting from the uniform assumption.

5.4 Choice of distribution assumption for use in proposed tolerance analysis algorithm

We have seen that, in the proposed tolerance analysis algorithm, all variates are considered to be normally distributed, and that the propagated errors result in final variations that are normally distributed as well. These final variations are described by trivariate probability density ellipsoids constituting good approximations of the various models of reality we have proposed.

In the present problem of a peg in a hole, the above considerations imply that, regardless of the distribution shape assumption on which the conditional variance of the peg axis position was computed, the proposed algorithm will normally distribute it. Given this fact, an important question remains: since the conditional variance of the peg axis position is different depending on its distribution shape assumption, which assumption results in a variance that best approximates reality when used in the tolerance analysis algorithm?
To help us answer this question, we have evaluated the actual distributions of the peg position resulting from both assumptions. The first step of the process is to compute the conditional variances using each assumption. Then, for each of these variances, we evaluate the actual distributions by generating a great number of data points using a Monte-Carlo simulation. To simulate the proposed tolerance analysis algorithm, we use a random number generator that produces normally distributed data points based on a specified mean and variance.

To see if these simulated distributions comply reasonably with the nature of a clearance fit, we have also simulated the peg and hole sizes, resulting in the actual distribution of $X_R$, as defined by equation 5.2. We know that $X_R$ is equivalent to the (variable) radius of the peg position variation zone. As for the shape of the distribution for the sizes, we have simulated both normal and uniform distributions. A subject of ongoing discussions, the distribution on size of circular features finds realistic justifications for either shapes. For example, one could consider the cutter tool to wear in such a way that the part diameter varies gradually between the minimum and maximum accepted values, resulting in a uniform distribution. On the other hand, it is possible that the rate of wear of the tool is not linear, creating a distribution shape drifting away from uniform, possibly towards normal. In any case, this work does not intend to answer the question, and we can observe the impact of both shape on our simulations.

Figure 5.3 describes the parameters used in the Monte-Carlo simulations, while figure 5.4 shows the results for the four cases considered. In all these cases, a normal distribution of the peg position was generated (along any radial direction of the
variation zone); the normal distribution simulates the proposed tolerance analysis algorithm. The four cases differ only by the assumption used to compute the conditional variance of the peg position, and by the shape of the size distributions. So, the peg position is simulated with a normal distribution using a variance computed with:

a) the assumption of a normal peg position distribution with:
   i) the sizes normally distributed for the clearance simulation;
   ii) the sizes uniformly distributed for the clearance simulation; and

b) the assumption of a uniform peg position distribution with:
   i) the sizes normally distributed for the clearance simulation;
   ii) the sizes uniformly distributed for the clearance simulation.

For all these cases, the same numerical example was used: a peg of nominal diameter 0.99 (any length units) in a hole of nominal diameter 1.00. The nominal clearance is therefore 0.01, the clearance ratio alike, and the nominal radius of the variation zone
**Figure 5.4** Simulated density curves of peg position and radius of variation zone.

For the peg position is half the clearance, i.e. 0.005. We note immediately that the actual values have little impact on the results: numerical examples with clearance ratios that were one order greater or one order smaller resulted in the same conditional variance values, leading to the same simulated densities. Table 5.1 summarizes the results which are discussed below.
\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
ASSUMPTION & a) NORMAL ASSUMPTION & b) UNIFORM ASSUMPTION \\
ON PEG & $\alpha_p^2 = \frac{1}{4} K'$ & $\alpha_p^2 = \frac{1}{9} K'$ \\
PSN DISTRIBUTION & & \\
\hline
i) NORMAL SIZES & 7.7% & 1.5% \\
\hline
ii) UNIFORM SIZES & 12.7% & 5.8% \\
\hline
\end{tabular}
\end{table}

$K = \sigma_{X_R}^2 + E(X_R)^2$

PROBABILITY OF INTERFERENCE BETWEEN THE NORMALLY DISTRIBUTED PEG AND THE HOLE GIVEN:
1) THE ASSUMPTION OF THE PEG POSITION DISTRIBUTION USED TO COMPUTE THE CONDITIONAL VARIANCE $\alpha_p^2$;
2) THE DISTRIBUTION SHAPE OF THE SIZES OF THE PEG AND HOLE.

**Table 5.1** Simulation of tolerance analysis algorithm for peg-hole mate.

The probability density curves of figure 5.4 give an idea not only of the distribution of the peg axis position and of the variation zone radius, but also of the probability of theoretical interference between the peg and the hole. On each of the four plots, the curve centered at 0 is the density of the peg axis position, while the curve to its right, centered at the expectation of $X_R$, is the density of the radius of the variation zone in which the peg axis has to lie. The intersection between the two curves represents half the probability of interference between the peg and the hole. For example, if, for a particular instance, the axis position and the variation zone radius are such that the latter is smaller than the former, it is as if portions of the two solid components were into each other, an impossible state in practical situations of clearance fits. Therefore, to best simulate reality, we want to choose the assumption (for computation of the conditional variance) that minimizes the probability of interference.
Such an assumption is one of uniform distribution of the axis position, the result of which creates an interference probability of 1.5% (when sizes are normally distributed), about the fifth of what the normal assumption results in, at 7.7%. Of course, in a true model of reality, there would never be interference between the peg and the hole: we are now trying to simulate a situation where the peg is already mated with the hole. To simulate true reality, we would have to 1) distribute the peg position uniformly in a range obtained from a conditional variance computed with the assumption of a uniform distribution of the peg; and 2) distribute the sizes (diameters of peg and hole) simulating a finite distribution (unlike the normal one). No Monte-Carlo simulation results are presented for this (reality) model, given its predictability.

