Proc. of SPIE Vol. 0406, Optical Specifications: Components and Systems, ed. W.J. Smith, R.E. Fischer (Apr 1983) Copyright SPIE

Tolerances and Techniques in High Precision Optical Assembly

Eugene K. Thorburn

Optical Instruments Corp.
6450 Dale Street, Buena Park, California 90621

Abstract

There are excellent computer programs available for use by the lens designer who needs to establish tolerances for an optical system. The tolerances derived from the use of these programs can have considerably greater significance if the designer understands how the components are to be manufactured and how the system is to be mounted. The easily achieved tolerances can be tightened in order to permit loosening of those which are more difficult to come by. Selective assembly may be required in order to avoid setting tolerances at an economically unfeasible level.

Introduction

This paper will try to cover a number of important considerations in the tolerancing of optical assemblies. The way in which I have decided to approach the subject is from the "historical" perspective of my own 44 years in precision optical design and manufacturing. I hope that this approach will serve to illustrate not only the evolution of my understanding of the problems of tolerancing, but also some of the changes which have taken place in the industry as a whole.

The Simple Life 1939-1941

My first job in optics in 1939 was in the fabrication of astronomical telescopes for amateur astronomers. The tolerances were applied essentially to one or two components, namely the primary mirror and the diagonal mirror of the Newtonian, or the secondary mirror of the Cassegrainian. The concept of a tight tolerance was confined to the working surfaces only, one at a time. This was determined by Foucault knife edge testing at the center of curvature of the primary or in a Hindle test of the hyperbolic secondary. The bible was "Amateur Telescope Making," Volumes I and II. Wedge in any element was unimportant, since the alignment of each surface was made in assembly. Tape measure calibration of focal length was sufficient. Life was fairly simple.

Life Gets More Complex 1941-1944

The small company which had fabricated one mirror surface at a time was, early in 1941, suddenly awarded an Army contract for 3000 sets of 6 x 30 binocular components. The tolerancing of these components had been done by some "old hands" at Frankfort Arsenal. These engineers had obviously had hands-on experience in the fabrication and assembly of optical systems. They not only had put tolerances on their prints, but in addition they saw to it that each component supplier had a set of Arsenal manufactured test fixtures to check each of these components. Most fixtures consisted of a collimator with an illuminated resolution and alignment reticle plus one half of a binocular, except for that one missing element which was the one to be tested. To test the component it would be inserted into the fixture assembly in place of the missing part. If the total assembly met the requirements for resolution, path length and alignment, the part was deemed acceptable.

Let us look at one component in that assembly and see how it was toleranced and tested. The part in question was a 45° - 90° prism which made up one half of a type 1. Porro set. The top was cut down and the ends were rounded off, as shown in figure 1. An illuminated go-no go template fixture was supplied which checked the shaping of the ends and the mechanical length.

There was an angular deviation tolerance of 6 arc minutes. This was determined less by the effects of field tilt than by the ability to adjust the boresight of one side of the binocular to the other. In visual systems field tilt is rarely a problem and in actual practice the prism's angular deviation was never measured, only the lateral image displacements in both x and y. From the illustration (fig. 2) it can be seen that an image displacement due to angular error can be compensated for (in a convergent beam) by a longitudinal shift of the prism.

Image displacements perpendicular to the one shown can be achieved by an identical type of displacement of the second prism in the erecting set. Rotation of the prism around its circular end was used to line up the vertex of the 90° angle perpendicular to the desired displacement plane, and thus insure against image rotation.

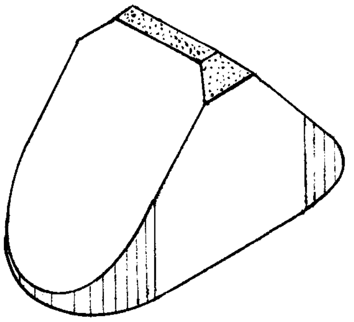

FIGURE 1. PORRO PRISM

The shop practice was to fabricate the prism to near the top of the tolerances for physical length and optical path. To inspect, it was put into the test fixture to see if the on-axis image fell within the tolerance circle. If it did not, it was turned end for end and was rechecked. If it again failed, the appropriate end could often be ground down enough (within the length tolerance) to shift the prism as in figure 2., and bring the image into the tolerance circle. If the minimum length was reached without reaching the tolerance circle, then one other heroic measure was available short of correcting angles. If the path length through the prism was sufficiently large, and the axial image was nearly into the tolerance circle, one surface could be reworked without changing the angles. This was equivalent to a lateral shift of one surface of the prism, producing a lateral image shift.

