Willey, R.R. and Parks, P.E. “Optical Fundamentals”
Optomechanical Engineering Handbook
Ed. Anees Ahmad
Boca Raton: CRC Press LLC, 1999

Optical Fundamentals

Ronald R. Willey and Robert E. Parks

1.1 Introduction

This handbook consists of chapters written by authors of considerable experience in the practical application of **optomechanics** and covers a broad range of related subjects. In this chapter, we introduce general background information, techniques, and concepts which may be useful to the practitioners of optomechanics. These topics include definitions, fundamentals of geometrical optics, drawing standards, tolerancing concepts, and environmental effects.

What Is Optomechanics?

We will draw upon the definition of optomechanical design from Dan Vukobratovich who is a contributing author in this handbook. We will define **optomechanics** as the science, engineering, and/or art of maintaining the proper shapes and positions of the functional elements of an optical system so that the system performance requirements are satisfied.

Vukobratovich also points out in his chapters why optomechanics is different from conventional mechanical engineering in that the emphasis is on strain or deformation rather than stress.

The recent difficulties with the Hubble Space Telescope (HST) have been attributed to an optomechanical error in the relative positioning of two mirror surfaces in the null corrector used to test the primary mirror. Great pains were undoubtedly taken in the design and construction of the HST which was a major achievement in the field of optomechanics. Aside from its well-publicized difficulty, it still is a good example of the application of optomechanics. The shape and the position of the functional elements such as the primary and secondary mirrors must be maintained very precisely in order for the instrument to obtain results not previously achieved. It is hoped that the repairs made to the HST in 1993 will allow most of the original scientific goals to be attained.

Another example of optomechanics at the other extreme are the eyeglasses that many of us wear. The frame must hold the lenses such that their principal points and astigmatic correction are in the right position and orientation with respect to the user's eyes to within appropriate tolerances. The frame must also interface to the user's head in a comfortable and reliable way, and the whole system must perform and survive in an appropriate variety of environments.

Role of Optomechanical Design and Its Significance

The optical design of an optical instrument is often less than half of the design work. The mechanical design of the elements and their support and positioning also are at least as critical. If we look at the typical astronomical telescope consisting of two or three mirrors, there are many more parts in the mechanical structure that position the mirrors and attempt to do so without distorting the functional surfaces. It can be seen that the mechanical design or the **optomechanics** plays a **major** role in any optical instrument development, particularly where the optics and mechanics interface.

It can be seen from the above examples of the HST and eyeglasses that optomechanics is significant to our lives over the whole gamut from the mundane to the sublime. An out-of-tolerance condition in our glasses can cause us headaches and other pains. A similar condition in the HST caused the scientific community emotional and fiscal pain. On the other hand, proper eyeglasses enhance our individual abilities and comfort, and a proper HST can expand our knowledge of the universe. The reader's own imagination can provide a myriad of other examples of the role and significance of optomechanical design. These might include adjustable rearview mirrors in cars or very sophisticated military optical systems.

The optical and mechanical designers of instruments have by far the greatest influence on the ultimate cost and performance of an instrument. All others, including the manufacturing operations, cannot have as much influence as the designers to change the potential satisfaction of the user and profitability of the producer. Therefore, once the instrument's function is satisfied, economics is of great significance in optomechanical design. Figures 1.1 and 1.2 illustrate these effects and show some of the typical steps in the process of developing an optical instrument.

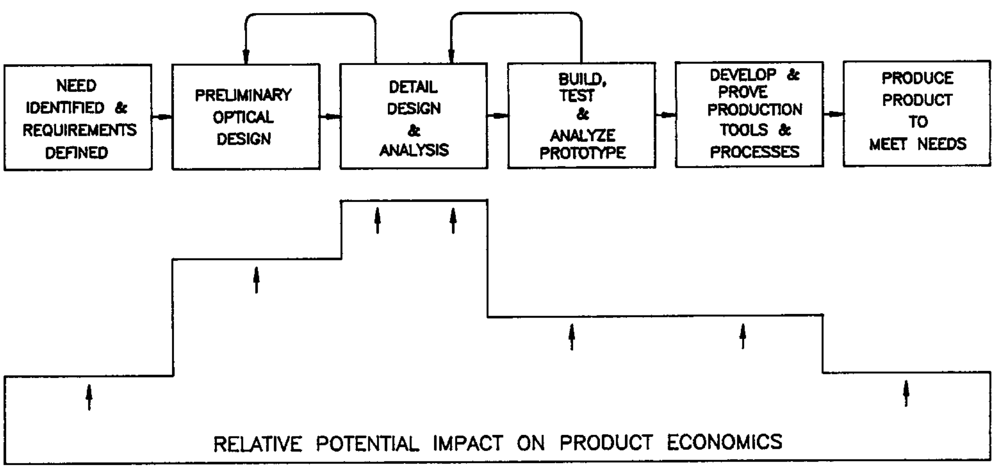

FIGURE 1.1 Overall process to develop and produce a new optical product.

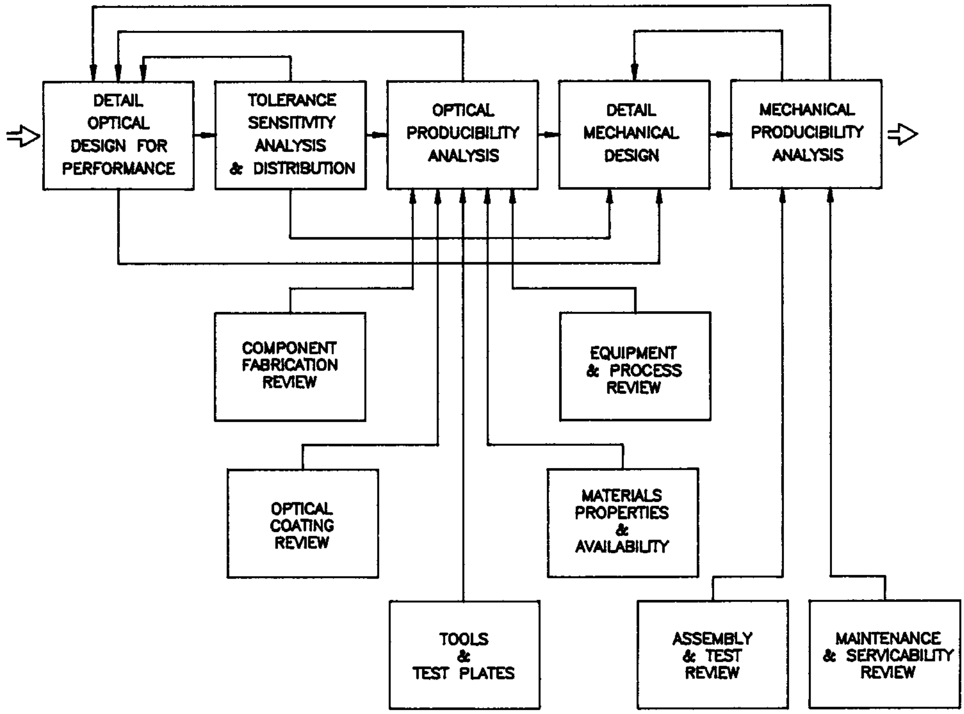

FIGURE 1.2 Detail design and analysis process of an optical product.

1.2 Geometric Optics Fundamentals

Basic Terminology

In this section, we will discuss some of the more pertinent optical terms and definitions which are often needed by the optomechanical design engineer. These terms will be highlighted in **bold** where they first occur in the discussion below. There are many good texts such as Smith's[1] for more detail if needed. We will provide some tools and concepts here which we think are useful for the designer.

Graphical Tools in Geometrical Optics and System Layouts

Image Position and Magnification

In many instances, an assembly of lenses such as a 35-mm camera lens, a telescope, a magnifier, etc. can be treated as a **black box** where only three parameters are known about the lens system. These three parameters are the **effective focal length** (EFL) and the positions in the box of the **front and rear principal points** (P_1 and P_2). Figure 1.3 illustrates such a black box lens system. We will assume throughout this discussion that the optical system is rotationally symmetric about the **optical axis** which passes through the two principal points and that there is air or a vacuum on both sides of the lens system. With this information, we can find the position and size or **magnification** of the image of any object which the lens system can image. This is the first major area of concern in an optical system which the optomechanical designer can easily work out. The second area has to do with how much light can get through the system as a function of the angle relative to the optical axis. We will address these issues subsequently.

Everything that we will deal with here is referred to as **first-order optics.** Departures from the answers which first-order calculations give are **aberrations** or deviations from these answers. These are higher-order effects which lens designers attempt to reduce to practical values in their detail

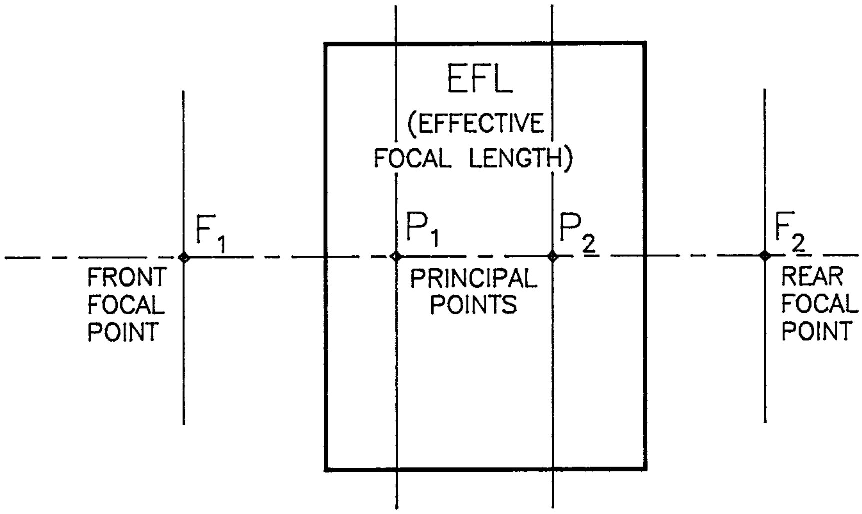

FIGURE 1.3 The three properties which define a "Black Box" lens: effective focal length (EFL) and the two principal points (P_1 and P_2).

design processes. We will not discuss aberrations in detail here since they are not something that the optomechanical design process is expected to improve upon. However, we will briefly introduce the subject. The famous scientist Hamilton viewed optical aberrations as being of three types. The first aberration is composed of the effects which cause the image of a sharp point (stigma) object not to be a sharp point. This he referred to as **astigmatism.** The second Hamiltonian aberration is that the image of a flat plane perpendicular to the optical axis is not on a flat plane but a curved surface. This he called **field curvature.** The third is that the mapping from the object plane to the image plane has **distortion.** This would make a rectangle look like a pincushion or a barrel. Optical designers today tend to divide astigmatism into several parts called **spherical aberration, coma, astigmatism,** and **longitudinal and lateral chromatic aberration.** This latter definition is much less inclusive than Hamilton's astigmatism. These details are mostly of only academic interest to the optomechanical design engineer.

The term "first-order" in optics can be described as simplified equations which are derived by using only the first terms of the series expansions of the sine of an angle and the coordinates of a spherical lens surface as shown in Equations 1 and 2.

$$\sin a = a - a^3/3! + a^5/5! + \dots \tag{1}$$

$$z = y^2/2R + y^4/8R^3 + y^6/16R^5 + \dots \tag{2}$$

In the equation for the sphere, the y is the **zonal radius** from the intersection of the optical axis with the surface of the lens, or its **vertex.** The R is the radius of curvature of the lens surface. The z is the **sagittal** distance or height in the direction of the optical axis from the plane containing the vertex and which is perpendicular to the optical axis. If we use only the first term to the right of the equal sign in each case, we have first-order optics. If we include the next term in each case, we have the basis of third-order optics and the associated aberrations. Seidel worked out the third-order relations whereby all of the above-mentioned aberrations can be calculated. The next terms give fifth-order aberrations etc. With the availability of computing power in modern times, we find only first-order and rigorous calculations to be useful; the higher-order aberrations are of mostly academic value. For our present purposes, no calculations are required, rather only graphical constructions based on first-order principles are needed.

With reference to Figure 1.3, we define a plane containing a principal point and perpendicular to the optical axis as a **principal plane.** The convention is that light passes in a positive direction through a lens if it moves from left to right. The first principal plane is the one which the light reaches first and its principal point is P_1. Similarly, the second principal plane and point P_2 is where light reaches after passing the first. When light is dealt with before it reaches the first principal plane it is said to be in **object space** because the object to be imaged is in that space. After the light has passed through the lens and departed from the second principal plane it would then be in **image space.** The first **focal point** (F_1) is at a distance EFL to the left of P_1 and the second focal point (F_2) is EFL to the right of P_2 on the optical axis. The **focal planes** are planes which contain the focal points and are perpendicular to the optical axis. With only these data, we can construct the position and size of images formed by any centered optical system. Although we will illustrate this with simple examples, almost all image-forming systems can be reduced to an effective focal length and its two principal points and thereby treated by this same technique.

The principles needed for this construction are simple and are illustrated in Figure 1.4. First, when a ray parallel to the optical axis in object space intersects the first principal plane at a given height, it will exit the second principal plane at the same height and pass through the second focal point F_2 in image space. Similarly, a ray parallel to the optical axis in image space must first pass through the first focal point F_1. Second, any ray passing through the first principal point will exit from the second principal point in the same direction (parallel to the first ray). These are really the **nodal points** which correspond to the principal points as long as the lens is bounded on both sides by the same medium (usually air or vacuum).

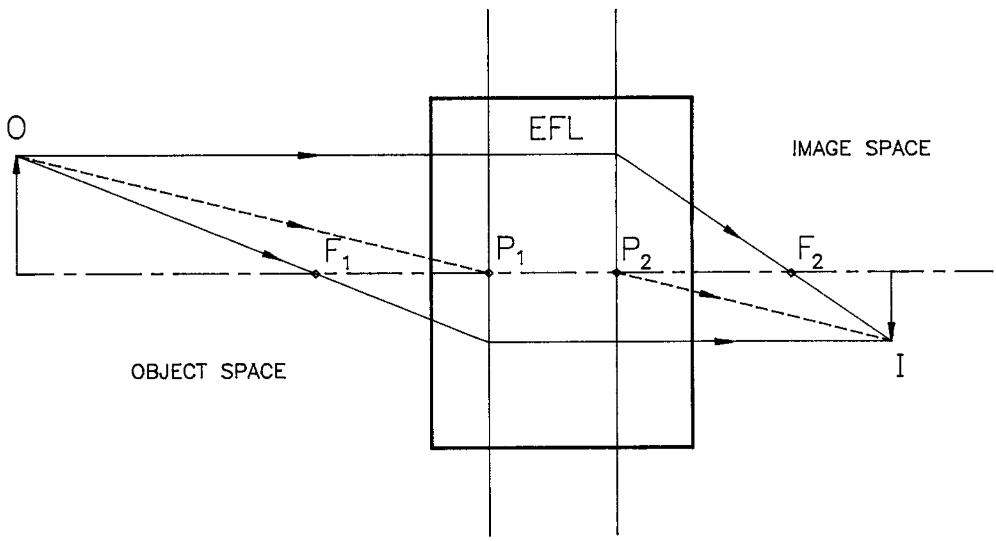

FIGURE 1.4 Construction of an image point from a general object point to find the size and position of an image.

In Figure 1.4, the image point I of an object point O is found by using the above rules on the ray from the point O parallel to the axis and the ray from O through F_1. The rays through the nodal points are also shown (dashed) as an alternative or check. These principles can be applied to any object point. Note that a point on one principal plane is imaged as a unit magnification on the other principal plane.

There are cases where the principal planes are "crossed" such that the first is to the right of the second. In such cases, the rays are still traced from the object point to the first principal plane and emanate from the second at the same height. The principal planes may also be outside of the physical lens system. This is particularly true of some telephoto lenses where, by definition, the lens system is shorter than its effective focal length. We, therefore, have the simple tools we need to find the position and size of the image of any object.

For an afocal telescope as a whole, the above rules have no clear meaning since the focal length is infinite. However, an afocal system can be dealt with when broken into front and rear sections of finite focal length whose focal planes are coincident.

Amount of Light Through a Lens System

The amount of light which can get through a lens system at a given angle from the optical axis is determined by the pupils, apertures, and vignetting. We will neglect the effects of the transmittance of the lenses and reflectance of the mirrors in the system and only address the relative difference of some angle off-axis from the on-axis light. The **entrance pupil** of a lens is the aperture viewed from object space which can pass light to the image space. The **exit pupil** is the aperture viewed in image space which passes light from object space. Both of these pupils are the images of the same **aperture stop** as viewed from object and image space. In a photographic lens, the aperture stop is typically an iris diaphragm which is of adjustable aperture for light brightness control. The **F-number** of a lens is the effective focal length divided by the entrance pupil diameter. The **numerical aperture** is another statement of the same quantity where it is 1/(2*F-number) when objects are at infinity and in air or vacuum. The greater the numerical aperture, the more light will pass through the lens.

We will now use a simple lens to illustrate pupils and the use of the previous principles to find an exit pupil. Figure 1.5 shows a meniscus lens with the aperture stop well in front of the lens and gives the principal and focal points (known as **cardinal points**). In this case, the aperture stop is also the entrance pupil because there are no other lenses in front of it. We want to find by construction the size and position of the exit pupil, which is the image of this stop. We use the point at the top of the stop as an object point to find where it is seen in image space. The ray through that point and the front focal point F_1 is extended until it intersects the first principal plane, and then a line parallel to the optical axis is drawn from that height in the second principal plane and on to the second focal plane. Note that this line extends from minus to plus infinity and we will find that it intersects the next ray on the left of the lens. The next ray is the ray through the stop parallel to the axis in object space which passes through the second focus F_2 in image space. When the image space ray through F_2 is extended backward to the left, it intersects the first ray at the image of the top of the stop in image space. By symmetry, this gives the position and size of the exit pupil. Even though it is to the left of the lens, it is still in image space because it is composed of light that has passed through the lens. Therefore, this exit pupil is larger than the entrance pupil and is farther to the left, but note that the numerical aperture or F-number is still correct.