As for the distribution of the peg and hole sizes, we observe a widely known fact from figure 5.4: the convolution of two uniform distributions of similar ranges results in a triangular shape, approaching the normal bell-like shape. But most importantly, a uniform distribution on the sizes will increase the probability of interference; this is a predictable fact if we remember that the peg itself is limited to a normal distribution (from the tolerance analysis algorithm). In the case of uniform assumption for the peg distribution, the probability of interference is 5.8% if sizes are uniformly simulated, while it is 1.5% if sizes are normally simulated. Since the proposed tolerance algorithm does not make direct use of size information, it will result in relatively conservative propagated errors (larger than reality) since it "allows", based on our simulations, up to 5.8% interference.

Keeping in mind that we are looking for the best approximation of what reality results in, we suggest the use of a uniform assumption when computing the conditional
variance of the position of a peg in a hole $\sigma_{cp}^2$, given by equation 5.12. This conditional variance, used in conjunction with the proposed closed-loop tolerance analysis algorithm (limiting the variables to normal distributions), is the best approximation to the real, uniformly distributed, peg position model.

5.5 Matrix transform representation

The variation zone, in which the axis of a peg can find itself without creating interference with the hole, is of cylindrical shape. Accordingly, we can represent this variation with errors associated with an identity matrix transform similarly to the case of positional tolerance of circular features, described by equation 4.12. In defining the new errors, the position tolerance $T_p$ of equations 4.12 is replaced by a conditional position tolerance $T_{cp}$, which we define with

$$T_{cp} = 6 \sigma_{cp} = 6 \sigma_x$$

(5.14)

given the (actual) normal distribution of the peg position. $\sigma_x$ is equivalent to $\sigma_{cp}$ and is computed with equation 5.12. As can be seen in figure 5.1, the homogeneous matrix transform from the Cartesian frame $F_0$, attached to the (fixed) middle of the variation zone defined by the (variable) clearance, to the frame $F_1$, attached to the peg axis, will remain the identity matrix as in equation 5.11, but the associated errors in +/- format become
where $T_{cp}$ is computed with equations 5.14 and 5.12, $L$ is the depth of insertion of the peg once assembled, and the optimizing factor for a cylindrical tolerance zone $f_{opt}$ is equal to 0.92, as computed in section 4.1.1.

**Conclusion**

In this chapter, we have seen how to represent an important type of mate variation representation, a peg in a hole, using a homogeneous matrix transform and its associated errors, given by equation 5.15. We have seen that, since the variation zone in which the axis of the peg has to lie is variable in diameter (defined by the clearance between the peg and the hole), we must compute a conditional variance $\sigma_{cp}^2$ for the peg axis position. This variance results in a different value depending on the assumption of the shape of the position distribution. We know that the proposed tolerance analysis algorithm distributes all errors normally, yet the assumption of a uniform distribution for the peg position axis is more realistic. Monte-Carlo simulations have been performed to help decide under which assumption should the actual conditional variance be computed.

The results are such that we suggest the use of a uniform assumption, which results in less probability of theoretical interference. The conditional variance is given
by equation 5.12, and is translated into a range for appropriate description of the errors associated with the matrix transform representation of this mate variation.

In the next chapter, we make a case study showing the use of many matrix transforms (and associated errors) derived in this and previous chapters. We also discuss the use of part mating theory to evaluate the probability of success of a particular assembly step.
Chapter 6

Case Study
Introduction

In the previous chapters, we have described a method to represent component tolerances and assembly mate variations using homogeneous matrix transforms and their associated errors. This enables us to use a closed-loop form tolerance analysis algorithm developed for robotic applications by [Veitschegger and Wu-86], and adapted for assembly evaluation by [Jastrzebski-90]. In chapter 2, we have introduced the reader to part mating theory, whose knowledge allows to define the range of linear and angular errors, between a mating peg and a hole, outside which failure of the assembly step is probable.

In this chapter, we unite tolerance analysis based on matrix transforms and part mating theory results through a case study. We look at an impeller that contains a small number of parts so that our approach's validity can be verified with intuition and manual computations. This is presented in section 6.2, but first, section 6.1 presents in more details a method we can use to evaluate the probability of success of critical peg-hole assembly steps, one of which we find in the impeller assembly example.
6.1 Evaluation of assembly success

We have introduced the reader to the concept of fictitious "worst case plane" in section 2.5 (figure 2.5.4), containing a zone of acceptable error combination. This plane allows to plot the angular error $\theta_0$ against the lateral (or linear) error $\varepsilon_0$ between the entry of a hole and the mating extremity of the peg just before assembly (before one-point contact phase). We have suggested that these errors could be considered as being the ones in the directions of maximum errors found after performing three-dimensional tolerance analysis. The tolerance analysis method we have proposed implements statistical propagation of errors resulting in final errors that are normally distributed. This means that, for each axis of the worst case plane, we are given errors in terms of normal probability density, as shown in figure 6.1. One could find the resulting normal joint probability density, as in figure 4.9, but it might be easier, for a first evaluation, to consider the individual and independent (by assumption of the tolerance analysis algorithm) densities. If the two statistical variables are independent, we find that

$$\text{prob}\{ (- [W+C] \leq \varepsilon_0 \geq [W+C]) \& (- \frac{c}{\mu} \leq \theta_0 \geq \frac{c}{\mu}) \} = \text{prob}(- [W+C] \leq \varepsilon_0 \geq [W+C]) \times \text{prob}(- \frac{c}{\mu} \leq \theta_0 \geq \frac{c}{\mu})$$

(6.1)

where $[W+C]$ and $\frac{c}{\mu}$, as shown in figure 2.5.4, are the maximum values that $|\varepsilon_0|$ and $|\theta_0|$ can take to prevent unsuccessful mating of the peg and hole. $W$ is the sum of chamfer widths on the peg and hole, $C$ is the clearance between them, $\mu$ is the coefficient of friction, and $c$ is the clearance ratio. Equation 6.1 states that the probability of a successful mate is equal to the product of the probabilities that each error is within its
Figure 6.1 Independent normal probability densities for lateral and angular errors.

respective limits. These last two numbers are easily found using cumulative normal probability data from standard tables. The 1, 2, and 3σ probability curves of the joint density could still be plotted on overlay of the acceptable zone for a rapid visual evaluation of the situation, as shows figure 6.2. For this case study, we will consider no slopes on the top and bottom edges of the parallelogram: The resulting rectangle is very close to the actual zone.