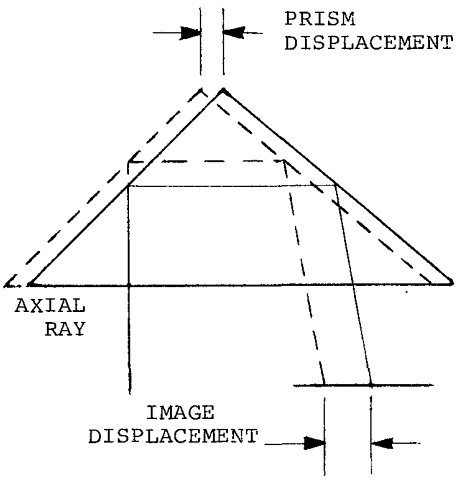

FIGURE 2. PRISM DISPLACEMENT

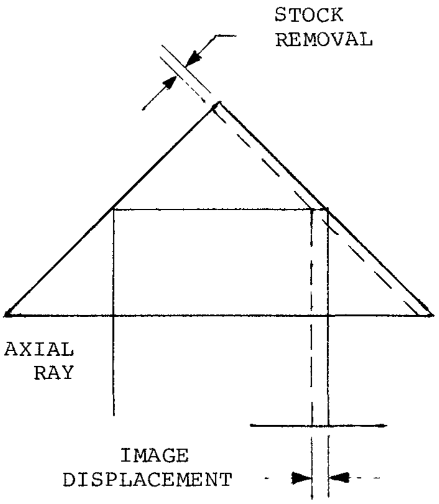

FIGURE 3. SURFACE REWORK

After passing inspection, the pivoting end of the prism which was held in contact with the fixture, was identified by a glass marking ink for future assembly.

This was my first encounter with the utilization of "technique" as opposed to total tolerancing of optical components for random assembly. Needless to say, the tolerances were broadened considerably and the rework time did not add up to the extra hours required to make a more precise part. The reduction in scrap was also important for the saving of raw materials.

I can defend this practice of trading image position by translation, in place of image position through angular accuracy, by a simple analysis. The major defect introduced into the binocular due to angular prism errors is image tilt. The WW II 6 x 30 binoculars used a Kellner type eyepiece, which for ease of computation we will assume to have had a focal length of 1.0 inch (10X). Let us assume that it was scaled from the Kellner shown as example 14.5 in Mil-Hdbk.-141. That example has a full field of 40°. Ignoring distortion, we can say that the real field diameter is 2 f tan 20° = .728 inch. Suppose we were to allow only ± ¼ diopter of defocus due to field tilt (½ diopter from edge to edge). The focal shift in inches is given by,

$$f = df^2/39.37$$

where:

f = focal shift in inches
f = focal length of eyepiece in inches
d = diopters of defocus

using:

$f = 1.0$
$d = \frac{1}{2}$

then,

$f = .0127$ inch (edge to edge)

The angular tilt corresponding to this amount of defocus is given by,

$$\tan \alpha = \frac{.0127}{.728} = .0174$$

$$\alpha = 1.0 \text{ degree}$$

How serious is a ¼ diopter defocus? Obviously, it can't be too serious, when a 65 year old observer can accommodate by a full ½ diopter. Another basis for comparison may be realized by noting that the Kellner design which we chose as our example has, at 20° off axis and at a focal length of 1.0 inch, tangential field curvature of 1.6 diopters and an astigmatic focal difference of 2.6 diopters.

It is reasonably obvious that the original tolerancing of the deviation limit of each prism at 6 arc minutes was made because of image position requirements rather than image degradation due to field tilt.