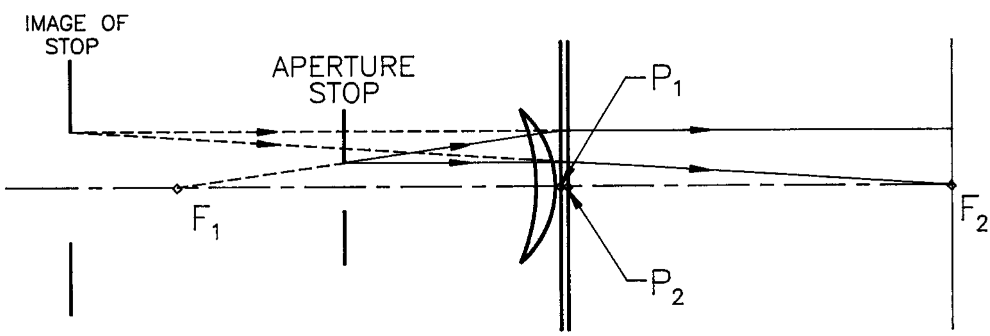

FIGURE 1.5 Finding the exit pupil of a system by constructing the image of the aperture stop.

As seen in Figure 1.6, we now trace rays, which are parallel to a line from the object to the front node (P_1), from the top and bottom of the exit pupil to the first principal plane. Then, from the same height on the second principal plane as those intersections on the first principal plane, we draw rays to the image point (determined by the line from the second node [P_2] to the focal plane which is parallel to the line from the object to the first node). This shows where this off-axis beam will be on the lens. Note that a ray (dashed) through the nodes of the lens to this off-axis image does not pass through the stop; it is **vignetted.** This is not a problem because the nodal ray is just used to find image position from an input angle and vice versa. If we go to a greater angle off-axis as in Figure 1.7, the aperture of the lens itself limits the rays which can pass through it; this is then a **limiting aperture.** The interaction of this limiting aperture and the exit pupil creates a vignetting pattern also shown in Figure 1.7. The only light which can pass is where the two apertures overlap, in this case about 50% of the on-axis value. More complex lenses will typically have front and rear limiting apertures which interact with the exit pupil to give vignetting. When these block all of the light coming to the focal plane, we reach the absolute limit of the **field of view** (FOV). Vignetting is sometimes used to block highly aberrant rays so that the image is sharper even though it is correspondingly dimmer off-axis. In many photographic or television systems, the film or detector (CCD) sizes are usually the limiters of the FOV rather than the vignetting. In visual instruments like binoculars or telescopes, there may be a physical aperture at the final or intermediate focal plane which limits the FOV and is called a **field stop.**

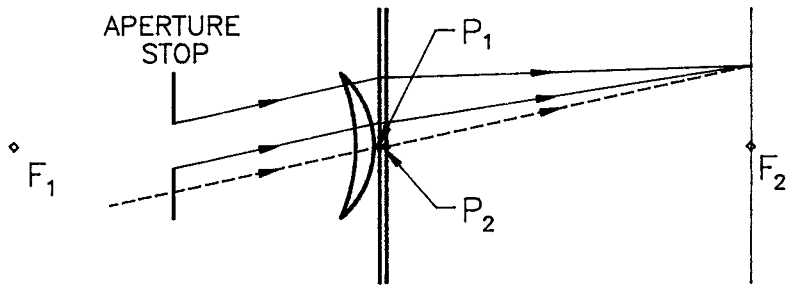

FIGURE 1.6 Construction of the ray paths from an off-axis object through the aperture stop or pupil to the image plane.

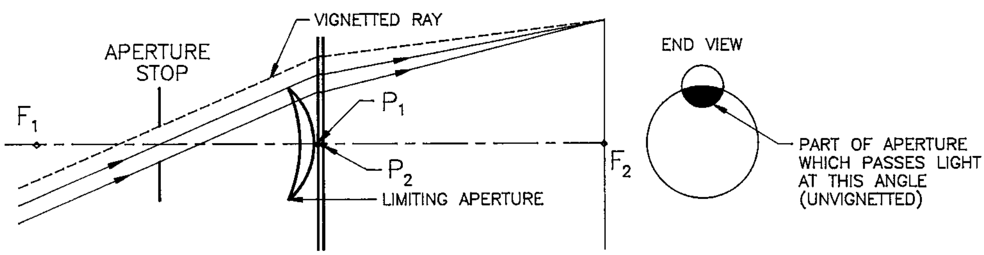

FIGURE 1.7 Off-axis beam which is partially obstructed by a limiting aperture showing its vignetting effect and limiting the ultimate field of view.

The display of the vignetting of the apertures can be easily constructed. In Figure 1.8 we project the limiting aperture onto the plane of the exit pupil. This is done by a first ray from the image point through the center of the limiting aperture to the exit pupil plane. This defines the center of the projected circle. The second ray is traced from the image point through the top or bottom of the limiting aperture to the exit pupil plane. This defines a point on the circumference of the circle whose center was found above. The common overlapping area pattern of this circle and the circle of the exit pupil (or image of the aperture stop) give the vignetting of the system for this image point as illustrated to the left in Figure 1.8. This same result can be obtained by projecting the exit pupil onto the limiting aperture plane. The whole procedure could also have been equally

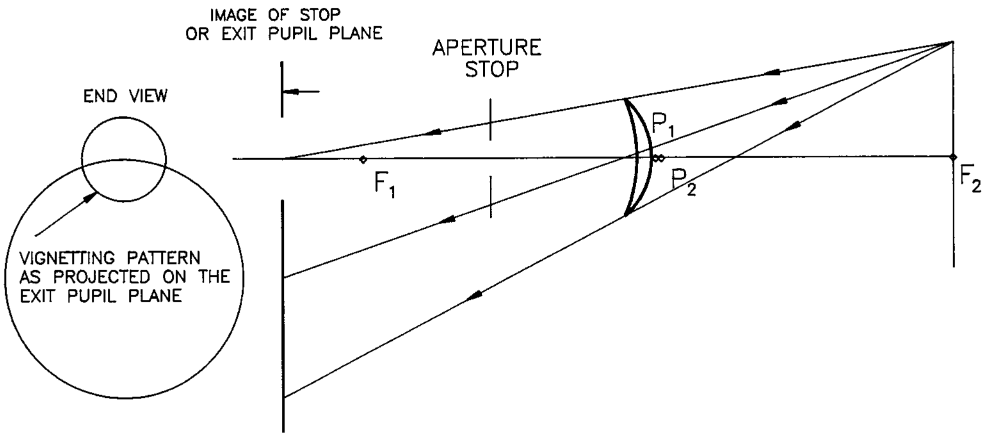

FIGURE 1.8 Construction of the vignetting pattern at a given field angle (image position) in image space.

well applied in object space instead of image space. In this latter case, the vignetted pupil is viewed from the object point.

Additional CAD Techniques for Design and Ray Tracing

It is now practical to trace rays that are rigorously correct using computer-aided design (CAD) tools. This has been true for over four decades with respect to ray tracing and lens design where equations are used to calculate ray paths in three-dimensional space. In fact, optical design may have been the most active CAD process over the majority of that period. However, today's drafting CAD tools allow additional ray tracing to be done with relative ease in a graphical setting. This would not be a replacement for optical design in the usual sense, but it allows the optomechanical designer to find the exact path of a specific ray when needed. The data required is the same prescription which the lens designer produces. This includes: the surface position in three dimensions, its radius, the index of refraction on each side of the surface, and the incoming ray coordinates. The relation which we will execute graphically is **Snell's law** as given in Equation 3.

$$n \sin i = n' \sin r \tag{3}$$

The index of refraction on the side where the ray is incident is **n.** The angle of the incident ray to the surface normal at the point where the ray intersects the surface is **i.** After refraction, the ray is in a medium of index **n′** and makes an angle **r** with the surface normal. The refracted ray lies in the plane which is defined by the incident ray and the surface normal.

The steps to trace the refracted ray as seen in Figure 1.9 are as follows:

Define (draw) the refracting curved surface in a plane which contains its center of curvature and the incident ray.

Draw the incident ray to the point where it intersects the refracting surface and draw the surface normal at that point. This is a line through that point and the center of the circle.

Draw two circles about the intersection point of the incident ray and the refracting surface whose radii are the same multiple of n and n′. It might be convenient to use 1.000″ for an index of 1.000 and 1.517″ for an index of 1.517 or possibly 2× these values.

Through the point where the incident ray intersects the n-circle, draw a line parallel to the surface normal which intersects the n′-circle. This could be a copy command offsetting from the incident ray intercept with the refracting surface to the ray intersection with the n-circle.

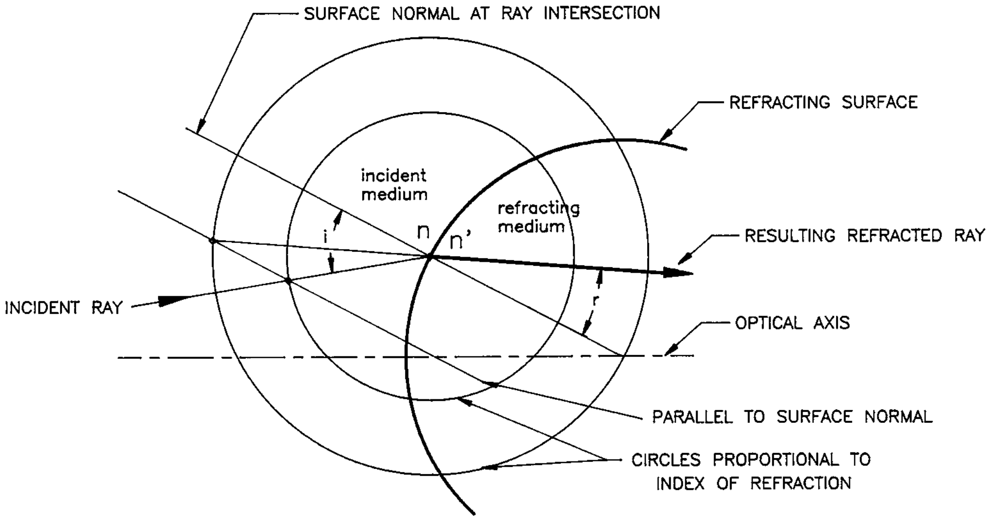

FIGURE 1.9 Rigorous ray tracing by construction of a refracted ray.

From the intersection of this parallel line with the n′-circle, trace a line which passes through the point where the incident ray intersects the refracting surface. This line, when extended beyond that point, is the refracted ray which we wanted to construct.

Trim away any excess lines and the process is done for that refraction of the ray.

It can be shown that this construction satisfies Equation 3. Because of the rigor of modern CAD drawing systems, this should provide real rays that are as accurate as the system can draw.

If exact ray trace is available, the ray can be drawn by connecting intercept points on surfaces. This may be more precise than the above-described graphical means.

If we need to find the path of a reflected ray, the process is even simpler on CAD. The "MIRROR" function of CAD will provide the reflected ray by mirroring the incident ray about the surface normal at the point of intersection.

Rays could be propagated through entire systems this way if needed. The most likely use might be only occasional, such as to check the path of rays near the edge of a lens mounting cell.

All of the tools and concepts described above are general and can be applied with or without CAD, but the availability of CAD has made them that much easier and more useful.

1.3 Drawings of Optical Components and Systems

Units of Measurement

The optical industry has probably been the first in the U.S. to be most familiar with the metric system. Most optical shops work in inches or millimeters with almost equal facility. Some recent U.S. Government military contracts even require the drawings to be in metric units, millimeters, etc., which is a change from the practice of the past. The choice of the units of measurement will probably continue for some time to vary with the customers and the contract specifications. The purpose of the new standards described below is to make specifications more uniform and to reduce ambiguity.

ISO and ANSI Drafting Standards

There are far more international (ISO) and national (ANSI) voluntary mechanical and optical technical drawing standards available than most people are aware of. In this section, we will describe most of these standards, explain why the tendency now is toward using the ISO standards, and where copies of these standards can be obtained.

First a word about why these standards are not better known. At least in the U.S., the majority of optical drawings were either in-house drawings for commercial products made on site and proprietary standards were used, or the drawings were done for military optical systems and were done to military standards. Thus there was little need for optical firms to use voluntary national or international standards. Now, however, there are few commercial optics being made in the U.S. and orders for military optics have decreased dramatically. Instead, commercial systems are being designed in the U.S., Europe, and Japan for manufacture elsewhere and the value of standards that are understood in all parts of the world is increasingly valuable.

Another point to keep in mind about optomechanical drawing standards is that they explain how to indicate on a drawing what features and dimensions for the features are desired on the finished product. The standards, in general, do not tell what values should be used for various desired features. In this sense, this section is merely a lead in to the next section on tolerances, where the actual numerical values to be used on the drawing will be discussed.

It should be pointed out, however, that the ISO standards do offer some guidance on suggested values for certain features, a guidance that is not often found in ANSI standards. For one, virtually every feature that should be considered in the drawing is listed so there is a reminder to think about whether this item needs to be considered or not. In addition, the ISO optical drawing standard contains a section that lists default tolerances. If, for example, chamfers are not called out on a drawing, this section governs the widths of chamfers on the finished components.

Before describing the available standards, we would like to re-emphasize why, when there is a choice of using either ANSI or ISO standards, we would recommend using the ISO ones. More and more, we are dealing in a global economy and the use of methods that make international commerce easier will lead to greater productivity. The ISO standards have been written more recently than most of the ANSI standards, are much more thorough in their treatment of features to be indicated on drawings, and they also deal with a broader range of possible features than the existing national standards.

Another reason to adopt these standards is that they are being built into optical design software. Several of the major lens design code suppliers have added or are adding the ISO indications to their optical drawing software packages. In addition, these same suppliers are working with the electronic data transfer standards people doing IGES/PDES and the more comprehensive STEP work for transfer of CAD/CAM data files. In other words, by working in the ISO environment, drawings will be forward compatible into the newer methods of electronic data transfer.

There is one final reason for using the ISO approach. There has been a concerted effort to make optical drawings virtually noteless, a big change from the past U.S. practice. Most of the ISO indications on drawings use alphanumeric symbols which stand for certain features or parameters of the features. Once the code for the features and parameters is learned, and whatever may be the language of either the designer or the manufacturer, the drawings can be interpreted by workers having almost any language background without the need for translation. An optical drawing created in Japan or Russia using the ISO notation should be readily usable in the U.S., for example, without any need for translation and vice versa. Figure 1.10 is an illustration of the application of the ISO standards to the drawing of a lens. This example was provided by Sinclair Optics and was automatically generated by their optical design software package.

Mechanical Drawing Standards

The principal U.S. national standard covering mechanical drawing practice is ANSI Y14.5M, *Dimensioning and Tolerancing*. This standard explains how to represent on drawings concepts such as maintaining parallelism between two surfaces or that a hole is located a certain distance from a pair of right angle edges. In addition to defining the symbols needed to express these ideas, the standard gives examples of what one should expect in terms of a finished part when a given set of symbols and dimensions are put on a drawing. Two other related standards used in conjunction

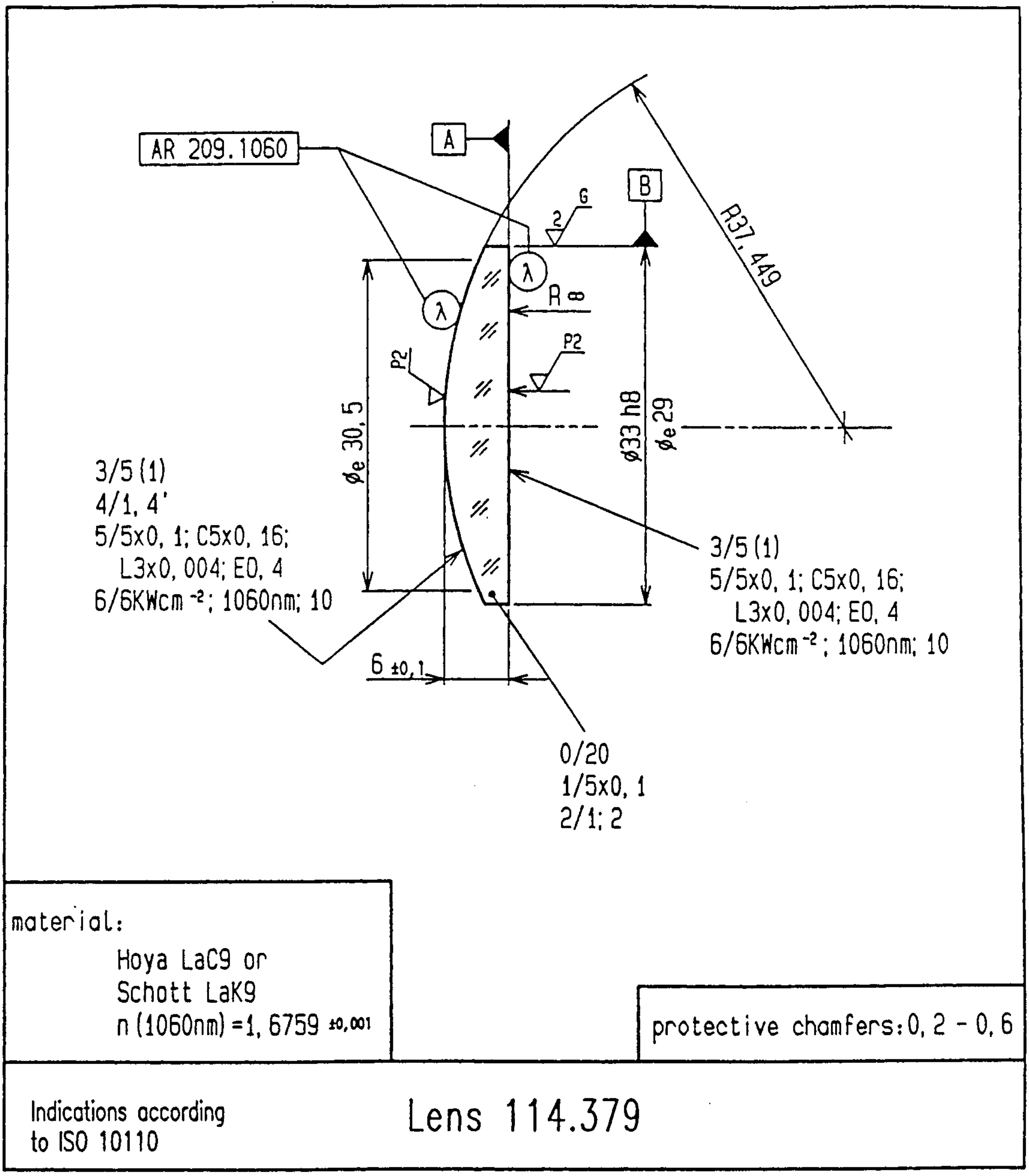

FIGURE 1.10 Example of a lens drawing to the ISO standards. Note the limited notes and dependence on any particular language (i.e., English).

with ANSI Y14.5M are ANSI Y14.36, *Surface Texture Symbols,* and ANSI B46.1, *Surface Texture,* dealing with the finish of machined parts.