Two details remain to fully describe our method of assembly success evaluation. First, we want ε₀ and θ₀ to be the errors between the peg, usually supported by the manipulator, and the hole, usually at the end of a sub-assembly supported by a fixture on a work table. To obtain this, we must propagate all the errors through a sequence of matrix transforms that includes not only the sub-assembly, but the assembly system as well. This is shown schematically in figure 6.3.
Second, the standard transforms and associated errors we have suggested in Chapters 4 and 5 typically result in linear and angular errors at the middle of the axis of circular features. This is clearly shown in Figure 4.2, where we see the final Cartesian frame $F_2$ attached to the middle of the axis of the circular feature, describing the case of a cylindrical tolerance zone. This simply implies that a homogeneous matrix transform should be added at both ends of the complete transform sequence so that the propagated errors match $\varepsilon_0$ and $\theta_0$, as shown in Figure 6.4. In this figure, we see that $\varepsilon_0$ represents the statistical lateral error, with 0 means, between origins of the Cartesian frames $F_{i+1}$ and $F_{j+1}$, measured along the nominal direction of the Y axis; while $\theta_0$ represents the statistical angular error, with zero means, between the same frames, which are nominally at $180^\circ$ of each other on Figure 6.4.
Figure 6.3 Subassembly and assembly system described by a unique matrix transform sequence (adapted from [Nevins & Whitney-'89]).

6.2 The impeller assembly

In section 2.2, we have introduced the reader to assembly sequence analysis (ASA). In doing so, we have also shown the implication of geometric variations on the choice of assembly sequences. The simple impeller example served as a basis to describe these important concepts. In this section, we use this example again as a case study to help the reader understand the use of homogeneous matrix transform representation of tolerances, and to use the concepts of part mating theory on an actual assembly. We will try to establish a relation between the specified component tolerances and the
probability of success of a critical assembly step involving a peg and a hole. The knowledge of this probability will help making a decision on which sequence and fixturing scheme to use.

![Diagram of assembly process]

*Figure 6.4 Statistical linear and angular errors between frames.*

### 6.2.1 Sequence with no refixturing

Figure 6.5 a) shows the same assembly graph as in figure 2.2 c). There are 5 possible assembly sequences for this simple four component product. The design is symmetric in the way that parts A and B (retainer rings) are identical and mated to similar surfaces on part C (shaft), therefore some of the possible sequences are practically identical. We therefore reduce the assembly graph to the one in b), where
we have chosen to eliminate the retainer B as a possible component for initial mate. This leaves 3 possible sequences. Starting the assembly with parts C or D will automatically imply a refixturing (to reorient the subassembly), which we will consider only as a last resort. Liaisons 4 and 5 are therefore eliminated as possible first mates, resulting in the single-sequence assembly graph in c).

This "refixturingless" sequence can be evaluated to see if all assembly steps are possible with the specified tolerances and assembly system variations. We can intuitively identify a critical assembly step in this sequence: it's the mating of the impeller D and the shaft C, as described in section 2.2. We will focus our analysis on this step.
Propagation of errors

Figure 6.6 shows the dimensioned and tolerated components separately. Figure 6.7 shows the retainer and shaft subassembly and the impeller as it is approached to the shaft for mating. We have identified the Cartesian frames between which matrix transforms, with or without associated errors, are to be specified as input data to Tola, the name given to the software implementing a three-dimensional, closed-loop tolerance analysis algorithm. As explained in section 6.1, we create a sequence of matrix transforms that includes the assembly system up to the part held by the gripper. For this example, we assume a perfect manipulator and a known relative position between the manipulator and the fixture holding the subassembly. This implies that we can describe the nominal position and angle between the frame attached to the fixture $F_3$ and the one attached to the gripped surface $F_2$, by a single matrix transform with no associated errors. Setting a convention, we start the sequence of transforms at the mating point of the part manipulated by the assembly system, and it finishes at the mating point of the feature to mate on the subassembly sitting on the worktable.

We start with the Cartesian frame $F_0$, attached to the bottom extremity of the axis of the actual impeller hole, from which we specify a matrix transform to go to the Cartesian frame $F_1$, attached to the middle of the same axis. This transform consists only of a translation along $Z$ of $0.905/2 = 0.4525$, the nominal value of half the length of the hole. The associated error is $+/- (0.91 - 0.90)/2 = 0.005$, the variation allowed on the full length.
Figure 6.6 A  Impeller specifications.

Figure 6.6 B  Retainer specifications.
Figure 6.6 C  Shaft specifications.