Microscope Objectives 1946-1950

My next experience with tolerancing vs. technique came when I was designing microscope objectives in the late forties. These objectives were turned out at the rate of over one thousand per month in spite of the fact that the equipment used to center the components both in and out of the assembly were totally inadequate. Also, the shop could not hold thickness tolerances of a small enough magnitude to insure that the quarter wave aberration acceptance critereon could be met with a random assembly of components. Selective assembly was used extensively. First, one airspace was selected as a variable to correct the spherical aberration introduced by the element thickness tolerance build-up. Second, selective rotation of the individual cells holding the separate components or doublets was used to reduce the coma caused by decentration. With these two adjustments the combined effects of spherical aberration and central coma on the focused wavefront were reduced to $\lambda/4$ or less. If the coma could not be brought within tolerance by this process, elements were substituted one by one until a suitable correction could be obtained. Those elements which had been replaced prior to achieving a good assembly, were rejected. At one point in a study of the problem, a precise photoelectric centration measuring device was constructed. With it, the design engineers were able to show that the distribution of centration error in the "rejected" elements was identical to the distribution in the untested elements directly from manufacturing. Obviously, the procedure was a "directed monkey on the keyboard" process using highly intelligent monkeys who typed out the "Gettysburg Address" with startling regularity.

Since that time, the same manufacturer has become a little smarter, and now one component is deliberately displaced to provide central coma to compensate that contributed by the rest of the decentered elements. This is still selective assembly, but with a more intelligent approach.

Reconnaissance Objectives 1953-1961

In working with the design and fabrication of reconnaissance lenses, some of the old problems remained or were intensified, but the solutions remained much the same. With the longer focal lengths and slower speeds, the effect of variations in thicknesses on spherical aberration became less of a problem. However, the greatly increased field angles meant that element thickness and airspace variations could affect off-axis aberrations very significantly. This was especially true of astigmatism. It was then that I learned that the classic method of "tweaking in" the astigmatism correction in a Double Gauss design was to vary the central airspace. Generally the ray angles in the axial bundle are quite small in the central stop space. A small change in that space will make only a very small difference in the way the axial bundle traverses the rear half of the lens. This means that a change in this airspace leaves spherical aberration and longitudinal color virtually unaffected. The off-axis angles are much larger in the central space than they are in object space and hence these off-axis rays will traverse the rear half in a significantly altered fashion, after even a small spacing change. Inasmuch as the third order astigmatism contribution of a surface is proportional to the square of the angle of incidence of the principal ray, that aberration is most strongly affected. Coma and lateral color vary linearly with the angle of incidence of the same ray, as does one of the two terms in the distortion contribution. This limits the change in these aberrations to an amount which (in those lenses) was usually tolerably.

Centration was, and still is, a big problem in the fabrication of reconnaissance lenses. In many cases, this was accentuated by having to supply the lens in two separate cells which were then threaded into a shutter by someone else, usually the camera manufacturer. If the front and rear pilots on the shutter were not concentric, then we had a lateral shift between cells. Those high angle off axis rays were displaced differently on the opposite sides of the axis in the rear cell. This resulted in a different distribution of coma and astigmatism in the two sides of the field, along with the introduction of ordinary field tilt.

Ideally, the best solution in the fabrication of a recon lens mount would be to machine both front and rear cells as a unit in a single chucking without ever removing the assembly from the lathe chuck. Spacers would be held in an expanding mandrel and machined from both sides in the same set-up, thus assuring parallelism. No hand deburring of seating surfaces would be permitted.

High Resolution Reduction Lenses (In the Sixties and Seventies)

I will start by defining a high resolution reduction lens as one exhibiting resolution 1000 line pairs per millimeter or better. One such lens, a 2.1 inch focal length, 0.4NA objective resolved 1600 lp/mm on a resist over a field of 5.0°. The tolerancing and methods of manufacture of this lens will be briefly described.

At the time the lens was designed and fabricated (1968), tolerancing programs were not nearly as sophisticated as they are today. Primarily, what was looked at was the change in third order aberrations with variations in lens parameters including radii thicknesses, airspaces and decentrations. It was determined that the permissible element wedges were in the range of 1 to 5 arc seconds, the thickness tolerances were in tenths of thousandths rather than thousandths of an inch and permissible index changes were correspondingly small.

Here again, the methods by which the precision was achieved was more important than the tolerances themselves. The design was revised for each change of melt, thus eliminating one set of tolerances. The lens elements were made to the best precision shop practice with tight 1-2 spherical fringe (and ¼ fringe irregularity) test plate fit, and ± .001 to ± .003 inch thickness tolerances. The thicknesses were then measured and the system with its actual thicknesses was reentered into the computer for reoptimation, using the airspaces as variables. Each assembly was customized in this manner to achieve an optimum performance, differing very little from the performance of the original design.