We should mention two books that are tutorials on ANSI Y14.5M. These are *Geo-metrics II,*[2] written by the vice chairman of the ANSI Y14.5 committee when the latest version of the national standard was published, and *Design Dimensioning and Tolerancing,*[3] which is written like a textbook and is a good place to start if one knows little about ANSI Y14.5M.

The ISO standard analogous to ANSI Y14.5M is ISO 1101–1983, *Technical Drawings — Geometrical Tolerancing — Tolerancing of Forms, Orientation, Location and Run-Out Generalities, Definitions, Symbols, Indications on Drawings.* In fact, there are so few differences between these two documents that the differences are easily listed in the Preface of ANSI Y14.5 and the members of both committees are working to remove even those.

A companion to ISO 1101 is ISO/TR 5460–1985, *Geometrical Tolerancing — Verification Principles and Methods — Guidelines*. For every feature designation in ISO 1101, ISO/TR 5460 shows in schematic form, kinematically correct methods for verifying that each feature meets the specified criteria. For every kind of feature control call out, this standard shows several examples of how to fixture the part to a surface plate and how to set up an indicator to verify that the feature is in tolerance. The effect is to make clear in a functional way what each feature control call out means in terms of the instruments used to inspect the feature.

The above-mentioned ISO drawing standards are bound together along with 58 other standards related to mechanical drawings and metrology in a 600-plus-page book called *ISO Standards Handbook 33, Applied Metrology — Limits, Fits and Surface Properties*. This, along with an earlier publication, *ISO Standards Handbook 12, Technical Drawings*, is a convenient and relatively inexpensive way of obtaining all the ISO standards on mechanical drawings published through 1988. These handbooks as well as more recent ISO standards are available through ANSI,[4] the U.S. sales agent for ISO.

Before leaving the mechanical drawing standards, we should mention that the official system of units for ISO standards is the SI system (the International System of Units). While ISO mechanical drawing standards allow for the use of English units on drawings if they are identified on the title block, English units are neither extensively used in world trade nor in optics, in general. Thus, it is suggested that the reader have a copy of ISO 1000–1981, *SI Units and Recommendations for the Use of Their Multiples and of Certain Other Units*. This standard is available along with 14 others dealing with definitions of physical quantities and units in *ISO Standards Handbook 2, Units of Measure*.

Optical Drawing Standards

There are a handful of U.S. voluntary standards relating to optical drawings and optomechanics. These are listed in Table 1.1 along with the issuing organization and page length of the standard. The only one of these standards that has had much if any practical impact on domestic optics is ASME Y14.18M-86, *Optical Parts*, and that only in its original incarnation as MIL-STD-34, *Preparation of Drawings for Optical Elements*. The ANSI standards on the list are all the product of the photographic industry and were written about the time when all camera manufacturing ceased in the U.S.

TABLE 1.1

	Ref.
ASME Y14.18M-86 Optical Parts, 37 p	5
ANSI PH3.617-80 Appearance Imperfections, Test for, 22 p	
ASTM F1048-87 Surface Roughness by TIS, 6 p	6
SAE AMS 2521B-89 Antireflection Coating for Glass, 8 p	7
ASTM F1128-88 Abrasion Resistance of Coatings Using Salt Impingement, 3 p	
ASTM D4541-85 Pull Strength Using Adhesion Testers, 7 p	
ASTM F768-82 Measurement of Specular Reflectance and Transmittance	
ANSI PH3.616-90 AR Coatings for Photographic Lenses	
ANSI PH3.713-85 Environmental Testing of Photographic Equipment, 15 p	
ASTM D2851-86 Liquid Optical Adhesive, 4 p	

On the international front, ISO Technical Committee 172, Optics and Optical Instruments, was founded in 1979 and has been writing optical standards ever since. Over 150 standards are either being worked on or are now published. The two standards that are of greatest interest are ISO 10110, *Indications in Optical Drawings*, and ISO 9211, *Optical Coatings*. ISO 10110 is similar to ASME Y14.18M but has 13 parts and is over 100 pages long in draft form. There is no U.S. equivalent of ISO 9211.

We will briefly go over the contents of the two standards as they bear directly on optomechanical drawings. Part 1 covers the mechanical aspects of optical drawings that are specific to optics and not already covered in one of the ISO mechanical drawing standards. Parts 2 to 4 cover material-related parameters such as stress birefringence, bubbles and inclusions, and inhomogeneity. Part 5 concerns figure measurement and differentiates between a figure measured visually with a test plate and that measured with a phase measuring interferometer. Part 6 deals with centering errors and allows either an entirely mechanical method of tolerancing or an optomechanical one.

Part 7 is the equivalent of what we now call the scratch and dig or surface beauty specification and appears to be a more workable method than present schemes. Part 8 concerns ground and polished surface texture and is unique to this standard. Part 9 tells how to indicate that a surface will be coated, but not what the specifications of the coating are. The latter are covered in ISO 9211. Part 10 tells how to describe the parameters of an optical element in tabular form and is the foundation of the effort to be able to transfer data about optical elements electronically.

Part 11 is a table of default tolerances on optical parameters so that if a particular parameter is not specified, it should then be made to the tolerances given in this table. Part 12 defines how to describe an aspheric surface and the method has been coordinated with the major vendors of lens design software so the definitions are consistent. Finally, Part 13 tells how to specify a laser power damage threshold on an optical component, again a parameter that goes far beyond any other existing standard.

The four-part ISO 9211, *Optical Coatings* standard exceeds any existing standard in its thoroughness and detail. Part 1 covers definitions of coating terminology and the definition of ten coating types by function. It also has an extensive table of types of coating imperfections including diagrammatic illustrations.

Part 2 concerns the optical properties of coatings and outlines the properties of the coating that need to be specified to be a complete description. It also shows graphical formats for specifying the transmission or reflection properties of coatings. Several example illustrations of coating specifications are given.

Part 3 is about the environmental durability of coatings as a function of intended use. There are five categories of use ranging from completely sealed within an instrument to surfaces exposed to severe outdoor conditions and unsupervised cleaning. There is a list of 14 different environmental tests for coated surfaces ranging from abrasion to mold growth. Part 4 defines environmental test methods specific for coatings. These are abrasion and solubility tests. The abrasion test includes specifications for the cheesecloth and the eraser for the tests.

It should be mentioned that there are many other ISO optical instrument standards that involve various aspects of optomechanics. For example, there are subcommittees working on standards for binoculars and riflescopes, microscopes, geodetic instruments, medical instruments such as tonometers and endoscopes, and all types of lasers and instruments using lasers. The laser subcommittee has been one of the most active groups and is working on such things as standardizing the diameters and thicknesses of optical elements used in lasers.

Information on these international standards is available from several sources. As has already been indicated, published ISO standards, Handbooks, and catalogs of standards are available from the "Foreign Order Department" at ANSI.[4] For draft versions of ISO standards and lists of drafts being worked on, contact the Optics and Electro-Optics Standards Council (OEOSC),[8] the secretariat for U.S. participation in ISO optical standards writing activities and administrator for the domestic ANSI OP Committee on Optics and Electro-optical instruments. There is a growing list of military, domestic, foreign, and international optical standards on the OEOSC Internet Homepage at http://www.optstd.org as well as calendars of standards meetings and other publically accessable optical standards information. SPIE[9] and OSA[9] have optical standards information on their Homepages as well. Finally, the Optical Society of America has published a handbook[10] designed to be a companion to the ISO 10110 optical drawing standard and an aid to its interpretation.

1.4 Dimensional Tolerances and Error Budgets

The focus of this section is concerned with the practical aspects of tolerancing the designs of optical instruments which are intended for production in large or even small quantities. Certain performance is required of an instrument in all applications. The design and tolerancing aspects of the process have a major effect on the life cycle cost and efficiency of the system. We will discuss what factors make up the cost of a lens and the effects of tolerances and other factors on that cost. This, incidentally, will result in a lens cost estimation formula. We will describe the interactions of lenses and lens cells/mounts from the tolerance viewpoint. We then explain the principles whereby the system tolerances can be determined to minimize the cost of a system, which meets the performance requirements.

The assignment of tolerances to various dimensions and parameters of an optical system is a **critical** element in determining the resulting performance and cost of the system. Much of the tolerancing of systems in practice has been done by art and experience rather than by scientific calculation. Here, we attempt to make the **engineering** principles as simple and clear as possible so that these may be applied in a straightforward manner. We use the term "engineering" to imply that practical approximations based on empirical data are used to reduce the problem to practical terms that can be handled in the real world. In the production of an optical system, random errors in parameters occur. These cause the results to be statistically predictable, but not exactly calculable. Therefore, the use of reasonable engineering approximations is appropriate and justifiable.

It is sometimes possible to tolerance a system such that each of the components is fabricated to an accuracy which will ensure that the instruments will be adequately precise and aligned to give the required performance by simple assembly with no alignment or adjustment. This may be the case with certain diamond-turned (a high accuracy process) optical assemblies of high precision, or with systems having low performance requirements with respect to the process capability. The other extreme is where almost every parameter of a system is loosely toleranced but can be adjusted, with proper skill and labor, to allow the system to deliver the desired performance. However, neither of these approaches is usually the least-cost method to meet the performance requirements. We next discuss philosophical principles and practical ways of approaching the least-cost solutions and give an illustration of the application of the techniques.

In 1982 we started a series of investigations[11-13] of how to achieve the least-cost tolerancing of an optical system. Since that time, Wiese[14] has compiled a very useful collection of papers covering many aspects of tolerancing. This volume includes a paper by Plummer[15] which played a role in our earlier work and a paper by Adams[16] which judiciously utilized some of our findings. Fischer[17] did a more recent survey based on Plummer's work, which we compare with some of our updated cost vs. tolerance data below. Parks[18,19] and Smith[20] have papers in Wiese's collection that have many practical and helpful discussions on the subject. The referenced works form a good general background for this section, but we will reiterate the salient points below for the convenience of the reader. We will also cite other specific references as they apply. It is the authors' experience and opinion that a great deal of resources have been wasted in the past due to poor tolerancing "art". A rigorous and all-encompassing treatment of all but the simplest system can be **very** complex. It is our aim to move the practice of tolerancing from the art stage to the engineering stage with as much simplification as is reasonably justifiable. Warren Smith[20] has made significant contributions to move the status of the practice in this direction. We are attempting to move another step in this direction by adding the real influence of cost into the tolerancing process.

Effect of Tolerances on Cost

Base Costs

We will first review the concepts of base costs from our previous work, and then present the results of later work for the estimation of the base costs directly from the data on a drawing and/or the specifications of a given component.

Let us take the example of fabricating a single lens. For fabricating a biconvex lens of glass, we would typically have to go through the following steps:

- Generate (or mill) a radius on the two sides.
- Mount the lens on a spindle.
- Grind and polish the first side of the lens.
- Remount the other side of the lens on a spindle and grind and polish the second side.
- Edge the lens.

There are obviously a few other minor steps such as obtaining the materials and grinding and polishing tools, dismounting, cleaning, etc. We have said nothing to this point about adding other specifications such as diameter, radii, thickness, and tolerances. Even without these parameters, there is a minimum cost in time, materials, and equipment necessary to make the biconvex glass lens. This is what we define as the base cost. As we get more specific about the lens and add more restrictive tolerances to the parameters, more care, time, and equipment will probably be required to make the lens to the new specifications. Therefore, the cost will increase with increasingly stringent requirements/tolerances/specifications.

Our previous work[11-13] went into some detail on the relationships of the increase above the base cost for the changing tolerance values. However, neither we nor Plummer[15] quantified those base costs. We will quantify the base costs here in order to enhance the usefulness of the technique and the accuracy of the results. It was not made clear in the earlier work that the base costs for a given cost vs. tolerance case are not the same for the grinding and polishing as they are for the centering and edging. As we shall show later, the centering and edging operation is not influenced by the grinding and polishing costs or tolerances and vice versa. Not incorporating this concept can lead to some error in the application of our earlier work and Adams'[16] extension of it. There is a different base for the increase if tolerance is to be applied for the two classes of operation. Figures 1.11 and 1.12 show the effect of tolerances of lens diameter and eccentricity on the costs as a percentage of base cost for the centering and edging operations. This applies to the centering and edging base cost (CE). Figures 1.13 through 1.18 show the tolerance costs as a percentage of the base grinding and polishing costs (GP) for tolerances of radius of curvature, irregularity, diameter-to-thickness ratio, center thickness, scratch and dig, and the glass stain characteristics. It was also interesting to us to recognize that the milling or generating costs are not particularly affected by tolerances with today's equipment and they are not part of the base cost affected by tolerance factors. The milling costs are, therefore, part of the base cost, but only as a function of material and dimensions and not tolerances.

Before we discuss the cost vs. tolerances in detail in the next section, we will show the development of the base cost formula. The present view and cost-estimating scheme are the authors' best effort to date resulting from their experience and discussions based on the work and experience of Stephen Cupka, Manager of Estimating, and Reinhard Seipp, Assistant Manager of Optical Manufacturing at Opto Mechanik (OMI). These people also draw upon their experience from one or more other shops. Let us call the total base cost to make a lens MT, which is in either time or money, which differ only by some multiplicative factor. If we call the milling or generating cost MG, we can represent the total base cost MT as the sum of milling, grind and polish, and centering and edging costs as given in Equation 4.

$$MT = MG + GP + CE \tag{4}$$

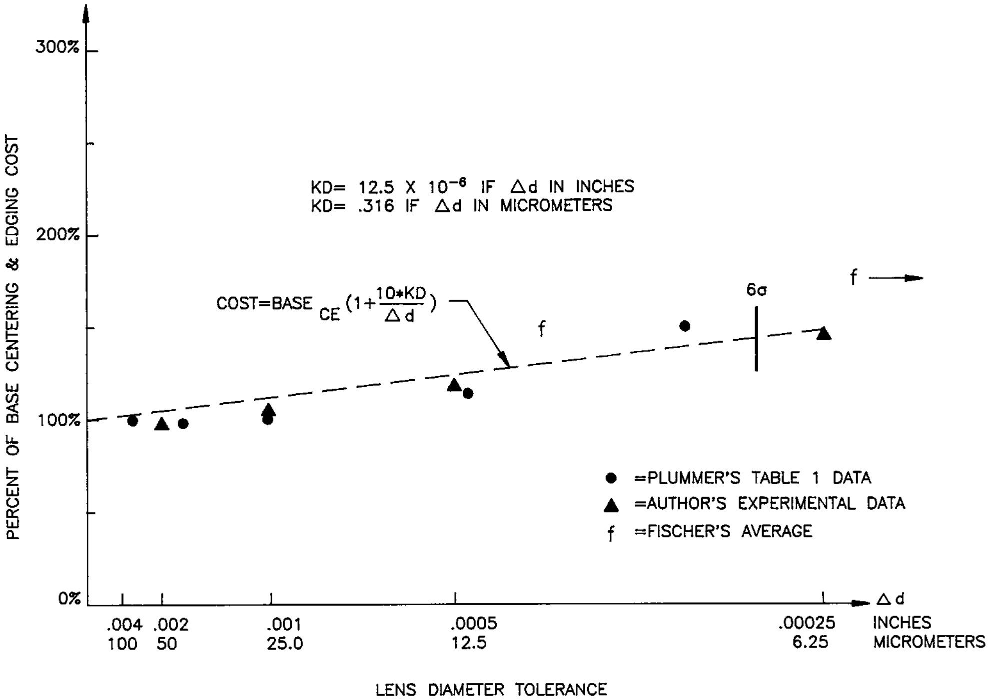

FIGURE 1.11 Relative cost vs. reciprocal tolerance according to various authors concerning lens diameter.

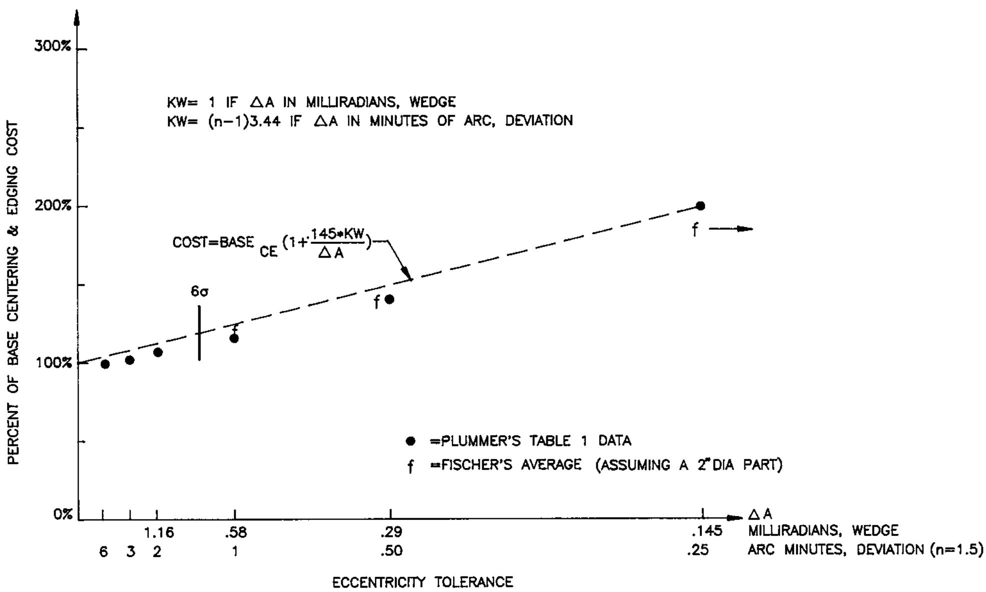

FIGURE 1.12 Relative cost vs. reciprocal tolerance according to various authors concerning eccentricity.