From frame $F_1$ to frame $F_2$, attached to the axis of the gripped surface denoted a datum A in figure 6.6 A, we use the runout specification of the hole surface. Since the hole (interior surface of the impeller) is specified with respect to the exterior surface (datum A), we consider that their relative runout specification will still be respected if we apply the specification to the exterior surface while using the hole as datum. Therefore, between $F_1$ and $F_2$, we specify an identity matrix transform since they should nominally coincide, and we represent the runout tolerance using equation 4.14.
The errors associated with the identity matrix transform become

\[
\Delta X = .92 \times .001 / 2 \quad \Delta \theta_x = .92 \times .001 / .905
\]
\[
\Delta Y = .92 \times .001 / 2 \quad \Delta \theta_y = .92 \times .001 / .905
\]
\[
\Delta Z = 0 \quad \Delta \theta_z = 0
\]

We have assumed the assembly system (work table, manipulator, programming, etc.) perfect for this case study, so the nominal transformation with no associated errors from \( F_2 \) to \( F_3 \), attached to the middle of the axis of the interior surface of the fixture, is just a translation of along \( Z \) of \(-(.905/2+.895+0.13/2) = -1.4125 \) so that the mating extremities of the hole and peg (\( F_0 \) and \( F_9 \)) are at the same \( Z \) level.

These three added terms are the nominal values of respectively half the impeller length, the full length of the largest diameter shaft portion, and half the length of the threaded portion of the shaft mating with the retainer ring.

The frame \( F_3 \), attached to the middle of the axis of the interior surface of the fixture, is the starting point of the sequence of transforms that identify the errors through the subassembly. That circular surface has a known diameter and we consider its axis perpendicular to the bottom surface of the fixture, the surface on which the retainer lies. These assumptions are justified, for now, by the fact that the fixture position and size can often be measured and accounted for in the manipulator programming. The retainer ring has been positioned in the fixture with success and we now consider its mate variations along the \( X \) and \( Y \) axes: although the ring and fixture look like a peg-hole mate as discussed in chapter 5, the ratio height/diameter of the retainer is small enough that its back surface is in stable contact with the bottom surface of the fixture. This implies that, when using equation 5.15, describing a peg-
hole mate variation, no angular errors are considered for the axis of the retainer's exterior surface, on which $F_4$ is attached. Furthermore, in equation 5.2, the value of half the clearance $X_R$ (used to compute the conditional variance $\sigma_{cp}^2$ and finally $T_{cp}$) inherits only the variance of the retainer diameter since the fixture diameter is known and fixed.

![Diagram of full assembly and fixture specifications]

Figure 6.7 Sequence with no refixturing: the peg-hole mate.
The nominal values to go from $F_3$ to $F_4$ are null, resulting in an identity matrix. For the associated errors, the value of $T_{cp}$ is computed with

$$T_{cp} = 6\sigma_{cp} = 6 \times \frac{1}{2} \sqrt{\frac{\sigma_{x_n}^2}{4} + \left(\frac{C}{2}\right)^2}$$

which is a combination of equations 5.3, 5.12, and 5.14, and where $\sigma_{x_n}^2$ is the variance of the retainer diameter, and $C$ is the nominal clearance computed by subtracting the expectation of the retainer diameter from the fixed diameter of the fixture. The corresponding numerical expression for $T_{cp}$, the diameter of the variation zone for the retainer axis is

$$T_{cp} = 3 \sqrt{\frac{0.0033^2}{4} + \left(\frac{0.03}{2}\right)^2} = 0.0453$$

where $\sigma_{x_n} = .0033$ has been computed with $(.73-.71)/6$, the range of the retainer diameter divided by 6 since we consider it the $6\sigma$ range. The full set of errors associated with the identity transform is

$$E_{3,4} : \quad \Delta X = .92 \times .0453 / 2 \quad \Delta \theta_x = 0$$
$$\Delta Y = .92 \times .0453 / 2 \quad \Delta \theta_y = 0$$
$$\Delta Z = 0 \quad \Delta \theta_z = 0$$

computed with equation 5.15.
We now define the transformation between $F_4$ and $F_5$, which is attached to the axis of the threaded retainer hole. This hole is specified as a datum for the outside surface, but we can invert the specifications between these two features without modifying their relative position: the outside surface is used as the datum to the hole, and the outside surface position specification is used by the hole. Since this specification is using a modifier, we use equation 4.24 where we compute $T_{cp}$ using equation 4.22 and 4.23

$$T_{cp} = 6\sigma_{cp} = \sqrt{\sigma_s^2 + \sigma_p^2 + [T_p + \text{size}_{\text{max}} - E(\text{size})]^2}$$

where the subscript $s$ or the word size refers to the hole size, while the subscript $p$ refers to the position tolerance zone now associated with the hole. The numerical values are

$$T_{cp} = \sqrt{(1E-3)^2 + 0.0017^2 + [0.01 + 0.203 - 0.200]^2} = 0.013$$

where $\sigma_s = 1E-3$ is computed with $(.213-.208)/6$, the sixth of the minor pitch diameter range; $\sigma_p = .0017$ is the sixth if the position tolerance zone; and where the word size refers to the minor pitch diameter. The set of errors associated with the identity matrix transform is

$$\Delta X = .92 \times .013 / 2 \quad \Delta \theta_x = 0$$
$$E_{4,5} : \quad \Delta Y = .92 \times .013 / 2 \quad \Delta \theta_y = 0$$
$$\Delta Z = 0 \quad \Delta \theta_z = 0$$
from equation 4.24, where here, the angular errors have been eliminated because the 
hole orientation is considered as a secondary contact surface with the shaft: the latter's 
orientation is probably not affected by the hole orientation. It is the top surface of the 
retainer that is the primary contact surface with the shaft and we consider the resulting 
variations next.

We first specify a transform from \( F_5 \) to \( F_6 \), attached to the middle of the top 
surface tolerance zone, with only a translation in \( Z \) of half the thickness of the retainer 
\((.14/2 = .07)\). We do not specify an associated error to this translation since the next 
transformation represents a planar size type tolerance and includes the error.