Achieving the tight centration tolerances was a different problem. No matter how closely the centration tolerances are held in a normal component, some clearance is required between the outside diameter of the element and the inside diameter of the bore in the mount, otherwise the element cannot be installed. This clearance permits an unwanted lateral shift of the element and in our case, even a few tenths of a thousandth of an inch clearance could permit an element decentration which produced intolerable image degradation. The selection of the appropriate mounting technique solved this problem.

A vertical mounting lathe was set up to permit final machining of the mount at assembly. Once assembly started, the mount was never removed from the lathe. This meant that each machined surface was parallel and/or concentric to every other surface within the runout of the lathe bearing. To insure that the runout of the bearing itself was minimized, we decided to use an air bearing. In 1968 good, large, stiff air bearings were not as readily available as they are today, so we designed and fabricated our own using optical lapping methods to produce accurate bearing surfaces. The resulting lathe was not particularly beautiful to behold, but it provided the necessary .000001 to .000002 inch maximum runout in both the radial and thrust directions and was stiff enough to give smooth machined surfaces on the mounting surfaces in the lens cells. It is shown in figure 4.

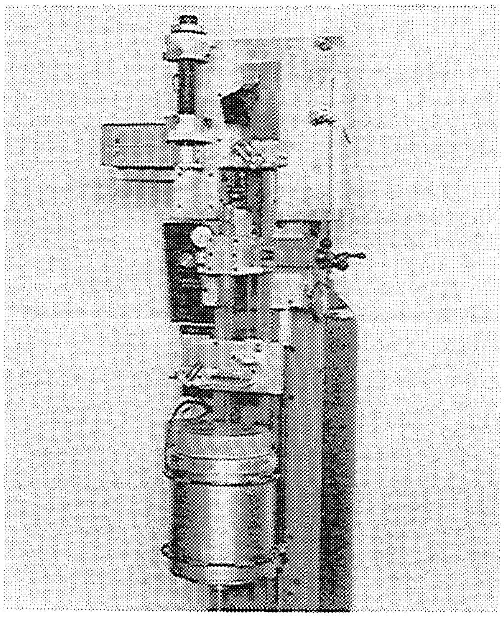

FIGURE 4. AIR BEARING LATHE

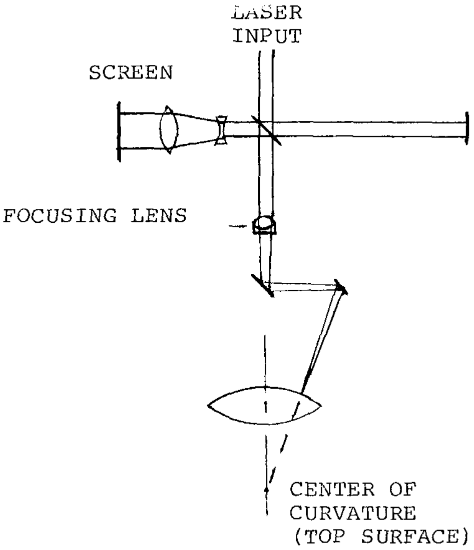

FIGURE 5. CENTRATION INTERFEROMETER

No attempt was made to retain the lenses by their outside diameters. Instead, the bores in the mount were deliberately made considerably oversize. The lens elements were precentered to a normally good tolerance (less than .0001 inch wedge). The seating surface for each lens was machined to a dimension measured with respect to the vertex of the previously installed element. This airspace dimension was tailored in each assembly to give the separation required by the redesign for element thicknesses. The element was then seated with the appropriate polished curve contacting the freshly machined surface in the cell. It was temporarily held by a retainer ring with four screws at 90° intervals around the periphery. This ring became the adjusting device by which the lens was centered in assembly. By rotating the lens about its spherical surface on a true-running seat, we could bring the top surface into alignment. Once aligned, the element was potted in place using a very low shrink epoxy. Once the epoxy was cured and the centration was rechecked, the operator repeated the procedure on the next element.