Milling/Generating Costs. We find a reasonable fit with experience for the milling cost to be given by Equation 5, where LM is the number per lot to be milled and d is the diameter in inches. The milling of both sides of the lens is included here.

$$MG = 4 + 90/LM + 0.1 \times d^2 \tag{5}$$

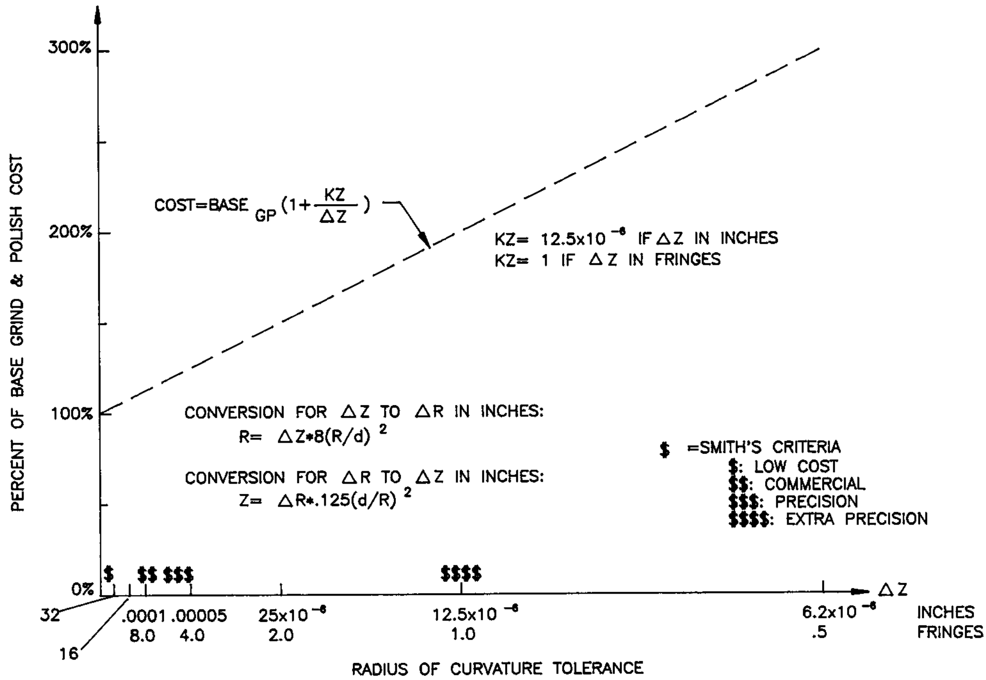

FIGURE 1.13 Relative cost vs. reciprocal tolerance according to various authors concerning radius of curvature (when expressed as sagittal error).

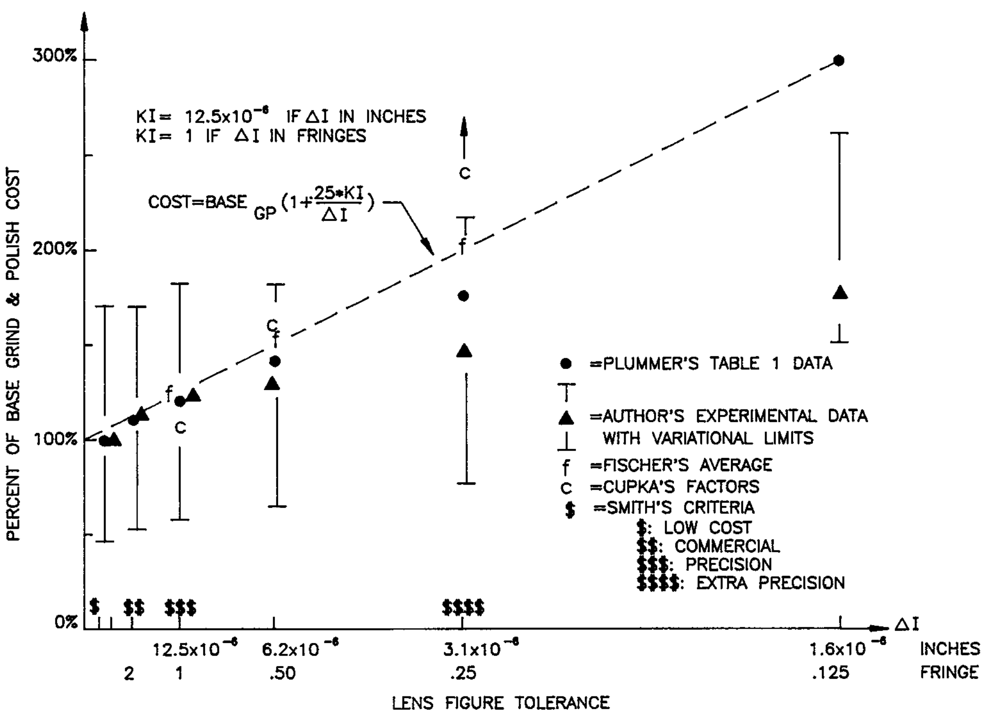

FIGURE 1.14 Relative cost vs. reciprocal tolerance according to various authors concerning lens figure irregularity.

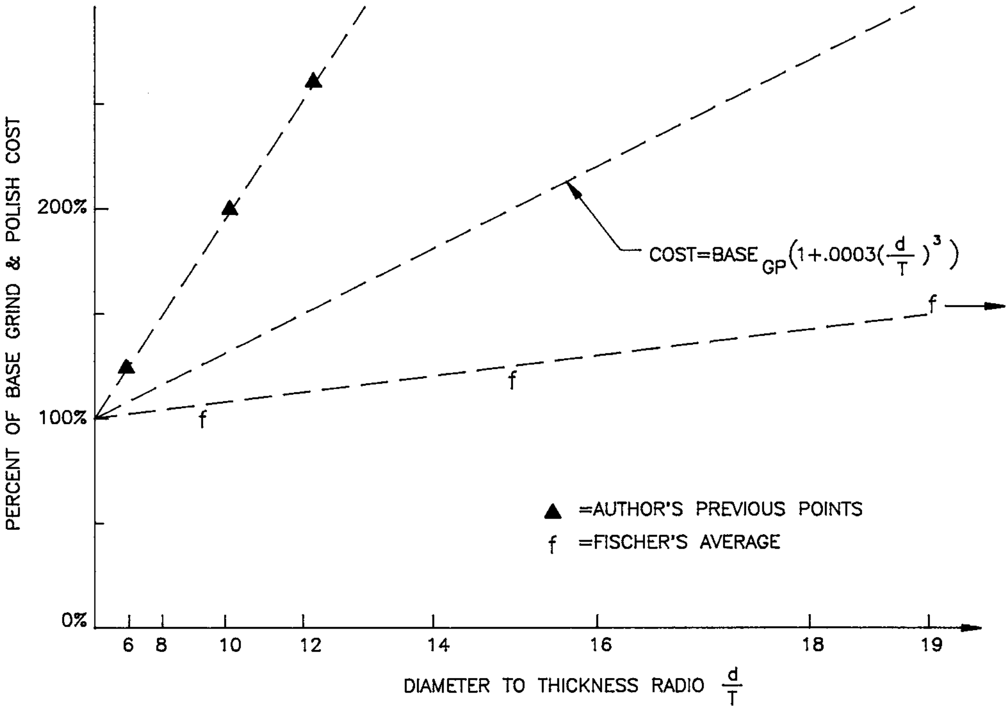

FIGURE 1.15 Relative cost vs. reciprocal tolerance according to various authors concerning lens diameter-to-thickness ratio.

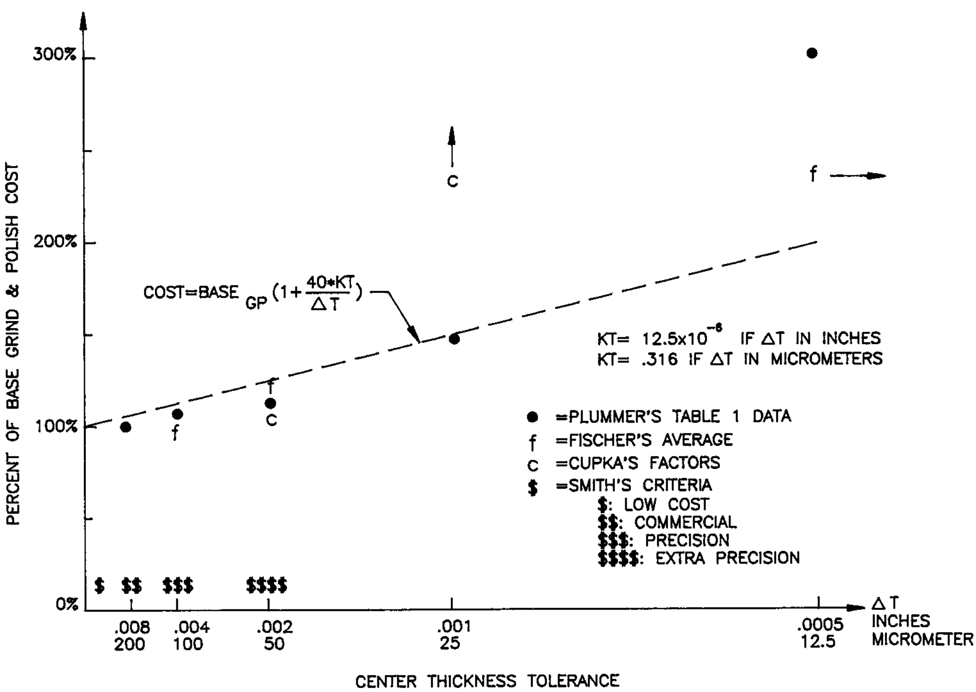

FIGURE 1.16 Relative cost vs. reciprocal tolerance according to various authors concerning lens center or axial thickness.

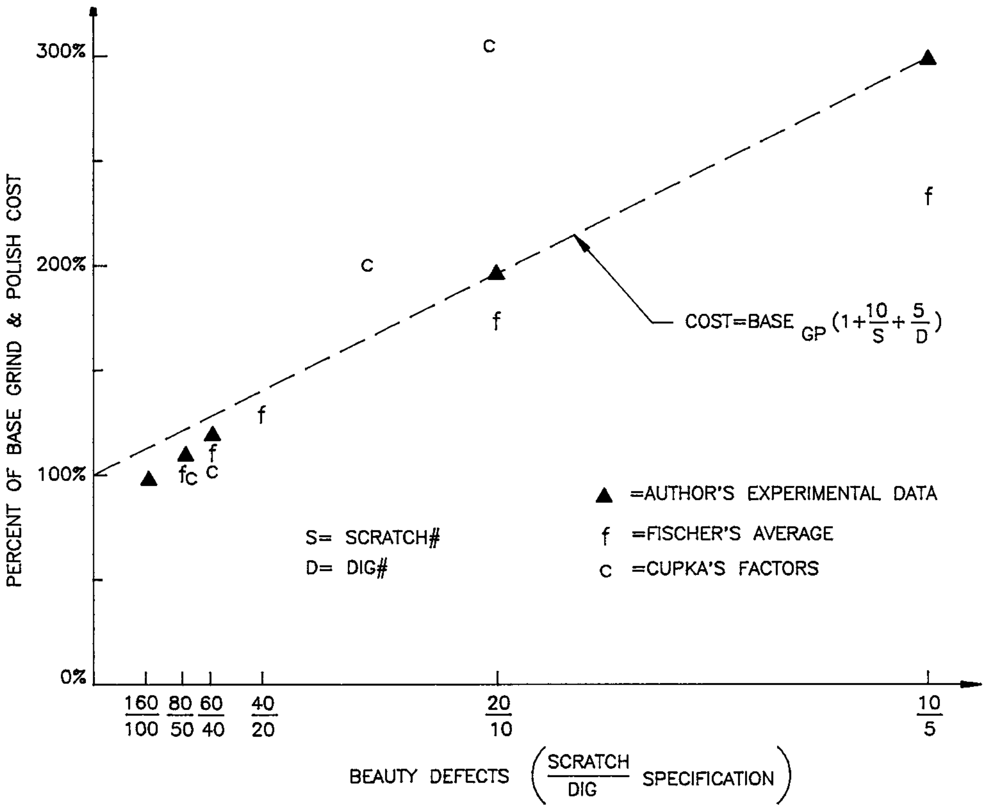

FIGURE 1.17 Relative cost vs. reciprocal tolerance according to various authors concerning surface finish or scratch and dig.

This implies that there is some base cost to mill the optic plus some setup cost divided by the number of parts to be run from that setup plus a factor due to lens size. Since the material to be removed from a molded blank is usually about the same thickness independent of blank diameter, the generating time is only a function of a blank's area (d^2).

Centering and Edging Costs. We will deal with centering and edging (CE) costs first because they are simpler. The cost is a function of the number in the lot LC to be centered in one setup, the diameter d, the number of chamfers C, and the number of flats F (planes perpendicular to the lens axis). Equation 6 represents our collective best estimate.

$$CE = \left(2 + d + C + F\right)/3 + \left(30 + 10 \times C + 15 \times F\right)/LC \tag{6}$$

This accounts for a setup cost for the diameter, chamfers, and flats plus the edging of each lens.

Grinding and Polishing Costs. The GP cost is a very strong function of the number of lenses which can be ground and/or polished at one time on a block. If the radius is short, the number NS which can be blocked for that side is determined by the radius R and the lens diameter d. If the radius is long, the number which can be blocked is determined by the maximum block diameter G and the lens diameter d. The precise calculation of this number can be performed when flats, chamfers, and center thicknesses are properly accounted for. For simplicity and practicality, we will use a conservative approximation without showing its derivation here. If the factor R/d is smaller than 0.87, only a single lens can be polished at one time. At least three per block can be

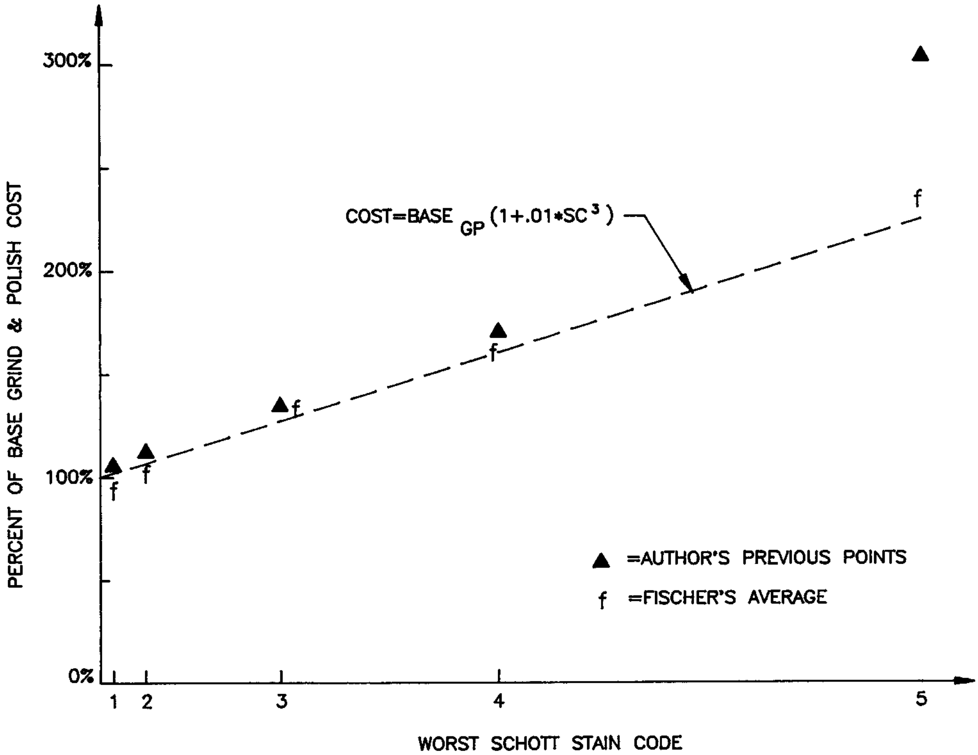

FIGURE 1.18 Relative cost of glass stain characteristics according to various authors.

processed if the ratio is greater. If the radius is short, the number per block NS will be given by integer value in Equation 7.

$$NS = INT\left[2.9 \times \left(R/d\right)^2\right] \tag{7}$$

Here R is the radius and d is the diameter of the side in question, and N1 would be NS for side 1 and N2 would be NS for side 2. If the radius of the side is long, the number per block will depend on the diameter G of the largest block which can be used as shown in Equation 8.

$$NS = INT\left[0.64 \times \left(G/d\right)^2\right] \tag{8}$$

Whichever NS is smaller for a given side (for R or G) will be the number which can be ground and polished per block for that side. Our experience is that less than the maximum number will often be used when the radius is long. This is not of great cost consequence, however, because the change in cost for a few parts, when NS is large, as given in Equation 9 is small relative to other costs. Our collective experience with the GP per lens and per side in a block of NS lenses is given in Equation 9.

$$GP = 7 + 14/NS \tag{9}$$

It is also usually appropriate to consider a yield-related factor Y due to scrappage etc. This factor Y multiplied by the number of lenses to be delivered is the number of lenses to be started to give the required yield. The factor Y is then actually an inverse yield. When this factor is applied and both sides of lenses are considered, the total GP base costs are given in Equation 10:

$$GP = Y \times \left[14 + 28 \times \left(1/N1 + 1/N2\right)\right] \tag{10}$$

where N1 and N2 are the numbers of lenses per block for side 1 and side 2. We now have all of the components of the base cost (in relative units) for a lens based on a collective empirical history and in a fairly workable form. We only need to apply the effects of tolerances and other influences such as material properties and diameter-to-thickness ratio to these base costs in the next section and we can then predict the cost to produce a given lens to its specifications. The derivative of that cost for a change in any tolerance or parameter will be used to determine the distribution of tolerances which will minimize the system cost and ensure the desired yield.

Effect of Tolerances and Other Factors on Costs

We have updated and included here the cost vs. tolerance data presented in our previous work[13] in Figures 1.11 through 1.18. In these figures, the dots are Plummer's work,[15] the triangles are our previous experimental and/or estimated values, the c's are factors from Cupka's earlier estimates, and the f's are the averages from Fischer's survey.[17] The lines drawn are the functions which we currently choose to use as the best estimates from experience and the functional equations are indicated on each figure. In the figures, where appropriate, we have included information from Smith.[20] He has defined tolerances that are "low cost" which are plotted as "$". His "commercial" tolerances are shown as "$$", "precision" tolerances as "$$$", and "extraprecise" tolerances are shown as "$$$$". These are in general agreement with the other data. For a more extensive discussion of the previous work, the reader is encouraged to review the references.

Figure 1.11 shows the cost effect of lens diameter tolerances which is not a strong factor up to the limit of the capabilities of the edging process. The various authors' data are in good agreement. We have shown the tolerance scale in inches and micrometers on this and some of the other figures for the readers' convenience. Figure 1.12 shows the cost effect of the lens centering tolerances. The centering tolerance is sometimes expressed as light deviation, wedge, or total indicator runout (TIR) at the edge of a lens or window. These different versions of the requirements can be reconciled in the following way. The wedge angle in radians is the same as the TIR divided by the diameter being measured. The light deviation depends on the wedge times the refractive index minus one and converted to arcminutes as needed. The KW-factors in Figure 1.12 account for this. Plummer's and Fischer's data are in good agreement and are well represented by the approximation which we use.