To go from the frame \( F_6 \) to the frame \( F_7 \), attached to the actual surface of the 
retainer, we use the error format associated with a planar size specification, given by 
equation 4.15. We specify an identity matrix with the following associated errors:

\[
\begin{align*}
\Delta X &= 0 \\
\Delta Y &= 0 \\
\Delta Z &= .95 \times .01
\end{align*}
\]

\[
\begin{align*}
\Delta \theta_x &= .95 \times 2 \times .01 / .72 \\
\Delta \theta_y &= .95 \times 2 \times .01 / .72 \\
\Delta \theta_z &= 0
\end{align*}
\]

The face of the shaft (datum \( F \)) is in contact with the retainer's surface: we 
assume that both surfaces are coincident and that the shaft orientation is affected mostly 
by this relation. Hence, we now go from the shaft face to the datum axis \( C \) using the 
perpendicularity specification, which we use as if it was associated to the axis of datum 
\( C \). The related datum then becomes the shaft face.
We need to describe a tolerance that specifies the perpendicularly of an axis with respect to a surface. The resulting tolerance zone is a (reversed, in our case) cone, the point of it being coincident with the origin of frame \( F_7 \), and its base coincident with the top face of the shaft. We have not developed the matrix transform representation of this tolerance zone but we can easily adapt equation 4.12, the set of errors associated with a position cylindrical tolerance zone. All we do is put the zone reference frame at the bottom of the zone instead of at the middle of it, eliminate linear variations, and specify angular variations that the actual axis will use up the full diameter of the zone at the other end of it. The result of such variation specifications is a conical tolerance zone, as shows figure 6.8. The resulting set of errors in +/- format associated with an identity matrix transform between the nominal reference frame \( F_0 \), with its Z axis along the cone axis, and the frame attached to the actual feature axis \( F_1 \), becomes

\[
\begin{align*}
E_{F_0F_1} : & \quad \Delta X = 0 \quad \Delta \theta_x = f_{opt} \times T_{pe} / 2L \\
& \quad \Delta Y = 0 \quad \Delta \theta_y = f_{opt} \times T_{pe} / 2L \\
& \quad \Delta Z = 0 \quad \Delta \theta_z = 0
\end{align*}
\]

where \( T_{pe} \) is the perpendicularly tolerance specification, which is the diameter of the cone base when the specification applies to an axis, where \( L \) is the height of the conical tolerance zone, and \( f_{opt} \) is the optimizing factor for a cylindrical zone computed in section 4.1.1. We note that this set of errors are used when linear errors (position of the axis) have been considered through other transformations.

The above set of errors represents the variations allowed on the 6 degrees of freedom of an axis which has been specified using a perpendicularly tolerance specification with respect to a datum surface. Referring now to our impeller example,
Figure 6.8 Perpendicularity of an axis to a surface: a conical tolerance zone.

we recall that we have proposed to switch the perpendicularity specification from the shaft face, a surface, to the axis of datum surface C. This implies that the values of both $T_{pe}$ and $L$ are related to the surface, not the axis. Therefore, the errors associated with the identity transform to go from frame $F_7$, attached to shaft face (coincident with the retainer surface), and frame $F_8$, attached to the actual axis of datum surface C, are

$$E_{7,8}: \begin{align*}
\Delta X &= 0 \\
\Delta Y &= 0 \\
\Delta Z &= 0 \\
\Delta \theta_x &= .92x.0001/(2x.375) \\
\Delta \theta_y &= .92x.0001/(2x.375) \\
\Delta \theta_z &= 0
\end{align*}$$

We note that $T_{pe} = .0001$ is the height of the planar tolerance zone resulting from a perpendicularity specification on a flat surface, as opposed to a diameter if it would have been initially associated with an axis. $L = .375$ is the expectation of the shaft diameter, $(.37+.38)/2$. 
Finally, we go from frame $F_8$ to $F_9$, attached to the top of the actual axis of the hole-mating portion of the shaft, with a translation along $Z$ of $(.88+.91)/2 = .895$ and an associated error of $+/-(.91-.88)/2 = .015$. This linear transformation is in effect a powerful lever for the angular variations that we have specified in the previous transformations.

There is an important question about how various shaft features are tolerated with respect to the datum surface $C$, the surface of concern for the critical mate we are analyzing. For example, two matrix transform sequences are possible to relate the threaded feature of largest diameter and the datum surface $C$: i) from the threaded surface to the datum surface $F$ through a perpendicularity tolerance, and then to datum surface $C$ through another perpendicularity tolerance; or ii) from the threaded surface directly to datum surface $C$ through a runout tolerance.

If our representation of tolerances with homogeneous matrix transforms was exact, this component could be considered as "over-toleranced", i.e. there would be confusion as far as which tolerance specification has precedence. But since, for example, since our representation of the runout tolerance only extracts position and orientation variations, even if runout also includes form and concentricity, the presence of these apparently redundant specifications might be justified. When considering the shaft alone, we could consider the transform sequence representing the tightest tolerances after trying both. However, using the largest tolerance allocation would lead to more conservative results in the total analysis.
But in this case study of an assembly process, we cannot consider the components individually; we have to evaluate the mating relationships as well and choose the appropriate sequence of matrix transforms according to the complete picture. This is what we have done between frame $F_5$, attached to the axis of the threaded retainer hole, and $F_9$, attached to the shaft datum surface $C$: we have considered variations of the threaded mate as of secondary importance in comparison to the flat surface mate from the retainer and the shaft. This evaluation has imposed a specific matrix transform sequence in our tolerance analysis, i.e. the one that goes from the retainer hole, to the retainer surface, to the shaft face with no variations, finally to the shaft datum surface $C$ with a perpendicularity tolerance. Note that this sequence is not even part of the choices described above, if the shaft would have been considered individually, and if the threaded shaft-retainer mate would have been considered as the primary mate.