In order to achieve the necessary precision in centration, we measured the surface runout with a laser interferometer. We considered using a Fizeau type spherical interferometer for this purpose. We rejected this partly because it would require frequent changes in focusing optics to cover the large range of surface radii, but mainly because we wanted to also look through the element and measure the runout of the seated surface. Our experience had taught us that microscopic burrs and/or dust particles on the seat could cause that lower surface to run out well in excess of our tight tolerances. The interferometer we used is shown in figure 5. It is basically a Twyman-Green in which the measuring leg is focused onto the surface being checked, and is normal to that surface near the outside diameter. The f-number of the focused beam is very large, being of the order of f/250. The aberrations introduced into such a slow beam are very small even when going through the edge of a highly curved element. Even when these aberrations are visible, they do not bother our measuring precision, since we use the device like we would a dial indicator, measuring surface runout near the edge of the element. What we see when the air bearing is rotated is a sinusoidal variation in the position of fringes across the fringe field as the surface moves up and down. The high spot on the surface is easily recognized and the centering adjustments to reduce runouts down into the microinch region can be made in a surprisingly short time.

A substantial number of high performance lenses have been assembled by this process and have met the high resolution specifications on each occasion. For example, the ten element lens which resolved 1600 lp/mm was assembled in this way. In this lens, the maximum surface runout at the clear aperture of the worst decentered element was .000005 inch TIR.

Video Disc Recording Lenses 1973-1983

I am not going to tell you how we have solved all of the problems of manufacture of Disc Recording Objectives. Although we have been making objectives for this purpose since 1973, I am not sure that we have solved all of the problems yet. In fact, each new design presents new problems in manufacture. Like their cousins, the microscope objective, they require impossibly small tolerances to be held at all stages of manufacture. Unlike the microscope objective, the disc objective must also have a very light weight. The light weight is necessitated by the fact that the lens must be driven to follow focus on a rapidly rotating disc, and in some cases must also be translated to follow the recording track.

Generally, the numerical apertures run from 0.5 to 0.8 NA which is comparable to those of the higher power dry microscope objectives. The performance requirements are usually somewhat tighter, requiring Strehl ratios of around 0.9, compared to 0.8 for the microscope objective. The light weight requirement means that we can no longer use sturdy, stable brass cells with easy to make centering adjustments. The follow-focus requirement means that the elements must be securely fixed in their mounts so that they are not decentered by the constant vibration.

Let us look at some typical tolerances which might be obtained from a normal, well run tolerance analysis. Element thickness tolerances as small as ± .001 inch. Airspace tolerances as small as ± .0005 inch. Element wedge as large as 10 minutes and as small as 1.5 minutes (on a 0.2 inch diameter lens). This doesn't sound too bad till you find that most elements must be held in lateral centration to 0.0001 inch TIR, or 0.00005 inch lateral displacement! This means that the inside diameter of the bore in the mount must be no more than .0001 inch larger than the outside diameter of the lens. Since the bore must be larger than the lens to permit installation, unilateral diameter tolerances on the parts might be + .00000, - .00005 on the lens and + .00005, - .00000 on the mount with a zero diametral difference at the two nominal diameters. This would mean that nominal parts would have an interference fit, hardly a healthy situation to work with. Equally unhealthy would be the imposition of smaller tolerances which would allow, say, a .00003 inch difference between nominal dimensions along with unilateral tolerances of .000035 inch in the parts. Manufacturing such parts would be very difficult, and assembly would be equally difficult.

An alternative to this situation is to center the individual lenses to a wedge tolerance of 20 to 40 arc seconds instead of 1.5 to 10 arc minutes and to apply the gain in performance to relieving the lateral shift tolerance. This type of element centration can be achieved in edging using specially made high quality ball bearing spindles. A further gain could be made by air bearing centration and edging, reducing the wedge to the order of 2 to 5 arc seconds. The third alternative is to use the air bearing assembly technology discussed in the previous section of this paper.

None of these techniques can be applied with any quarantee of immediate success. However, if the proper considerations are given to the available trade-offs between tolerances and techniques, a solution can usually be found to the most difficult optical fabrication and assembly problems.

Conclusions (1983-?)

Most designers who have been in the optical business for many years have and experienced the many of the trade-offs between tolerances and manufacturing and assembly techniques. However, many new optical designers are entering the field today from other disciplines, without having had practical experience in the optical shop or the optical laboratory. Some of these highly intelligent people are designing good optical systems and then are tolerancing them in a way which makes them virtually impossible to fabricate, or at least terribly expensive. As an optical manufacturer, I can only hope that each new designer (and some of the old ones) will step away from the computer long enough to learn something about how optics are made. Then he may find out that there are shop techniques available which, when properly applied, can make his tightly toleranced design practical to fabricate.