Figure 1.13 deals indirectly with the cost of radius of curvature tolerances. The work of Thorburn[22] reported in Wiese's collection made it clear that we should refine our approach for the cost effect of radius of curvature. We had previously just used the percentage of the radius as a measure of stringency of the radius tolerance. But it becomes apparent as a result of simple analysis that the change in sagitta over a surface is much more meaningful. This is what a spherometer or an interferometer can measure. One can derive the change in sagitta with a change of radius as approximated by Equation 11.

$$\Delta Z/\Delta R = 0.125 \times \left(d/R\right)^2 \tag{11}$$

This equation has been used to plot Figure 1.13. The graph is plotted as cost vs. the reciprocal of the delta sagitta ΔZ which is a function of delta radius ΔR, R, and d. The ΔR and R alone are not enough to estimate the real cost of the tolerance; the surface diameter must also be taken into account. It might be more appropriate to refer to the sagitta tolerance of a surface than the radius tolerance, but shops are used to seeing the radius tolerance spelled out on a drawing. Equations 12 and 13 can be used to convert either way between ΔR and ΔZ (typically in inches).

$$\Delta R = \Delta Z \times 8 \times \left(R/d\right)^2 \tag{12}$$

$$\Delta Z = \Delta R \times 0.125 \times \left(d/R\right)^2 \tag{13}$$

The chosen cost function is based on Thorburn's[22] statement that a tight tolerance on sagitta is 0.000005 in. and 0.0001 in. is loose.

Figure 1.14 shows the impact of surface figure irregularity tolerances. We described the earlier experimental data in Reference 12. We have added the points from Fischer and Cupka with a slight reformulation and labeling. The data from the various sources seem sufficiently consistent and are reasonably represented by the chosen function.

Figure 1.15 shows the effects of the diameter-to-thickness ratio of a lens. This has to do with the flexibility of the part when trying to hold a good surface figure and sometimes the temperature effects when working with thin negative lenses. Our simplistic treatments of this to date are probably not as adequate as we would like. This is probably worthy of further study. Fischer[17] reported a much less severe effect than our previous work.[13] In our previous work, we estimated that the effect was inversely proportional to the flexibility of a disk which goes as the cube of the diameter-to-thickness ratio d/T. We will now accept (conditionally) Fischer's data collection as possibly more representative and we have fit our function to a compromise between our old data and that of Fischer, where both are seen in Figure 1.15. There is no doubt that plano–plano windows can be worked very thin by contacting them to rigid plates which lend the effect of their own d/T ratio. In the case of lenses, this is not the practice, however. Structural shape, thermal effects, etc. need to be studied in more detail. We will have to use this function until we or someone else can refine it further.

Figure 1.16 deals with the cost of center thickness tolerances. A major cost factor here is that once the thickness goes under the tolerance limit, the lens is lost and must be replaced. We believe that this is what is reflected in the radical change in cost in Plummer's data as a 0.0005-in. tolerance is approached. At some point, a given tolerance would not allow a scratch or pit to be ground or polished out and the part would be lost due to the combination of the thickness and scratch and dig specifications. Plummer's and Fischer's data are in reasonable agreement, and Cupka only divided the costs into above and below a 0.002-in.-thickness tolerance.

Figure 1.17 shows the effects of scratch and dig tolerances. The various data cluster reasonably near the very simple function that we choose to employ.

Figure 1.18 shows the simple function of cost vs. glass stain code that we fit to our previous data and Fischer's data. We might comment that we feel that even the 5-code glasses can be worked on a fixed price basis now, but at a significant added cost.

The "polishability" factor should also be applied as a cost impact. In general, pyrex takes more time to grind and polish than BK7 and fused silica takes longer than pyrex, etc. In an attempt to include this factor, we have collected the estimated time factor for a variety of materials as compared to BK7. Some suggest that germanium has the same polishing time as BK7, and others think it takes significantly longer. This will, of course, vary from shop to shop and the procedures used. The authors believe that, typically, germanium lenses will polish to specifications in about the same time as BK7, but that the typical polishing specification may be 160/100 for Ge and 80/50 for BK7. Therefore, the Ge surface will take longer to meet all of the same specifications as a piece of BK7. The numbers in Table 1.2 reflect the authors' best estimates based on a variety of inputs, but each shop needs to examine which factors to use in its own case. The grind and polish time will then be multiplied by this material polishability factor P.

TABLE 1.2 Polishability Factor of Various Materials

	%
BK7	100
SF56	120
Pyrex	125
Germanium	130
Fused silica	140
Zerodur	150
ZnS, ZnSe	160
FK2, BaF2, Amtir	170
LaKN9, LaFN21	200
Electroless Ni	250
CaF2, LiF	275
MgF2, Si	300
Electrolytic Ni	350
Ruby	700
Sapphire	800

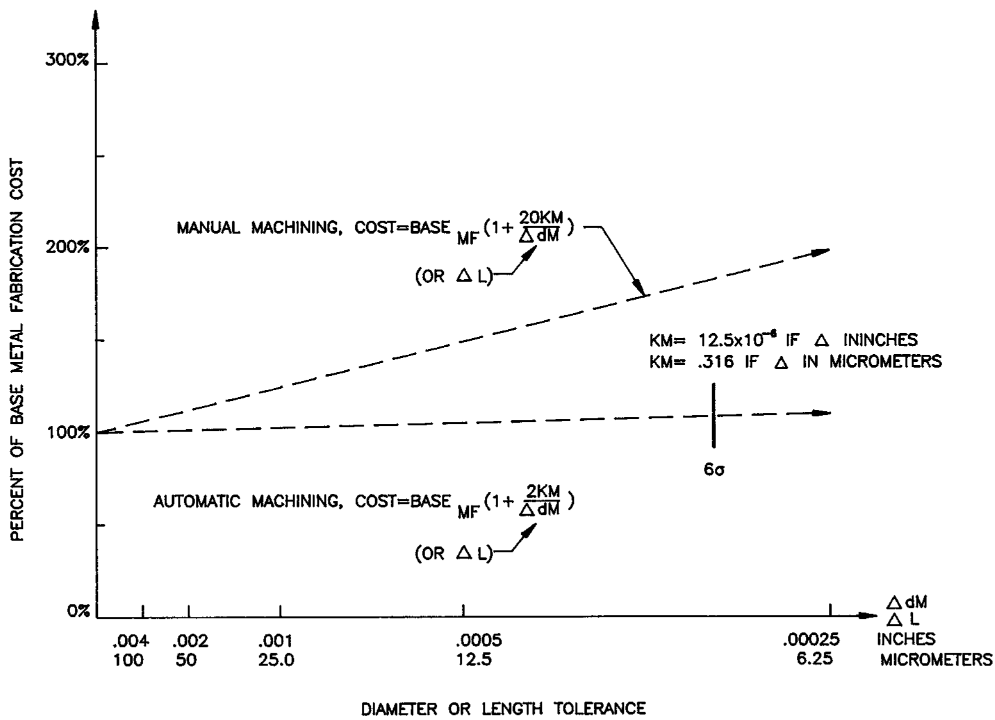

FIGURE 1.19 Relative cost vs. reciprocal tolerance concerning lens bore diameters and lengths.

Once the lenses are fabricated, they are typically mounted in metal cells. We have to coordinate the tolerancing of the metal and the glass to get the desired results. Similar cost vs. tolerance curves can be developed. Figure 1.19 shows the cost vs. lens cell diameter tolerance for both manual and automatic machining. The same curve applies to the length tolerances between bores as in spacers. In an automatic setup, the parts repeat within the capability of the machine with little change in the cost vs. tolerance. The manual operations are more and more labor intensive as the tolerances increase. Figure 1.20 shows the cost vs. tolerances of bore concentricity and length run-out such as tilt in a spacer. There is no difference between manual and automatic here. However, the big difference is whether both bores are cut without removing the part from its holder (chuck). If so, then the concentricity will be limited only by the accuracy of the machine. If the part must be rechucked for the second bore, much more time is consumed to hold a tight tolerance in the rechucking or mandrel-type operation. This points out the strong motivation to design the cell for single chucking as much as is practical.

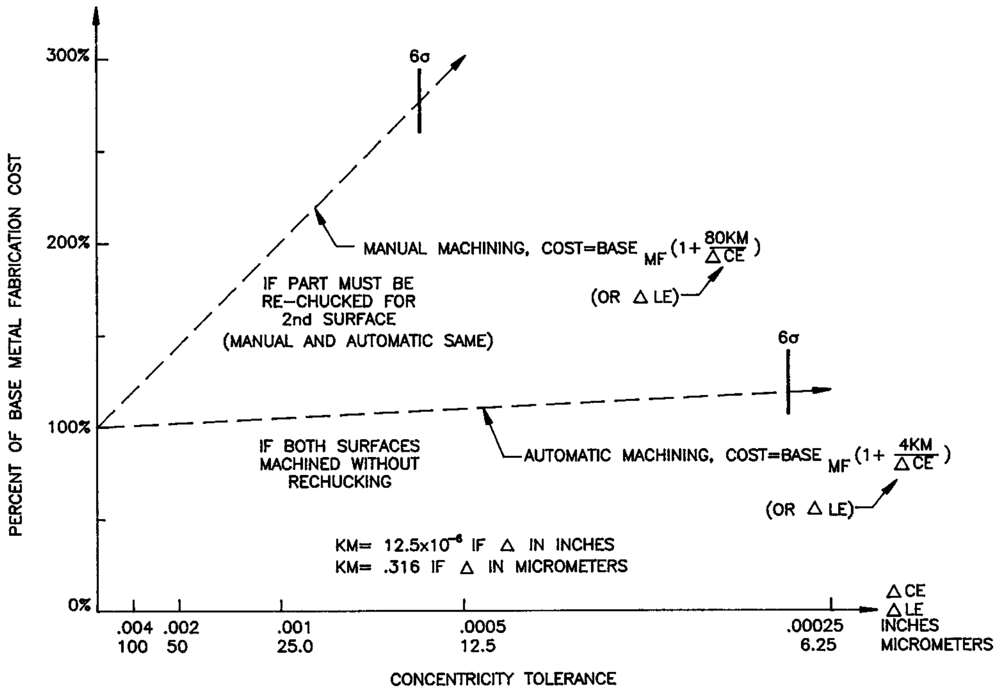

FIGURE 1.20 Relative cost vs. reciprocal tolerance concerning lens bore concentricity (ΔCE) and tilted and length run-out (ΔLE).

We will use the results of Figures 1.19 and 1.20 when we distribute the tolerances in the instrument design. The base cost for the machining operation will be multiplied by the metal tolerance factors in this process. For simplicity, we will ignore the setup costs and use a base machine fabrication cost of six (6) cost units per bore or spacer cut in the same denomination as those used for the lenses above. For an automatic setup, we will use a number of four (4) units. This aspect could be made much more complex and refined. This approximation is adequate to allow us to properly distribute the tolerances in the later sections, but cannot be used to estimate the total machining costs. These cost vs. tolerance and cost vs. other parameter functions are the major factors influencing the optical component fabrication cost. Now that we have them reasonably well characterized, we can apply them to the base costs specified earlier to estimate the total component cost. We can also find (to a sufficient approximation) the change in total component cost with a change in any of the parameters or tolerances. This, coupled with the sensitivity of the system performance characteristics to each of the parameter and tolerance changes, will then allow us to distribute the tolerances in such a way as to minimize the system cost while maintaining the required successful product yield. This is the major goal of the cost optimization process.

Total Lens Cost Estimation

When we combine the base costs given in Equations 5, 6, and 10 with the tolerance and other factors of Figures 1.11 through 1.18 and Table 1.2, we can estimate the total lens cost as given in Equation 14. This obviously is not scientifically rigorous, but is only a practical estimate for engineering or business purposes. The detail factors will vary from shop to shop and time to time.

It is the authors' belief that this type of estimation formula is more accurate and consistent than any other method now available.

$$
\begin{aligned}
&\text{TOTAL LENS COST} && = MT = \\
&\text{(GENERATING)} && + 4 + 90/LM + 0.1 \times d^2 \\
&\text{(PART SETUP)} && + Y \times 14 \\
&\text{(SIDE 1, G\&P)} && + (P \times Y \times 14/N1) \\
& && \times \{(1 + .25 \times KI/\Delta I1) \\
& && \times [1 + (R1/d)^2 \times 8 \times KZ/\Delta R1 + .0003 \times (d/T)^3] \\
& && + 40 \times KT/\Delta T \times (1 + 10/S1 + 5/D1) \times (1 + .01 \times SC^3)\} \\
&\text{(SIDE 2, G\&P)} && + (P \times Y \times 14/N2) \\
& && \times \{(1 + .25 \times KI/\Delta I2) \\
& && \times [1 + (R2/d)^2 \times 8 \times KZ/\Delta R2 + .0003 \times (d/T)^3] \\
& && + 40 \times KT/\Delta T \times (1 + 10/S2 + 5/D2) \times (1 + .01 \times SC^3)\} \\
&\text{(CENTERING)} && + CE \times (1 + 10 \times KD/\Delta d + .145 \times KW/\Delta A) \qquad (14)
\end{aligned}
$$

Although Equation 14 is extensive, it is not particularly complex. Table 1.3 defines the parameters of this equation.

TABLE 1.3

MT	Total lens cost estimate per piece (in relative units)
MG	Milling/generating cost from Equation 5
LM	Number of parts milled in one lot setup
Y	Yield factor of parts started/parts acceptable
P	Polishabilty factor from Table 1.2
N1	Number of parts/block for R1 side from Equation 7 or 8
N2	Number of parts/block for R2 side from Equation 7 or 8
R1	Radius of side 1
R2	Radius of side 2
G	Diameter of largest block to be used for grind and polishing
d	Diameter of lens
T	Thickness of lens
S1	Scratch number spec for side 1
S2	Scratch number spec for side 2
D1	Dig number spec for side 1
D2	Dig number spec for side 2
SC	Worst stain class of glass type Ks and deltas as in Figures 1.11 to 1.18
CE	Centering and edging base cost per Equation 6
C	Number of chamfers on the lens
F	Number of flats on the lens
LC	Number of lenses centered and edged in one lot setup

The generating cost and a certain portion of the setup costs are virtually independent of the tolerances of the part. The centering and edging operation and tolerances are virtually independent of the grinding and polishing operations and tolerances. The operations and tolerances of the two radii of the lens are in most cases independent. This is all reflected in Equation 14, and some logic has been applied to whether the tolerance factors are added or multiplied and to what they are applied. The irregularity cost factor is multiplied with the base cost of grinding and polishing each side. This irregularity cost is further multiplied by the radius tolerance cost factor plus the flexure (d/T) difficulty factor which makes the figure irregularity harder to achieve. The decision to add the d/T influence to the radius tolerance influence was made on the basis that they do not significantly affect each other, but they both affect the cost of meeting the irregularity specifications. The stain class interacts with and affects the scratch and dig requirements and they both increase the difficulty of holding the thickness tolerance, but they do not affect the irregularity degree of difficulty. The edging operation is affected by the diameter tolerance and the wedge or deviation tolerance. We will discuss the unique interaction of the diameter and centering tolerances in the next section. This formula then reduces the estimation of the fabrication cost of most lenses to a clerical task of entering the parameters from the lens drawing into a simple computer program. This could also, in principle, be worked into a CAD program to allow the designer to see immediately the cost impact of the design and tolerances. The key factor for tolerancing, however, is that we can find the partial derivatives of Equation 14 with respect to each of the reciprocal tolerances for use in finding the minimum cost tolerance distribution for a system as discussed below.

Interactions of Lenses and Mounts

A lens system typically consists of lenses in metal cells. The cells have bores that closely fit the lens diameters with "seats" or rings of contact between one or both of the radii and a metal locating surface in the direction of the optical axis. The relationship of one lens to another will be determined by the spacing dimensions and tolerances of the mount and the concentricity and tilts between the locating surfaces. The mounting cell interfaces must be toleranced to be compatible with the glass tolerances. In this section, we will address how to handle this task. There are several interesting references[21-23] dealing with tilts, decentrations, and rolls (by whatever names), but none seem to have addressed them in an analytic form appropriate to our cost minimization goals.

Lens Centering. Figure 1.21 shows a lens in a cell bore where only centering factors are considered, not tilts. It is easy to evaluate the effect or sensitivity of decentering a lens from the intended optical axis in most lens design software. This decentering in a system is the sum of several factors. The decentering of the optical axis of the lens with respect to the outside diameter of the lens is what the optical shop works on. The centering of the mounting bore with respect to the ideal axis is what the machine shop works on. There needs to be some fit clearance, f, for the assembler to insert the lens into the cell. The tolerances of the lens diameter d and the bore diameter dM, Δd, and ΔdM give rise to more potential clearance. These clearances will allow the lens to move to extreme positions in the cell and cause more decentering. Equation 15 expresses the possible total decentering td as a function of the lens decentering dc and the factors mentioned above.

$$\mathrm{td} = \mathrm{dc} + \left(\mathrm{f} + \Delta\mathrm{d} + \Delta\mathrm{dM}\right)/2 \tag{15}$$

The costs vs. tolerances have been defined and quantified above. We can see that the least cost distribution of tolerances for a lens in a bore with respect to total decentering will dictate a certain ratio between the dc, Δd, and ΔdM. We will then reduce the tolerancing of the "set" to a function of only Δd since the dc and ΔdM can be taken as the dependent variables. The dc of the lens is measured by the wedge ΔA in Figure 1.21 (here we work in milliradians) or in the arcminutes deviation that it causes which we relate by the factor KW in Figure 1.12. We can show that dc can be expressed in terms of ΔA as in Equation 16. (Some RSS possibilities will be ignored here.)

$$\mathrm{dc} = \left[\mathrm{R1} \times \mathrm{R2}/\left(\mathrm{R2} - \mathrm{R1}\right)\right]/\left(1000 \times \mathrm{KW}\right) \times \Delta\mathrm{A} \tag{16}$$

ΔA in terms of Δd is found to be

$$\Delta\mathrm{A} = \mathrm{SQR}\left(.145 \times \mathrm{KW}/10 \times \mathrm{KD}\right) \times \Delta\mathrm{d} \tag{17}$$

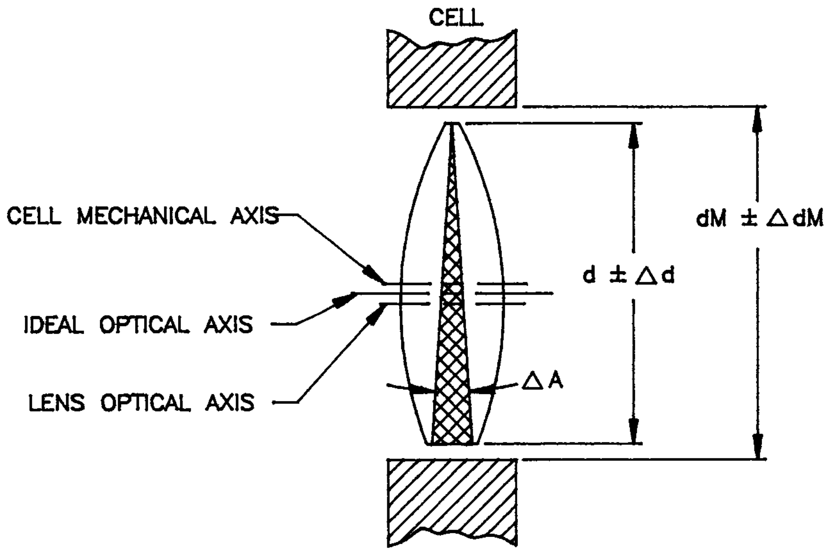

FIGURE 1.21 Decentering factors of a lens in a cell.