There is a total of 10 frames between which we have defined 9 homogeneous matrix transformations and associated errors. The sequence of transforms goes from $F_0$ to $F_9$ by arbitrary convention (it could be the other way around). These values are given as input to Tola, the software, as input data (radians are changed to degrees first). The result of the closed-formed, three-dimensional tolerance analysis, the output of Tola, gives the standard variation of the statistical errors for the six degrees of freedom: the linear errors along the $X$, $Y$, and $Z$ axes of $F_0$, and the angular errors about the same axes. They are

\[
\begin{align*}
\sigma_{\Delta x} &= .0107 & \sigma_{\Delta \theta x} &= .0088 \\
E_{0,9} : \quad \sigma_{\Delta y} &= .0107 & \sigma_{\Delta \theta y} &= .0088 \\
& \sigma_{\Delta z} = .0061 & \sigma_{\Delta \theta z} = 0
\end{align*}
\]
These errors are assumed normally distributed, and when taken three by three, their joint probability density is an ellipsoid, centered at the nominal position of $F_0$, and whose outside surface can be considered as the $6\sigma$ limit. Figure 6.9 shows the ellipsoids for the linear and angular triplets. We note that the angular error about $Z$ is null, which results in a two-dimensional ellipse (here a particular case of the ellipse where the principal axes are equal since the two non-null standard deviations are equal, resulting in a circle).

![Figure 6.9](image)

**Figure 6.9** Normal density ellipsoids for linear and angular errors of critical peg-hole mate.

Consolidation of propagated errors with part mating theory

To find $\varepsilon_0$ and $\theta_0$ of the fictitious "worst case plane", we look for the maximum pertinent errors. We take the maximum linear error between the ones along $X$ and $Y$
since the one along Z is not a critical factor for the success of the assembly step. This maximum error becomes the lateral error $\varepsilon_0$. For the angular error, we take the maximum between the ones about X and Y, again because the one around Z is not critical. For this initial evaluation, this becomes the angular error $\theta_0$ measured in the same plane as the lateral error: the worst case plane.

Now we find the acceptable zone in that worst case plane, inside which a combination of $\varepsilon_0$ and $\theta_0$ is small enough for assembly success. Along the angular error axis, we define the limits with

$$\frac{\varepsilon_0}{\mu} = \frac{(D - d)D}{D} = \frac{(0.385 - 0.375)/0.385}{0.40} = 0.0649$$

where $\varepsilon$ is clearance ratio, $D$ and $d$ are the diameters of respectively the hole (of the impeller) and the peg (shaft), and $\mu$ is an average coefficient of friction for metal-metal contacts. Along the lateral error axis, we define the limits with

$$\pm [W + C] = 2 \times 0.006 + 0.01 = 0.022$$

where $W$ is the sum of the chamfers, here assumed to be equivalent to about 75% of the nominal radial specifications (75% x 0.008 for both the peg and hole), and $C$ is the nominal clearance between the peg and hole. We see that since the diameters of the features and the size of the chamfers can vary, these acceptable zone limits are really stochastic; the zone should be represented by a set of boundary curves defined by the means and standard deviations of $\varepsilon$, $W$, and $C$. In this example however, we will consider them fixed at their nominal values.
Figure 6.10 visually shows the results of the analysis of the critical peg-hole mate assembly in the impeller example. The probability that both the lateral and the angular errors be, independently, within their acceptable limits is computed with equation 6.1 and is equal to 95.96%.

Figure 6.10 Results of peg-hole mate for sequence with no refixturing.

6.2.2 Alternate sequence: using a refixturing

Figure 6.5 b) shows three possible sequences, one of which we have just evaluated. The other two, starting the assembly with parts C or D, involve a refixturing to reorient the subassembly. These sequences could now be evaluated to see if the critical peg-hole mate step could be eased. However, we can find more possible sequences if we differentiate the various surfaces of each components as potential
fixturing surfaces. This means that the sequence graph can be expanded to include the sequences that are different only on the basis of which fixturing surfaces they utilize. [Abell-'89] has addressed that issue and has implemented the capability to do an expanded sequence graph which includes various fixturing and orientation possibilities for each mating sequences. The consideration of tolerances in ASA is closely related to the use of this expanded graph. In this example, we will assume that the expanded graph presents, among others, the sequence proposed in figure 2.4 c) and d). Again, we analyze the peg-hole mate to evaluate the probability of success.

Figure 6.11 shows the situation once the refixturing has been done. Starting with the assembly system and gripped impeller, we use the same transforms as in the first assembly sequence: this defines the transformations between frames $F_0$, $F_1$, and $F_2$ as we did in the precedent analysis, except that the nominal translation along $Z$ from $F_2$ to $F_3$ is $0.905/2 + 0.895 + 0.13 + 0.2625/2 = 1.4775$. These four terms are the nominal values of the length of respectively half the impeller, the mating portion of the shaft, the threaded shaft portion of big diameter, and half the threaded shaft portion of small diameter (not explicitly specified).

We consider as variation-less the mating relationship between the fixture's threaded hole and the small threaded shaft portion. This implies that the $Z$ axis of frame $F_3$ is coincident with the axes of both the fixture's hole and its shaft mating feature.