If we call the base machining cost MF, then ΔdM in terms of Δd is given by Equation 18.

$$\Delta dM = SQR\left(MF \times 20 \times KM / CE \times 10 \times KD\right) \times \Delta d \tag{18}$$

The minimum fit clearance factor f has to be determined by the assembly plans for the cell and whatever allowances are made for differential thermal expansion. At nominal temperature, it will allow a shift in an otherwise perfectly fitting cell of f/2. This decentering will have to come right off the top of the total decentering budget for this lens/bore set leaving the residual budget to be divided among Δd, ΔdM, and ΔA. We express the result in Equation 19.

$$\begin{aligned} td - f/2 = \Big[1/2 &+ SQR\left[\left(MF \times 20 \times KM\right)/\left(CE \times 10 \times KD\right)\right] \\ &+ SQR\left[.145/\left(10 \times KD \times KW\right)\right] \times ABS\left\{R1 \times R2/\left[\left(R2 - R1\right) \times 1000\right]\right\}\Big] \times \Delta d \end{aligned} \tag{19}$$

This factor multiplied by Δd is to be used in the tolerance allocation process with the decentering sensitivity to determine Δd. The Δd can then be used to assign ΔA and ΔdM (which are dependent on Δd) in a secondary operation using Equations 17 and 18. The diameter of the cell bore dM is also determined by this process as expressed in Equation 20.

$$dM = d + f + \Delta d + \Delta dM \tag{20}$$

Although Equation 19 might not appear to simplify anything, it is a relatively straightforward application of the cost and geometric factors which allows us to properly spread the cost and tolerances in the lens and its cell.

Lens Tilt and Roll. An otherwise perfect lens might be tilted with respect to the system's ideal optical axis because the metal locating surface of the cell is tilted by an angle ΔAT, as shown in Figure 1.22a. This can be simply dealt with using the cost vs. tilt curve in Figure 1.12 and the sensitivity of the system to tilt of the whole lens. This assumes that the tilt is not otherwise limited

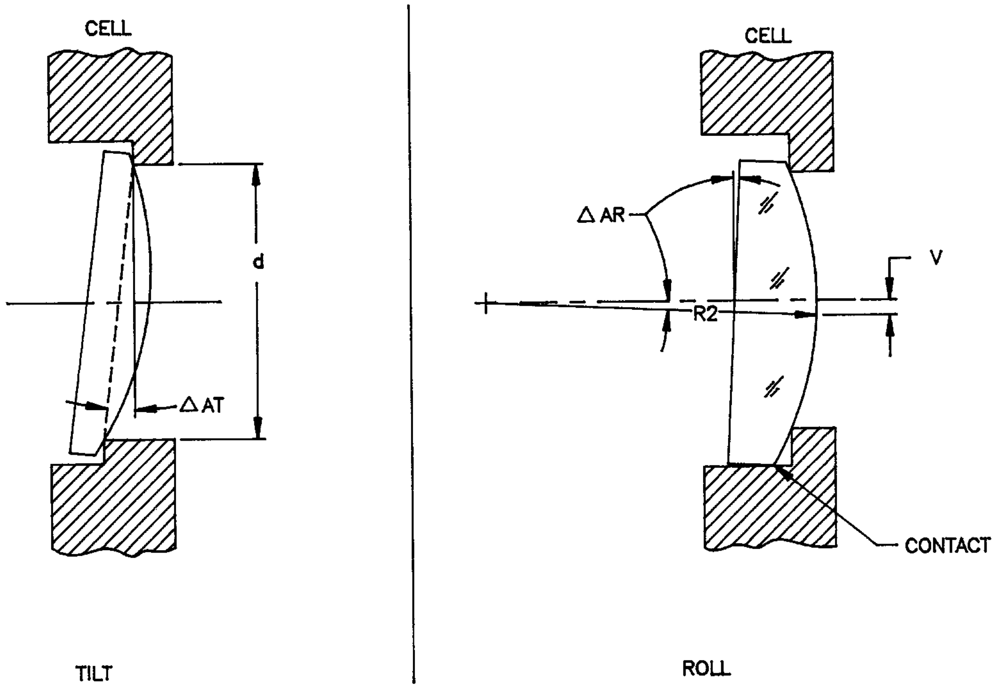

FIGURE 1.22 Tilt and roll factors of a lens in a cell.

by such factors as a retainer on the other side of the lens or the fit of the cylindrical lens diameter into the cylindrical cell bore which prohibits that much tilt. The perfect lens might also "roll" in an oversized bore as shown in Figure 1.22b. This shows that the left-hand surface tilts while the right-hand surface is correctly located against the "perfect" cell. The lens will roll about the center of curvature of the right-hand surface R2. The left-hand surface, which we have shown as plano for clarity, will tilt through an angle of ΔAR which is approximately V/R2 radians. We can show that:

$$V = \left(f + \Delta d + \Delta dM\right)/2 \tag{21}$$

We know ΔdM in terms of Δd from Equation 18. This allows us to express ΔAR1 of the left surface R1 as a function of Δd in Equation 22.

$$\Delta AR1 = \left(f + \left\{1 + SQR\left[\left(MF \times 20 \times KM\right)/\left(CE \times 10 \times KD\right)\right]\right\}\right)/ABS\left(2 \times R2\right) \times \Delta d \tag{22}$$

This and the system performance sensitivity to a tilt of surface R1 will allow us to allocate the tolerance budget for a tilt of R1. However, note that controlling this requires the control of Δd which is already determined by the decentering requirements! Generally, one or the other will be the more demanding on Δd. It would appear that we should find which is the more stringent and use it to determine the tolerance allocations. The other parameter would still make some contribution to the error budget, but not be independently determined. As in the tilt case, the ability of the element to roll to the full extent indicated in Equation 22 may otherwise be restrained by (but might also be caused by) a retainer ring, etc. Since this can be the case, it may be appropriate for the designer to use some judgment in the application of tilts and rolls after looking at the sensitivities and the mounting designs to decide how these equations may be applied. This unfortunately seems to bring us back a bit from the engineering toward the "art" in tolerancing. Having now dealt with the necessary elements of cost vs. tolerances and other parameters and the applicable geometric simplifications, we will next proceed with how we can apply the foregoing equation to distribute the tolerances for the maximum performance and cost benefit.

Allocation of Tolerances

The term **Six Sigma** (6σ) has come to be referenced extensively in recent years with respect to quality and tolerancing functions. Six Sigma is a tool and philosophy which can improve yield and reduce rework, and thereby improve costs and customer satisfaction. There are two major elements to it. First is to measure the routine capabilities of the processes used to achieve a result. This might include the part-to-part repeatability of a lens diameter. If there are things that can be done to reduce the variations, these will improve the capabilities and therefore yields. Statistical process control (SPC) is a part of this approach. The second major element is to make the tolerances as large as possible and still meet the system performance requirements.

We have indicated in Figure 1.11 where the 6σ point would be for the lens diameter. That is, a tolerance less stringent than ±0.0003 in. would give 6σ results or better. Anything more stringent (to the right in Figure 1.11) would be beyond the capability of that process to provide 6σ results. Figures 1.12, 1.19, and 1.20 also show the 6σ points for the processes of lens centering and metal machining.

Over the last decade, we have been heavily involved in finding methods to best distribute the tolerances in a way as to minimize the cost. This is essentially consistent with 6σ philosophy. One aspect has been to keep the designs as simple as possible and still meet the requirements. This is a key element of a 6σ design. The less complex the design, the lower will be the cost, risk, rework, etc.

Let us digress from the subject to describe our design philosophy. We first try the simplest design we think has a chance of working. If testing shows that it must be more complex, we change it. If we had made it more complex to start with, we would never know if it could be simpler and less costly. This is often a necessary part of the development process.

6σ graphically demonstrates the fact that it is important to identify the weak link in the capability chain and work to improve it. Even if the rest of the links in the process are better than 6σ, the results will still only be as good as the weakest link.

6σ is valuable in focusing the attention on tools and philosophies which can be used to continually improve the processes and designs. Both manufacturing and design organizations have been generally doing many of the right things from a 6σ point of view, but now have a better focus and understanding of how to measure and execute the process. The effects of 6σ will be evolutionary in the future.

Assigning Tolerances for Minimum Costs — An Example

We described the general principles of distributing the tolerances for a minimum cost in our first work[11] on this subject. Adams[16] made some significant additions to our work which we shall apply below. If there is more than one performance criterion that must enter into the tolerancing process, the solution to the equations is somewhat involved, but it is feasible. However, many systems, including the one which we will use as an example, have one performance criterion which dominates all of the others as it relates to the tolerances. That is, if the tolerances are chosen to meet that dominant performance requirement, then all other requirements will also be met. This reduces the computation considerably and makes it easier to visualize. For the balance of this discussion, we will use the single requirement case with the understanding that it can be extended to multiple criteria by the methods of the previous papers as needed.

Figure 1.23 shows the multifocal length tracking telescope (MFLTT) that we will use as an example of the tolerancing process. It has a catadioptric telescope section of 300-mm aperture and about 2000-mm focal length with a 25% central obstruction due to the secondary mirror. The telescope image is then collimated by a focus lens set. The afocal beam is then imaged by one of three imaging lenses to the final focal plane. These lenses are alternately positioned in the beam to give system effective focal lengths of 1000, 2000, and 4000 mm. Before the final focal plane, there is an auto-iris system of variable neutral density filters and a reticle projection unit (AIR). There is also a 500-mm system which is partially separated from the others to allow a larger field of view. The 500-mm system is folded into the same optical path as the others by a movable prism. There are sealing windows in front of the telescope and the 500-mm lens. In this complex telescope system example, the most stringent requirement of the system is the on-axis MTF at 30 lp/mm. When this is satisfied, the off-axis MTF at 30 lp/mm, the on- and off-axis MTF at 10 lp/mm, and the boresight, etc. requirements will all also be satisfied without additional tolerance requirements.

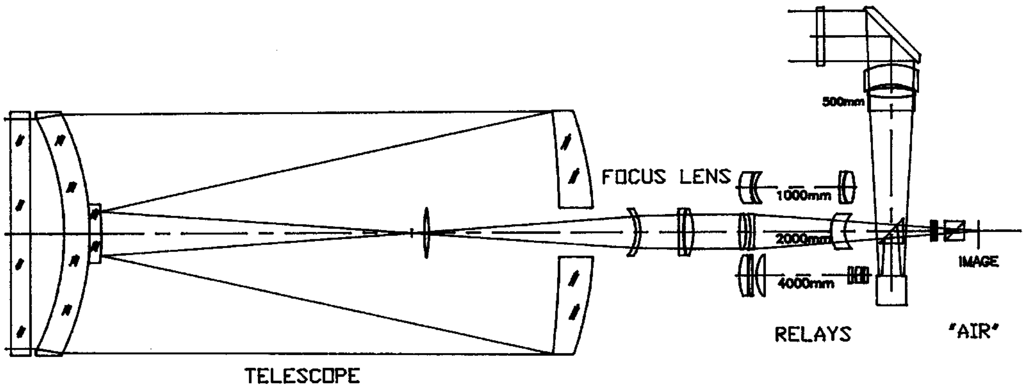

FIGURE 1.23 Example system: Multifocal length tracking telescope.

To be consistent with our previous report,[11] we will designate this performance requirement by E which represents the maximum permissible error in MTF from ideal for the system. We will actually convert this E to units of RMS wavefront error (RMSWE) for simplicity. The total E will eventually be partitioned among each of the tolerances which affect it. (To make a tractable example for this discussion, we will partition the total E among the various sections of the system.) The partial E will then be allocated to the parameter tolerances within one section based on the cost minimizing technique. This "divide and conquer" approach is needed here, plus any justifiable simplifications, in general, to reduce the overwhelming magnitude of the problems that may have multitudes of component tolerances to be determined. In the final analysis, it is best to "tolerance" the whole optical train from object to image in one operation. This will truly allocate the tolerances to achieve the required performance at the minimum cost. The simplifying partitioning will cause some deviation from the ideal result unless the estimate used in the partitioning was exactly correct. In the example used here, it would be best to tolerance the 4000-mm system from end to end, but the data would be too cumbersome to make a good illustration in this discussion.

Simplifying Approximations. The MTF of a system is often the best performance measure to use because it most directly relates in many cases to the performance of an overall system when it is used. It is, however, not generally possible to measure the MTF effect of each component lens of a system in the production process. The characteristics that are readily measured on a lens were discussed above, such as irregularity, radius, centration, etc. We chose to work here with the effects of each tolerance on RMSWE, because it can be reasonably related to the system MTF. We estimated the reduction in MTF per wave of RMSWE at 30 lp/mm for the 2000-mm effective focal length, f/8 system by introducing errors into the system and evaluating it for MTF and RMSWE. With parameter deviations, we produced defocus, spherical aberration, coma, and astigmatism. Defocus was introduced by evaluating the system at different focal planes from the best focus. Spherical aberration was introduced by varying the y^4 aspheric coefficient from the nominal. Coma was

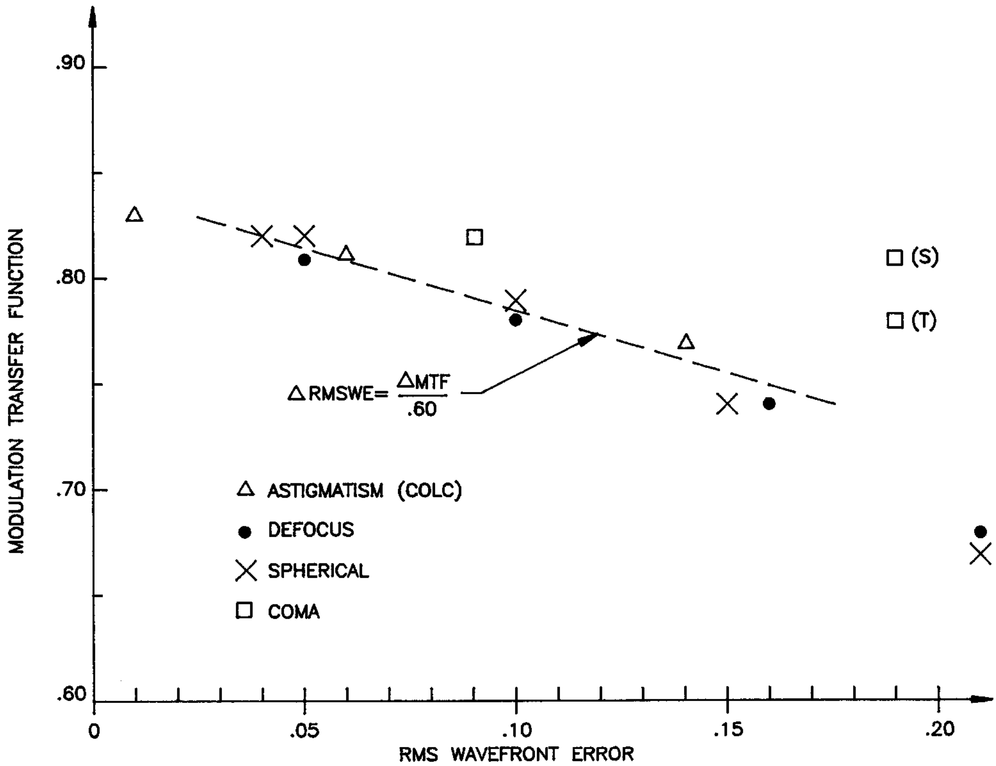

FIGURE 1.24 Diffraction MTF at 30 lp/mm vs. wavefront error for a 2000-mm, F/8 lens with various aberrations.

evaluated in an "equivalent" (f/8, 2000 mm) parabolic mirror system with the stop at the focal plane so that astigmatism was zero. The system was evaluated off-axis to introduce coma.

Last, astigmatism was introduced in an "equivalent" (f/8, 2000 mm) Ritchy-Chretien telescope where coma and spherical aberration were zero. The system was evaluated off-axis to introduce astigmatism. The results appear in Figure 1.24. All of the data form a reasonably consistent pattern except the coma. We do not presently understand this anomaly which may be worthy of a separate study. However, since the effects of coma are less severe than the others, we will ignore them and use the conservative numbers indicated by the others. Therefore we will use ΔRMSWE = ΔMTF/0.60 as the amount of reduction in MTF that will be accompanied by a corresponding RMSWE. This will allow us to work with the effects of tolerances on the RMSWE which we will assume are quasilinear in the regions where we are applying them. This may be a conservative estimate, but we would like to err on that side. Another approximation that we will draw upon comes from Smith[20] where

$$\text{RMS} = (\text{Peak-to-Valley})/3.5 \tag{23}$$

approximates the RMSWE expected from most types of error. It would seem that sharp departures over a small portion of the wavefront would violate this rule, but those are not usually encountered. We conducted a small investigation of our own by comparing the RMS and P-V data on many interferograms from a ZYGO interferometer. We discovered that the factor in Equation 23 might be more like 7 than 3.5 when small irregularities, such as those on spherical surfaces, are examined. For this example, however, we will use Smith's value. In the example system, the apertures were selected at the first-order stage to yield the required MTF when the diffraction effects of the obscuration plus one-quarter wave of design and fabrication errors were taken into account. This is not much error to spread across the many elements from the object to the focal plane. One benefit is the fact that certain compensating alignments can be made at assembly since such systems are not made in very large quantities. We will use the approximation of Equation 23 to establish a preliminary total error budget of 0.071 RMSWE (1/4 wave P–V) from all sources in laboratory tests. In the final application, obviously, atmospheric and other effects might influence the results further.