The next variation is specified by inverting toleranced feature and datum feature as done before: the runout tolerance associated with the threaded feature is now used for surface $C$, and the related datum becomes the threaded feature's axis. But before, we use
a transform with only a translation along Z of .2625/2+.13+.895/2 = .709 to get to the middle of the axis of the runout tolerance zone, on which frame F_4 is attached. The three terms in the Z translation summation are the nominal lengths of respectively half the threaded shaft portion of small diameter, the threaded shaft portion of big diameter, and the half the mating portion of the shaft. The allowed error along Z is not clearly given between the two frames. Since we already include a tolerance between frame F_5 and F_6, we will keep this one to zero. Note that this would have to be clarified if the Z direction error was more important than in this case study. This runout tolerance zone has been identified as cylindrical and can be represented by an identity matrix transform to the frame F_5, with associated errors defined by equation 4.14:

\[ E_{4,5} : \begin{align*}
\Delta X &= .92 \times .02 / 2 \quad \Delta \theta x = .92 \times .02 / .895 \\
\Delta Y &= .92 \times .02 / 2 \quad \Delta \theta y = .92 \times .02 / .895 \\
\Delta Z &= 0 \quad \Delta \theta z = 0
\end{align*} \]

A translation along Z of .895/2 = .4475 between F_5 and F_6 is used to go at the top of the shaft created by surface C. The associated error is specified as \( \pm \left( .91 .88 \right) / 2 = .015 \) since it is the maximum variation allowed for the length of the surface C.

There is a total of 8 frames between which we have defined 7 homogeneous matrix transformations and associated errors. The sequence of transforms goes from F_0 to F_6. The output of Tola is

\[ E_{0,6} : \begin{align*}
\sigma_{\Delta X} &= .0043 \quad \sigma_{\Delta \theta X} = .0068 \\
\sigma_{\Delta Y} &= .0043 \quad \sigma_{\Delta \theta Y} = .0068 \\
\sigma_{\Delta Z} &= .0033 \quad \sigma_{\Delta \theta Z} = 0
\end{align*} \]
The resulting worst case plane curves, using the same approach as for the previous sequence case, are shown on figure 6.12. The probability that both the lateral and the angular errors be, independently, within their acceptable limits is computed with equation 6.1 and is equal to 100%.

Figure 6.11 Sequence with refixturing: the peg-hole mate.
Figure 6.12 Results of peg-hole mate for sequence with a refixturing.

6.2.3 Choice of best sequence

The sequence involving a refixturing is clearly a better choice as far as probability of success of the critical peg-hole mate is concerned. However, the propagated errors in both sequences, when compared to the acceptable region defined by the physics of part mating, indicate that none of the sequence can be accepted without further investigation.

The factors to be further investigated include most importantly the variability on the width of the chamfers and clearance, affecting the size of the acceptable region in the worst case plane. In actuality, one can evaluate the standard deviations of such variations and plot the worst plane acceptable region using three squares, representing
respectively the 1, 2, and 3σ regions, similarly to the actual resulting ellipses. Such an analysis can help define appropriate specifications for the chamfer dimensions.

Another point to take into account is the relativity of the values of probability of success that were computed with equation 6.1. This equation assumes that the maximum linear and angular errors are independent, which may not be the case. Therefore, the numbers should be used with care, probably only as a relative appreciation of each sequences to compare.

Also, recall that the worst case plane is just that: a worst case. One should plot the resulting ellipses and acceptable squares in each direction to gain better understanding of the assembly step.

Finally, the fixtures whose dimensions we assumed perfectly known might impose a costly manufacturing approach. Increasing flexibility is required to respond to a more competitive market: just in time, production lots of 1, reduced inventory, etc. These concepts often require rapid changing of fixtures and tooling on a given assembly system; these rapid changes can require the allocation of a (controlled) variation on fixture position and orientation.

After these clarifications, a designer is better positioned to modify the specifications to respect the assembly process, or a process engineer is better positioned to choose the sequence that will not only maximize the likelihood of successful mating, but minimize the assembly cost. For example, the refixturing in the alternate sequence might be more expensive than the respect of a tighter specification on the chamfers.
Conclusion

We have shown how standard geometric D&T specifications could be represented in matrix transform, leading to a complete evaluation of assembly sequences through closed-loop tolerance analysis and part mating theory. Out of two assembly sequences, we were able to identify the one that would reduce the chances of unsuccessful automated assembly for a specific and critical peg-hole mate.

We have seen, however, that the identification of the matrix transform sequence (tolerance chain) is a tedious process that often required the judgement of a designer. A basic set of standard representations was derived in the previous chapters, but the presence of many types of situations in parts and assemblies would justify the derivation of an extended set.
Chapter 7

Conclusion
7.1 Recommendations for future work

Develop more matrix transforms

In chapter 4, we developed matrix transform representations for eight standard tolerance specifications. These transforms can be regrouped in four type families: location, runout, planar size, and orientation. Although all the basic specification relevant to assembly evaluation have been covered, we were able to see, through the impeller example in chapter 6, that particular cases of tolerances needed matrix transforms to be adapted to respect the designer’s intent. For example, we have defined explicitly the representation for the perpendicularity of a plane with respect to another one, but we had to adapt it to allow perpendicularity of an axis (of a hole) with respect to a plane to be properly represented. The development of an extended set of matrix transforms for more typical component specifications would help in reducing the need for human judgment in the process of representing Y14.5 tolerances in transforms and associated errors.

In chapter 5, we developed matrix transform representations for one important case of mate variations: a circular peg-hole mate. There exist many other mates that are often present in assemblies of individual components: two planar surfaces, square peg-hole mate, multiple surfaces simultaneously, to name a few. Initially, further development of the understanding of the physics and factors involved, and then representation of the possible and relevant variations with homogeneous matrix transforms are needed. This would increase the range of capabilities of full tolerance analysis in assemblies and assembly systems.
Simplify the Identification of tolerance chains

Identification of tolerance chains can be simplified in various ways: use a clear surface nomenclature, do a complete transform representation of each component before tackling the assembly representation, use the already existing assembly structure information, and develop single matrix transforms that replace frequently used sequences of matrix transforms.