Error Budgets. Next we need to decide how to distribute this 0.071 RMSWE among the many facets and tolerances of the system. Smith[20] describes how to work with the root sum of the squares (RSS) to combine error effects. McLaughlin[25] shows that RSS tends to be too pessimistic and Smith himself concludes that it may err on the conservative side. McLaughlin shows that the total system error will tend to be 0.42 times the RSS prediction if the fabrication errors have a Gaussian distribution which is truncated at the 2σ level. Although there is a major move at this time in industry to apply 6σ tolerancing as mentioned above, there is still work to be done along the lines of Adams'[16] contributions to properly incorporate it. For the present case, we have used 2σ in this case where individual adjustment and testing are required. We will therefore use McLaughlin's 0.42 factor for the fabrication errors. To simplify the example, we will partition the 4000-mm path of the system. In looking at Figure 1.23, we count 32 surfaces through the 4000-mm optical path. The authors chose to emphasize the sensitivity effects of mirrors by counting them twice to give 34 as the surface count. Of this 34, 8 are in the telescope, 8 in the focus system, 12 in the 4000-mm relay, and 6 in the AIR. The other paths are less complex. This one will be the critical path and set the pace for the telescope, focus system, and AIR tolerances.

We will allocate the budget to the four sections of the 4000-mm path (telescope, focus, relay, AIR) in proportion to the square root of the number of surfaces in the section divided by the total number of surfaces. This is an engineering estimate of the relative influence of each section. The division of the system into these sections is also logical because each section can be tested independently for RMSWE in production. Figure 1.25 shows the error budget broken down this way. The top level requirement was determined above to be effectively 0.071 RMSWE. We know from the design stage that the design has used up 0.030 RMSWE. Another analysis indicates that the effects of alignment focus errors and the laboratory environment should be on the order of 0.009 RMSWE. This leaves 0.0637 RMSWE to RSS with the other two (three) parts of that level of the budget to give 0.071 RMSWE. From McLaughlin's information and the assumption of Gaussian errors, we then divide the fabrication budget of 0.0637 by 0.42 to give 0.1517 RMSWE, which can be distributed over the four sections of the 4000-mm system. The bottom four boxes of the budget in Figure 1.25 show how these work out when the above argument is applied. We will work through the simplest section, the focus lens with 0.0736 RMSWE budget, as an example of the procedure for tolerance distribution to give the minimum cost while meeting the performance requirements (to within some statistical uncertainty).

Derivatives of Costs with Respect to Tolerances. In the assignment of tolerances for a minimum cost, it is necessary to have the partial derivatives of the total cost with respect to the reciprocal of each tolerance. These are basically derived from Equation 14 and the metal cell tolerance costs in Figure 1.20. Since we showed that the metal cell diameter and lens-centering tolerances can be made dependent on the glass diameter tolerance, we only have five types of tolerances to allocate in the framework which we have been discussing. These are irregularity, radius, thickness, diameter, and tilt of a lens due to the errors in the cell. Roll is shown in Figure 1.22b and can be derived from Δd. Tilt is just the cell run-out parallel to the optical axis (ΔLE of Figure 1.20) divided by the diameter d of the lens. We will call the partial derivatives of the total system cost with respect to the reciprocal of these tolerances \$I, \$R, \$T, \$d, and \$L. The \$L applies to both the metal cell bore diameters and axial run-out. The first three derivatives will be functions of the base grind and polishing costs with a common factor that we will call BP, which is defined in Equation 24.

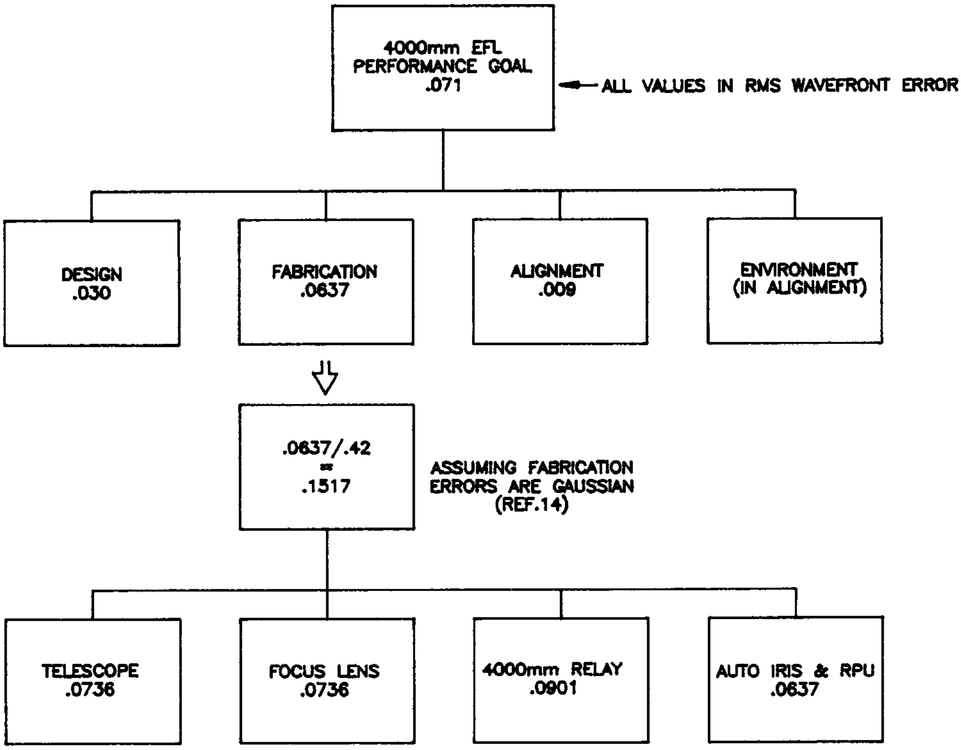

FIGURE 1.25 Error budget allocation for the 4000-mm channel of the multifocal length tracking telescope.

$$BP = P \times Y \times 14 / N\# \tag{24}$$

The N# is the N for the given surface number just as we will use ΔI#, ΔR#, etc. for those values associated with that surface number. The $d has the base centering cost CE as a factor, while $L has as a factor the base machining cost MF. There are additionally three "fudge" factors FI, FR, and FT associated with $I, $R, and $T which can be taken as unity for a simplifying assumption or calculated in each case as we will explain below. Equations 25 through 29 give these partial derivatives which will be needed to allocate tolerances.

$$\$I\# = BP \times FI \times 0.25 \times KI \tag{25}$$

$$\$R\# = BP \times FR \times 8 \times KZ \tag{26}$$

$$\$T = BP \times FT \times 40 \times KT \tag{27}$$

$$\$d = CE \times 10 \times KD \tag{28}$$

$$\$L = MF \times 80 \times KM \tag{29}$$

The factors FI, FR, and FT come from taking the partial derivatives of Equation 14 with respect to the reciprocal of the tolerances. They are due to the modifying effects of other parameters and are given in Equations 30, 31, and 32.

$$FI = \left[1 + 8 \times KZ \times \left(R\#/d\right)^2 / \left(\Delta R\#\right) + 0.0003 \times \left(d/T\right)^2\right] \tag{30}$$

$$FR = \left[1 + 0.25 \times KI / \left(\Delta I\#\right)\right] \tag{31}$$

$$FT = \left(1 + 10/S\# + 5/D\#\right) \times \left(1 + 0.01 \times SC^3\right) \tag{32}$$

There is a problem with Equations 30 and 31, however. The values of ΔR# and ΔI# are not known until after the tolerancing process. We can include the effects of FT in Equation 27 because all of its parameters are known at the start. We chose to include the effects of FI and FR after the tolerances are calculated by dividing the resulting tolerances for irregularity and radius by the cube roots of FI and FR, respectively. The reasons for this will be more apparent from Equation 35 below. The other alternative would be substituting FI and FR back in an iterative procedure. We now need only to address the simple allocation equations and procedures and we can finish the task of finding the least cost tolerances.

Tolerance Allocation Process. We showed previously[11] that the total error E was the sum of all of the contributions from each error source which contributions are the products of the tolerance value t_i times the sensitivity S_i of the performance to variations of that parameter, as seen in Equation 33.

$$E = \sum_{i=1}^{n} \left(S_i \times t_i\right) \tag{33}$$

(It may be found in the future that the application of 6σ to this process is simply accomplished by a proper adjustment of E, but we must "press on" without this understanding at this time.)

We have shown that the tolerances could be distributed for minimum cost by applying Equation 34 to each tolerance in turn. The A_i's are the coefficients of the derivatives of cost given in Equations 25 through 29. In Equation 34, the sum of all the square roots of the products of the cost coefficients A_k and the sensitivities S_k is divided into the total error budget E to get a constant which is multiplied by a function of the cost and sensitivity of each tolerance.

$$t_i = \sqrt{\frac{A_i}{\left|S_i\right|}} \frac{E}{\sum_{k=i}^{n} \sqrt{A_k \left|S_k\right|}} \tag{34}$$

Adams[16] pointed out that Equation 34 gave the solution which represented the tolerances all going to the worst-case limit condition, and this is obviously too severe. He showed that the RSS condition would be satisfied by the tolerance distribution if Equation 35 was applied instead.

$$t_i = \sqrt[3]{\frac{A_i}{S_i^2}} \frac{E}{\sqrt{\sum_{k=1}^{n} \sqrt[3]{\left(A_k S_k\right)^2}}} \tag{35}$$

The allocation is then very straightforward in the single E case. It is only necessary to develop a table of A_i and S_i for all appropriate i and process the data in accordance with Equation 35.

Table 1.4 gives these values for the example case, and the resulting t_i's which we seek are in Table 1.5. The example case has a total of four lenses which we illustrate in Figure 1.26. The field lens is close to the focal plane and has negligible sensitivity. We therefore remove it from consideration since we can assign minimum cost tolerances to it without affecting the rest of the task. This then reduces the example to a three-element lens where the surfaces and dimensions are numbered as in Figure 1.26. The necessary data processing is very conveniently set up and done using a spreadsheet program for data entry and all of the necessary calculations. Table 1.4 starts with the parameters determined from the "lens drawing" or design parameters. The number per block for each side is determined from Equations 7 and/or 8. The sensitivities of the system performance to each parameter are determined from the lens design program. The base costs and cost vs. the reciprocal tolerance derivative coefficients are calculated from Equations 25 to 29. The constant multiplier in Equation 35 is calculated. The individual factors from the individual sensitivities and cost derivatives are used to compute the tolerances for each of the toleranced parameters.

The adjustments to the I and R tolerances are made for FI and FR as mentioned in the previous section. Table 1.5 contains the resulting tolerances which needed to be determined. It is the set of tolerances for each of the toleranced parameters which will give the least cost solution and meet the performance with some statistical "RSS" certainty. The assumption per Adams[16] is that the errors will be distributed about the norm in a Gaussian manner and the tolerance limits will be 2σ.

However, a significant problem appears in Table 1.5. Many of the tolerances are well beyond what can be achieved in normal practice; they are off the chart in Figures 1.11 to 1.20! This is a disappointing result for the designer, but hopefully not the end of the road. Finding the problem at the design stage is not nearly as frustrating or expensive as finding it at the production stage. The lower part of Table 1.5 shows the application of the above formulas to compute the base cost for each of the lenses and the total costs when the tolerances are included. The total costs are about 15 times that of the simple base lens if they could even be made. "Off the chart" implies in most cases that it cannot be made or at least it would be much more expensive than the linear chart data would predict. We may have, therefore, identified here an impractical design. The designer then has the challenge of finding a design and/or an approach which will be less sensitive. The addition of more lens elements is not at all out of the question if they can reduce the sensitivity significantly. If the added cost of one or more elements reduces the tolerance costs sufficiently and all the lenses can be built, the total cost will be less than the first design. It might be possible to cement a doublet to get rid of a sensitive air space. It might be practical to make a centering adjustable and/or add other assembly tricks.[26] The designer can now evaluate the impact of design changes on cost by using the tools put forth in this discussion. The "bottom line" in Table 1.5 can tell him if he has improved the situation or not. What we see here is an example of the processes in Figure 1.2 where the tolerance sensitivity analysis and distribution feed into the producibility analysis which sends us back to the detail optical design for further work.

To solve this particular problem mentioned above, we first tested the ability to cement the achromats to reduce the critical air space sensitivity. It was found that the differential expansion of the glasses caused the cemented lenses to break at the extreme temperatures required for the system. The final solution was to design the lens cells so that the centering of each lens could be adjusted at assembly. This allowed the centering requirements on the lenses and cells and lens-to-cell fits to be significantly relaxed. Although the adjustment process was skill and labor intensive, an otherwise impossible result was achieved. The end result of this subsystem and the other subsystems in the telescope was essentially a diffraction limited performance.

A complex system such as the complete 4000-mm example system will have several times as many columns as Table 1.4, but the process is the same and relatively straightforward to apply. The most difficult aspect can be obtaining the sensitivities of the performance criterion to parameter variations. Existing lens design programs can do this with greater or less facility, but all are readily modifiable to generate the data needed. This data generation is computer intensive and time consuming, but probably unavoidable. The tolerancing program described above takes only a few seconds on a personal computer to calculate the tolerances for the six-surface case. It should

TABLE 1.4 Tolerance Data and Computations — Example Focus Lens Set

	Surface #					
	1	2	3	4	5	6
Radius	3.54	–6.504	–6.346	59.203	2.319	1.941
Diameter	2.35	2.35	2.31	2.31	2.04	2.04
Thickness	0.472	0.045	0.237	1.969	0.237	10.63
Scratch	80	80	80	80	80	80
Dig spec	50	50	50	50	50	50
Refractive index	1.5		1.6		1.5	
Stain ·	5.2	5.2	5.12	5.12	2	2
Polishability	3	3	1	1	1	1
Chamfers	1	1	1	1	1	1
Flats	0	0	0	0	0	1
Number/milling lot	7		7		7	
Centering lot size	7		7		7	
Yield	1.5	1.5	1.3	1.3	1-3	1.3
# Per block	6.58065	7	7	7	3	1
KI (fringes)	1	1	1	1	1	1
KZ (sag, in.)	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05
KT (thk, in.)	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05
KD (dia, in.)	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05
KM (met dia, in.)	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05	1.25E-05
KW (mrads wedge)	1	1	1	1	1	1
SI dRMS/fringe	0.03	0.03	0.06	0.05	0.03	0.04
SZ dRMS/in.	375	750	176.5	250	400	117.6
ST dRMS/in.	4	8	2	3	2	0.1
Std dRMS/in.	80		90		25	
SLE dRMS/radian	225		400		25	
f, clearance	0	0	0	0	0	0
MF (MFG base)	6	6	6	6	6	6
BP (Equation 21)	9.573522	9	2.6	2.6	6.066667	18.2
CE (Equation 3)	7.497619		7.484286		7.394286	
A$I (Equation 22)	2.393381	2.25	0.65	0.65	1.516667	4.55
A$Z (Equation 23)	0.000957	0.0009	0.00026	0.00026	0.000607	0.00182
A$T (Equation 24)	0.004787	0.0045	0.0013	0.0013	0.003033	0.0091
A$d (Equation 25)	0.000937		0.000936		0.000924	
A$L (Equation 26)	0.006		0.006		0.006	
[A(K)*S(K)]^2/3 I	0.172736	0.165767	0.114991	0.101829	0.127437	0.321136
[A(K)*S(K)]^2/3 Z	0.505111	0.769479	0.128164	0.161648	0.389033	0.357792
[A(K)*S(K)]^2/3 T	0.071561	0.109015	0.018904	0.024772	0.033258	0.009388
[A(K)*S(K)]^2/3 d	0.177792		0.192088		0.081117	
[A(K)*S(K)]^2/3 L	0.691039		1.02581		0.175497	
Constant E/SQR (sum) =		0.018141				
[A(I)/S^2]^1/3 I	13.85082	13.56855	5.65104	6.381321	11.89711	14.16387
[A(I)/S^2]^1/3 Z	0.001896	0.00117	0.00203	0.001609	0.00156	0.005089
[A(I)/S^2]^1/3 T	0.0669	0.041287	0.068772	0.052484	0.091213	0.969055
[A(I)/S^2]^13 d	0.005274		0.004873		0.011398	
[A(I)/S^2]^1/3 L	0.003697		0.002534		0.016765	
Delta I (fringes)	0.251274	0.246153	0.102518	0.115766	0.215831	0.256953
Delta Z (in.)	3.44E-05	2.12E-05	3.68E-05	2.92E-05	2.83E-05	9.23E-05
Delta AT (tilt, rd)	6.71E-05		4.6E-05		0.000304	

Note: E = 0.0736 RMSWE.

be approximately linear with the number of surfaces as long as only one performance criterion is to be considered. Multiple criteria would be more cumbersome to evaluate as we showed in our first work.[11] However, Adams[16] shows that it is most likely that one criterion is all that is needed, and more than two is highly unlikely.

TABLE 1.5 Resulting Tolerances and Costs

Tolerances	Surface 1	2	3	4	5	6
Delta R (in.)	0.000786	0.001644	0.003355	0.225109	0.000378	0.000839
Delta I′(fringes)	0.279124	0.32987	0.113803	0.13038	0.245058	0.268062
Delta T (in.)	0.001214	0.000749	0.001248	0.000952	0.001655	0.01758
Delta AR (roll, rd)	1.81E-05	9.88E-06	8.05E-06	8.63E-07	4.65E-05	5.56E-05
Delta td (in.)	9.57E-05		8.84E-05		0.000207	
Delta LE (in.)	0.000158		0.000106		0.00062	
Delta d (in.)	5.67E-05		4.51E-05		9.49E-05	
Delta dM (in.)	7.17E-05		5.71E-05		0.000121	
Delta A (mrad)	0.001931		0.001535		0.003231	
Delta A (′dev)	0.003322		0.003168		0.005558	
Metal diam = dm	2.350128		2.310102		2.040216	
Base Cost Computation Base = MG + GP + CE						
LENS #	1–2	3–4	5–6			
Milling = MG	17.40939	17.39075	17.2733			
Grind and polish = GP	39.57352	23.4	42.46667			
Center and edge = CE	7.497619	7.484286	7.394286			
Base cost total =	64.48053	48.27504	67.13426			
Computation of Total Lens Cost = MT						
Lens #	1–2		3–4		5–6	
Milling/generate	17.40939		17.39075		17.2733	
Part setup	21		18.2		18.2	
Grind and polish	35.68302	716.8573	15.47956	13.71988	21.21353	39.65565
Centering	586.8878		735.275		348.9406	
Total lens cost	1377.837		800.0652		445.2831	

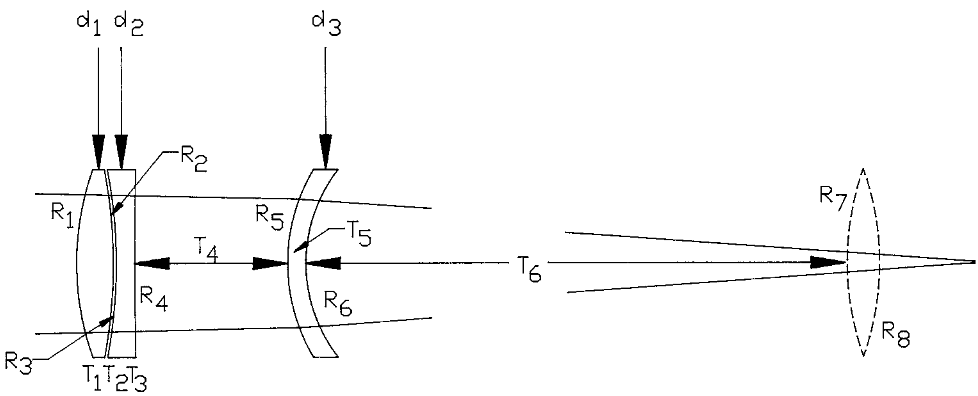

FIGURE 1.26 Focus lens set of telescope used as an example.

Tolerancing Summary

We had previously[11-13] shown the principles of how to assign tolerances to achieve a minimum production cost and we mentioned the possible application of the results to estimating total lens production cost. In this discussion, the previous data and principles have been refined and some of the results provided by Adams,[16] Smith,[20] Parks,[18] and Fischer[17] have been incorporated. The lens cost estimating formulas mentioned in the earlier work[13] have been developed into useful tools. A new analysis is presented of the interdependency of the lens and cell diameter tolerances as a result of the cost vs. tolerance knowledge. And finally, a minimum cost tolerancing procedure has been reduced to practice in a form which is straightforward and accurate enough for practical engineering application. It is practical to estimate the production cost of most lenses by entering the drawing data into a spreadsheet program. This is essentially an "expert system" estimator which can be applied by a person with very little training or experience to get as good or better cost estimates for most lenses than an expert. It is also now practical to distribute lens tolerances using a spreadsheet program such that the production costs are minimized and to find the impact of design changes on lens costs. Both of these tools have advanced what was an "art" to an engineering discipline but not quite a "science" because of the necessary simplifying assumptions. However, these tools are more accurate than what has been the common practice to date, and they are more accurate than the cost data as they typically are measured at this time. The application of the estimating program can reduce the business overhead cost of a production operation, and it can point to the cost drivers of any particular lens such as: setup, milling, polishing, or centering. The application of the tolerancing program can decrease the production cost of systems from what they typically have been in lens production, assembly, and testing. If the tolerances are unnecessarily tight, the lens production cost is wasteful. If the tolerances are not tight enough to give a good yield of deliverable parts, the assembly, rework, and testing costs are wasteful. This tool addresses the economic problem at the point of the most potential impact on the life cycle costs as shown in Figure 1.1, the detail design and analysis phase.

1.5 Environmental Effects

Laboratory optical instruments may operate in a relatively benign environment, but most others are subject to environmental effects which can be a major consideration in the design of an instrument. Astronomical instruments must operate satisfactorily over a broad range of nighttime temperatures, or daytime in the case of solar telescopes. Personal cameras and binoculars may experience a broad range of weather conditions. Military instruments probably have the most widespread and severe environmental exposure to temperatures, moisture, shock, vibration, dust, chemicals, etc. It is very relevant and important that all of the environmental requirements be properly dealt with in the optomechanical design and development process. Subsequent chapters will cover many of these aspects in more detail, but we will discuss a few things here and give some guidance with respect to available standards and specifications.

The U.S. government through its military procurement activities over the past century or more has evolved a very extensive set of specifications and test methods to ensure that the optical instruments (and anything else they buy) will perform as required in the expected environments. Most of the current U.S. optical industry has worked to these requirements as exemplified by MIL-STD-810.[27] This standard might be a worthwhile document to consult to check whether you have considered all of the possible effects that might be important to a given instrument under development and how to test it. For that same reason, we will also describe below the new international standards that exist or are in development.

Survivability under Temperature, Vibration, and Shock Loads

The quality of the image and pointing direction are two of the major performance factors of most optical instruments. These might also be referred to as resolution or modulation transfer function (MTF) and boresight. With a 35-mm camera, the MTF is the key factor and pointing is usually not an issue. In binoculars, one becomes concerned also with boresight so that the views through each eye are not uncomfortably divergent. Tilt and equality of magnification also are important in binoculars. A surveying instrument or military aiming sight places a great deal of emphasis on the repeatability of the line of sight or boresight.

The designer needs to be sure that the expected temperature changes, vibrations, and shocks will not damage or disable the instrument, but also will not degrade the image quality and boresight to unacceptable levels. For example, glass lenses in aluminum housings are at risk of becoming too loose at high temperatures and being "squeezed to breakage" at cold temperatures. This is, of course, due to the difference in the thermal coefficients of expansion (TCE). In the bonding of optical components, the TCE differences can cause major difficulties of distortion and glass fracture. Paul Yoder's chapter and others in this handbook touch on how to deal with some of these problems. Other materials can be used, but at penalties of weight, cost, and sometimes performance. In many instruments, thermal distortions of the shape of optical surfaces (particularly mirrors) can degrade the image quality severely. We have touched on some examples of how these factors are dealt with in the section on tolerancing, and later chapters will discuss many design techniques in some detail. The optomechanical designers' challenge is to develop an instrument which will survive and perform in all of the required environments.

Humidity, Corrosion, Contamination

Designing an instrument which will not degrade due to humidity, corrosion, and contamination poses another class of challenges. Whenever practical, instruments are sealed with an internal atmosphere such as dry nitrogen. This prevents internal condensation of moisture at low temperatures and optical coating degradations due to humidity. It also will keep out dust, contamination, and corrosive agents. A large astronomical telescope usually cannot be sealed. In such a case, the mirrors tend to become dusty and the coatings tend to degrade. The scattering of light by the contamination reduces the contrast (MTF) and can be disastrous when the telescope is used to look at faint objects while it is illuminated by bright objects. Handling contamination in large telescopes and designing to minimize its effects is a major specialty.

In sealed instruments, corrosion is only an issue for external surfaces and interfaces. The most difficult environment is usually salt fog. Unprotected metals such as aluminum and steel will deteriorate rapidly. Plating and/or painting is usually required. Unless specific treatments have been well tested before, it is highly recommended that samples be extensively tested for durability. Subsequent chapters contain some suggestions with respect to corrosion and contamination.

Environmental Testing Standards

To date there are two ISO standards dealing with the effect of environmental conditions on the performance of optical instruments, ISO 9022 — *Environmental Test Methods*, and ISO/CD 10109 — *Environmental Requirements*. ISO 9022 has been issued as a standard while ISO/CD 10109 is still being written. It is complete in outline form but the performance criteria for many instruments are still being defined.

ISO 9022 defines terms relating to environmental tests for optical instruments and for instruments that contain optical assemblies and components. In addition, it specifies the essential steps for conducting an environmental test and defines some 20 types of tests along with various subcategories of these tests.

ISO 10109 specifies the environmental requirements to be met regarding the reliability of particular optical instruments when exposed to various applicable environmental influences. It also defines the geographical and technological areas of applicability of the instruments. The standard does **not** apply to specifications for **packaging** for transportation and storage.

ISO 9022 — Environmental Test Methods

The 20 parts of ISO 9022 are listed in Table 1.6, along with an indication of how many subdivisions there are of the basic test. We do not list all the subdivisions but it is informative to see an example in Table 1.7 where the subdivisions of Parts 2 and 3 are listed.

TABLE 1.6 ISO 9022 — Environmental Test Methods

Part 1	Definitions, extent of testing
Part 2	Cold, heat, humidity — 7
Part 3	Mechanical stresses — 8
Part 4	Salt mist — 1
Part 5	Combined cold, low air pressure — 2
Part 6	Dust — 1
Part 7	Drip, rain — 3
Part 8	High pressure, low pressure, immersion — 3
Part 9	Solar radiation — 1
Part 10	Combined sinusoidal vibration, dry heat or cold — 2
Part 11	Mold growth — 1
Part 12	Contamination — 4
Part 13	Combined shock, bump or free fall, dry heat or cold — 6
Part 14	Dew, hoarfrost, and ice — 3
Part 15	Combined random vibration wide band — 2
Part 16	Combined bounce or steady state — 4 — acceleration, in dry heat or cold
Part 17	Combined contamination, solar radiation — 2
Part 18	Combined damp heat and low internal — 3 — pressure
Part 19	Temperature cycles combined with sinusoidal or random vibration — 3
Part 20	Humid atmosphere containing sulfur dioxide or hydrogen sulfide

TABLE 1.7

Part 2	**Cold, Heat, Humidity**
10	Cold
11	Dry heat
12	Damp heat
13	Condensed water
14	Slow temperature change
15	Rapid temperature change
16	Damp heat, cyclic
Part 3	**Mechanical Stresses**
30	Shock
31	Bump
32	Drop and topple
33	Free fall
34	Bounce
35	Steady-state acceleration
36	Sinusoidal vibration
37	Random vibration (wide band)

The beginning part of ISO 9022 defines basic terms relating to conducting these tests. For example, an **environmental test** is defined as a laboratory simulation of (usually severe) climatic, mechanical, and chemical influences likely to occur during transport, storage, and operation on a test specimen in order to quickly determine changes in the behavior of the specimen due to the influences. The act of subjecting the specimen to these influences is called **conditioning.**

Conditioning is considered to be the sum of external influences acting on the specimen during the test including the conditioning method (or particular environmental test), the degree of severity of the test, and the internal influences due to the state of operation such as motion and/or temperature change. Also defined are three **states of operation:** state 0 — in a storage or transportation container; state 1 — unprotected, ready for use, but not turned on; and state 2 — unprotected, turned on, and operating.

In order to evaluate what has happened during a test or conditioning, there are three types of tests or examinations. The first is simply a **visual examination** to see if, for example, some part of the specimen became loose during conditioning. The second is a **functional** test to see if the device still functions after conditioning. Finally there is the **measurement,** an objective determination of a physical quantity by comparison with a specified quantity.

A **test sequence** is defined as given in Table 1.8. While this table looks trivial, it does define precisely what is meant by a test sequence and that there are some important matters that must be noted before applying the conditioning so that changes may be recognized after the conditioning. In order to specify what test(s) are to be performed on a particular instrument, a one-line environmental test code is used for each type of test required. This test code is illustrated in Figure 1.27.

TABLE 1.8 Definitions — Test Sequence

1	Preconditioning — prepare specimen for testing
2	Initial test — state of device prior to testing
3	Conditioning — apply conditioning method at degree of severity and state of operation
4	Intermediate test — does it function in state 2
5	Recovery — bring back to ambient conditions
6	Final test — state of device after testing
7	Evaluation — determination if specifications met

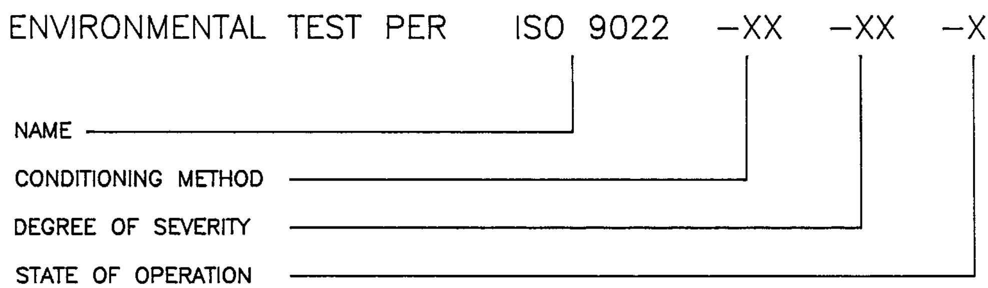

FIGURE 1.27 Definition of an environmental test code which can be communicated in a one line call-out under ISO 9022.

ISO 10109 — Environmental Requirements

Whereas ISO 9022 outlines the details and degrees of severity of nearly 100 conditioning methods, ISO 10109 is concerned with which of these methods and with what degree of severity they should be applied to a particular optical instrument designed for a particular type of use or for use in a particular climatic region.

ISO 10109 is an 11-part standard that specifies the types of testing needed to establish the suitability of an optical instrument for its intended conditions of use. The first part deals with definitions while the remaining ten parts are instrument-type specific sets of environmental requirements.

Part 1 starts by defining five **climatic zones** listed in Table 1.9. Zone 5 is a normal, protected environment in which most analytical instruments would be expected to operate. On the other hand, many types of optical instruments must survive rather severe climates and virtually all instruments must be unaffected by air shipment, much of the reason for zone 4, high altitudes. The severe climates can be thought of as Arctic (zone 1), Maritime (zone 3), and all other outdoor regions (zone 2).

The standard then goes on to define three instrument types: field instruments, instruments used in weather-protected locations, and these same instruments with the additional provision that they must be sterilizable. Also defined are ten groups of instruments along with subclasses as shown in Table 1.10. These begin with ophthalmic optics and cover the field to electro-optical systems.

TABLE 1.9 Definitions — Climatic zones

Zone 1	Not weather protected, cold and extremely cold
Zone 2	Not weather protected, global locations
Zone 3	Not weather protected, maritime locations
Zone 4	High altitudes to 30,000 m
Zone 5	Weather-protected, laboratory environment

TABLE 1.10 Definitions — Groups of Instruments

Group 1	Ophthalmic optics — 3 types
Group 2	Photographic instruments — 3 types
Group 3	Telescopes — 4 types
Group 4	Microscopes — 5 types
Group 5	Medical devices — 3 types
Group 6	Metrology instruments — 4 types
Group 7	Military instruments — 3 types
Group 8	Geodetic instruments — 3 types
Group 9	Photogrammetric instruments — 3 types
Group 10	Electro-optical systems — 3 types

Once this standard is issued, it will specifically define how well each of these instrument types must hold up under various environmental influences depending on the intended type of use the instrument will be subjected to in normal operation.

Summary of Environmental Effects

It can be seen that the optomechanical designer is usually required to give extensive thought during the design process to what environmental conditions will be encountered, and how to maintain the instrument performance under these circumstances. The available standards provide a good checklist to avoid overlooking a pertinent condition. They also provide guidance as to how the instrument might be tested to prove its performance.

REFERENCES

1. Smith, W. J. 1966. *Modern Optical Engineering*. McGraw-Hill, New York.
2. Foster, L. W. 1986. *Geo-metrics II*. Addison-Wesley, Reading, MA.
3. Wilson, B. A. 1992. *Design Dimensioning and Tolerancing*. Goodhart-Willcox, South Holland, IL. See also Gooldy, G. 1995. *Geometric Dimensioning & Tolerancing*, Prentice-Hall, Englewood Cliffs, NJ.
4. ANSI, 11 West 42nd St., New York, NY 11036, (212) 642-4900.
5. ASME, 345 East 47th St., New York, NY 10017, (212) 705-7000.
6. ASTM, 100 Barr Harbor Dr., West Conshohocken, PA 19428-2959, (610) 832-9585.
7. SAE, 400 Commonwealth Ave., Warrendale, PA 15096, (412) 776-4841.
8. OEOSC, 7913 B Fingerboard Rd., Frederick, MD 21704, (301) 607-4249.
9. SPIE's Homepage is at http://www.spie.org and OSA's at http://www.osa.org.
10. Optical Society of America, 2010 Massachusetts Ave., NW, Washington, D.C. 20036, (202) 223-8130.
11. Willey, R. R., George, R., Odell, J., and Nelson, W. 1982. Minimized cost through optimized tolerance distribution in optical assemblies. In *Optical Systems Engineering III, Proc. SPIE*, 389, Taylor, W. H., ed., p. 12. SPIE, Bellingham, WA.
12. Willey, R. R. 1983. Economics in optical design, analysis, and production. In *Optical System Design, Analysis, and Production, SPIE*, 399, Rodgers, P. J. and Fischer, R. E., eds., p. 371. SPIE, Bellingham, WA.

13. Willey, R. R. 1984. The impact of tight tolerances and other factors on the cost of optical components. In *Optical Systems Engineering IV, Proc. SPIE*, 518, Yoder, P. R., ed., p. 106. SPIE, Bellingham, WA.
14. Wiese, G. W., ed. 1991. *Optical Tolerancing*. SPIE, Bellingham, WA.
15. Plummer, J. L. 1979. Tolerancing for economies in mass production of optics. In *Contemporary Optical Systems & Component Specifications, Proc. SPIE*, 181, Fischer, R. E., ed., p. 90. SPIE, Bellingham, WA.
16. Adams, G. 1988. Selection of tolerances. In *Simulation and Modeling of Optical Systems, Proc. SPIE*, 892, Fischer, R. E. and O'Shea, D. C., eds., p. 173. SPIE, Bellingham, WA.
17. Fischer, R. E. 1990. Optimization of lens designer to manufacturer communications. In *International Lens Design Conference, Proc. SPIE*, 1354, Lawrence, G. N., ed., p. 506. SPIE, Bellingham, WA.
18. Parks, R. E. 1980. Optical component specifications. In *International Lens Design Conference, Proc. SPIE*, 237, Fischer, R. E., ed., p. 455. SPIE, Bellingham, WA.
19. Parks, R. E. 1983. Optical specifications and tolerances for large optics. In *Optical Specifications: Components & Systems, Proc. SPIE*, 406, Smith, W. J. and Fischer, R. E., eds., p. 98. SPIE, Bellingham, WA.
20. Smith, W. J. 1985. Fundamentals of establishing an optical tolerance budget. In *Geometrical Optics, Proc. SPIE*, 531, Fischer, R. E., Price, W. H., and Smith, W. J., eds., p. 196. SPIE, Bellingham, WA.
21. Willey, R. R. and Durham, M. E. 1992. Maximizing production yield and performance in optical instruments through effective design and tolerancing. In *Optomechanical Design, SPIE*, CR43, Yoder, P. R., ed., pp. 76–108. SPIE, Bellingham, WA.
22. Thorburn, E. K. 1980. Concepts and misconcepts in the design and fabrication of optical assemblies. In *Optomechanical Systems Design, Proc. SPIE*, 250, Bayar, M., ed., p. 2. SPIE, Bellingham, WA.
23. Beyeler, B. H. and Tiziani, H. J. 1966. Die optische Ubertragungsfunktion von dezentrierten optischen Systemen, *Optik*, 44(3), 317 (included in Ref. 4).
24. Kojima, T. 1979. Yield estimation for mass produced lenses based on computer simulation of the centering tolerance. In *Optical Systems Engineering, Proc. SPIE*, 193, Yoder, P. R., ed., p. 141. SPIE, Bellingham, WA.
25. McLaughlin, P. O. 1991. A primer on tolerance analysis, *Sinclair Optics Design Notes*, 2(3).
26. Willey, R. R. 1989. Optical design for manufacture. In *Recent Trends in Optical Systems Design II, SPIE*, 1049, Fischer, R. E. and Juergens, R. C., eds., p. 96. SPIE, Bellingham, WA.
27. MIL-STD-810, Environmental Test Methods.