In the case study of chapter 6, we have described all relevant tolerances between surfaces in matrix transform format. In doing so, we have repeatedly used the same surfaces for different transform sequences. Since fixtureing schemes can identify different assembly sequences often by just specifying a different fixtureing surface, it becomes important to clearly identify each surfaces of each component. [Abell-'89] has proposed a mating and fixtureing surface designation scheme, shown in figure 7.1 for the impeller assembly. This nomenclature allowed him, among others, to develop the use of an extended sequence list that included sequences different only in fixtureing surfaces used [Abell-'89]. This is increasingly important when assemblies are comprised of numerous and complex parts.

Once surfaces and features are clearly identified, one could establish the complete matrix transform representation of each components taken separately. This description should include all sequences of matrix transforms needed to represent the relationships between all possible and relevant pairs of surfaces and features in the parts. It should also respect the datums specified in the available tolerancing information. One possible approach is to define a principal reference frame for each single part; from this frame,
Figure 7.1 Mating and fixturing surface designation.
one could identify nominal (non statistical) transforms to all datums of the part, thereby creating relationships between the principal reference frame and each feature of surface-related reference frame (datum). Actual features could then be represented with transforms with associated statistical errors. With this approach, relationships between features of different parts could be easily related using only information about the mating relationships in the assembly.

In any model of complete product, there exists a documented assembly structure to describe how the parts fit together in a final assembly. However, the format of this document and of the information it contains can vary widely. Feature-based design, as described in chapter 2, is a way to describe a component and its corresponding assembly by extending the information beyond the classic geometric shape. What ever the method chosen to describe the assembly structure, it is desirable that the format used be compatible with (i.e. useable by) an algorithm that could translate the structure in the homogeneous matrix transform language. The relations between principal reference frames, as described above) of each parts would then be insufficient: tolerance analysis requires information about which features actually mate to assemble the two parts.

Finally, there is an interest to create a library of matrix transforms so that, by invoking a single transform with associated variabilities from this library, one could fully represent frequently used subassemblies such as gear assemblies, actuators of various types, or families of different bar linkages.
Simplify and automate the input of data in computer program

At the present moment, the computer-based implementation of the 3-D, closed-form tolerance analysis algorithm requires the user to input tolerance information already translated in homogeneous matrix transform format. In a first phase, a proposed task is to implement a user-interface that will facilitate the use of the existing tolerance notation, and make the software more compatible with existing standard and practices.

Without the various automated approaches suggested above, a particular assembly sequence now requires the user to enter both nominal values and allowed variations for the complete set of parts and mating relationships. This will constitute the tolerance chain to a particular feature, associated with a given step of a given sequence. It becomes obvious that, unless the task is simplified, there is a considerable amount of time required to enter the information for all possible sequences.

In an effort to reduce both time and inaccuracies when identifying and specifying this tolerance chain, the user-interface should allow a smooth transition of data between the Y14.5 standard and this new representation of tolerances. Each tolerance type has an associated matrix type, or family, and the user-interface should allow automatic construction of these matrices by implementing a “fill the spaces” approach. The information to put in these spaces will be closely related with the data used to describe a typical tolerance specification, using a feature control frame.
Eventually, this exchange of information between product design database and assembly process planning (and tolerance analysis for assembly evaluation) database could be done automatically. This could be done by sharing appropriate data fields in the complete information structure, including both geometric and attribute-oriented information. Interface applications could be written to make the appropriate format translations.

**Fully Incorporate tolerance analysis into ASA**

The recommendations outlined above are aimed at extending present advancements in tolerance considerations to support assembly sequence analysis (ASA), as proposed by figure 2.5. In other words, the goal is to automate rapid tolerance analysis of a many sequences related to the same product, minimize data input requirements, and provide useful information to help designers and process engineers make sound cost and quality driven decisions. Only after significant advancements in this direction can valuable comparison of assembly sequences be easily performed.

The computer-based tool could then help assign tolerance to satisfy the assembly system or inversely, help to elaborate specifications for the system. For example, guidelines for the assignment of tolerances could come from an identification of the specific variations that contribute most to the final feature's assembly difficulty. The closed-loop tolerance analysis would be rapid enough to permit "what-if's" evaluations as tolerances are adjusted.
7.2 Conclusion

After a brief overview of geometric dimensioning and tolerancing, in chapter 3, we have developed a homogeneous matrix transform representation for most standard tolerance specifications, in chapter 4, and for the important peg-hole mate variation, in chapter 5.

In such a representation, tolerance zones are represented with a matrix transformation between a Cartesian frame attached at the nominal position of the feature, which is the middle of the (usually bilaterally specified) tolerance zone, and the Cartesian frame attached to the actual feature, which has a statistical position and orientation error with respect to its nominal position. The nominal transform is therefore the identity matrix, and the standard deviations of variabilities, associated with the 6 degrees of freedom of the actual feature, are used in a 3-D, closed-loop tolerance analysis algorithm, presented in chapter 2. The choice of degrees of freedom and their respective variation range are chosen so to reasonably approximate the tolerance zone intended by the standard specification.

In the case study of chapter 6, we have demonstrated that this approach to tolerance analysis can successfully be used, in conjunction with part mating theory, to evaluate the probability of success of critical assembly steps. Possible assembly sequences can then considered and compared in light of this new information.

Tolerance representation and analysis based on homogeneous matrix transform is a useful and promising area of research leading to the development of integrated tools for
rapid and efficient assembly process evaluation. This representation should not completely replace other means to describe tolerances because it is not compatible with form tolerance definition (which is also the limitation of many other mathematical representations of tolerance). But the approach we propose has been developed for a clear usage: tolerance analysis for assembly evaluation. It is then suggested as a complementing representation of tolerances that can co-exist with other definitions which have been developed for other usage, as for example, quality control activities.
REFERENCES


