
Chapter
15

Optics in Practice

This chapter will briefly survey the factors involved in reducing an optical system to practice. A short description of the optical manufacturing process will be followed by a discussion of the specification and tolerancing of optics for the shop. The mounting of optical elements will be considered next, and the chapter will be concluded with a section on optical laboratory measurement techniques.

15.1 Optical Manufacture

Materials. The starting point for quantity production of optics is most frequently a rough molded glass blank or pressing. This is made by heating a weighed chunk of glass to a plastic state and pressing it to the desired shape in a metal mold. The blank is made larger than the finished element to allow for the material which will be removed in processing; the amount removed must (at a minimum) be sufficient to clean up the outer layers of the blank which are of low quality and may contain flaws or the powdery fireclay used in molding. Typically a lens blank will be about 3-mm thicker than the finished lens and 2-mm larger in diameter. A prism blank will be large enough to allow removal of about 2 mm on each surface. These allowances vary with the size of the piece and are less for a clean blank. When the blank is of an expensive material, such as silicon or one of the more exotic glasses, the blanking allowances are held to the absolute minimum to conserve material.

Although most blanks are single, a cluster form is frequently economical for small elements. A cluster may consist of five or ten blanks connected by a thin web which is ground off to free the individual

blanks. If molded blanks are unobtainable, either because of the small quantity involved or the type of material, a rough blank may be prepared by chipping or sawing a suitable shape from stock material.

Rough blanks can be checked fairly satisfactorily for the presence of strain (which results from poor annealing of the glass) by the use of a polariscope. An accurate check of the index requires that a plano surface be polished on a sample piece; however, if a batch of blanks is known to have been made from a single melt or run of glass, only one or two of the blanks need be checked, because the index within a melt is quite consistent. Since the final annealing process raises the index, the presence of strain is frequently accompanied by a low index value.

When the shape of a blank is such that there are large variations in thickness from center to edge, it is difficult to get a uniform anneal. A variation of index within the blank may result. This is especially true for certain of the exotic optical glasses which are difficult to anneal. Glass in slab form is easier to anneal uniformly and is thus more homogeneous; it is often required for especially critical lenses for this reason.

Rough shaping. The preliminary shaping of an element is often accomplished by using diamond-charged grinding wheels. In the case of spherical surfaces, the process involved is *generating*. The blank is rotated in a vacuum chuck and is ground by a rotating annular diamond wheel whose axis is at an angle to the chuck axis, as indicated in Fig. 15.1. The geometry of this arrangement is such that a sphere is generated; the radius is determined by the angle between the two axes and by the effective diameter of the diamond tool (which will usually overhang the edge of the lens). The thickness is, of course, governed by how far the work is advanced into the tool. Flat work can be roughed

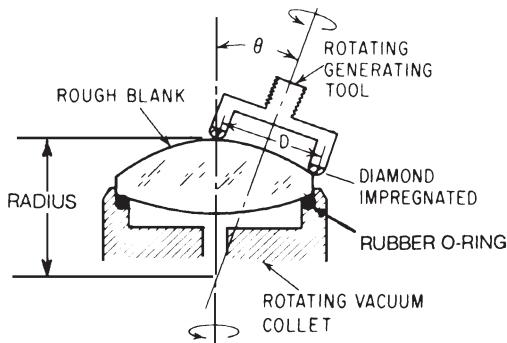


Figure 15.1 Schematic diagram of the generating process. The annular diamond tool and the glass blank are both rotated. Since their axes intersect at an angle (θ), the surface of the blank is generated to a sphere of radius $R = D/2 \sin \theta$.

out in a similar manner, with the two axes parallel. Rectangular shapes can be formed by milling, again using diamond tools.

Blocking. It is customary to process optical elements in multiples by fastening or blocking a suitable number on a common support. There are two primary reasons for this: The obvious reason is economy, in that several elements are processed simultaneously; the less apparent reason is that a better surface results when the processing is averaged over the larger area represented by a number of pieces.

The elements are fastened to the blocking tool with pitch, although various compounds of waxes and rosins are also used for special purposes. Pitch has the useful property of adhering tenaciously to almost anything which is hot and not sticking to cold surfaces. The pitch bond is readily broken by chilling the pitch to a brittle state and shocking it with a brisk but light tap. Typically the elements are fastened to the blocker by pitch buttons which are molded to the back of the elements (suitably warmed); the buttons are then stuck to the heated blocker, as indicated in Fig. 15.2. (The surfaces of the elements are maintained in alignment by placing the buttoned elements into a lay-in tool of the proper radius and then pressing the heated blocker into contact with the pitch buttons.)

The cost of processing an element is obviously closely related to the number of elements which can be blocked on a tool. There is no simple way to determine this number exactly; however, the following expressions (which are "limiting-case" expressions, modified to fit the actual values) are accurate to within about one element per tool.

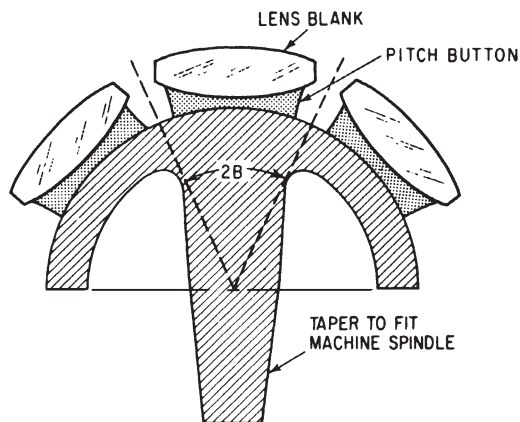


Figure 15.2 Section of a blocking tool with blanks fastened in place with buttons of blocking pitch. The maximum number of lenses that can be blocked on a tool is determined by the angle B (see Eq. 15.2).

For plano surfaces:

$$\text{No. per tool} = \frac{3}{4} \left(\frac{D_t}{d} \right)^2 - \frac{1}{2} \quad (15.1)$$

rounded downward to the nearest integer, where D_t is the diameter of the blocking tool and d is the effective diameter of the piece (and should include an allowance for clearance between the elements).

For spherical surfaces:

$$\text{No. per tool} = \frac{6R^2}{d^2} \left[\frac{\text{SH}}{R} \right] - \frac{1}{2} \quad (15.2a)$$

where R is the surface radius, d is the lens diameter (including a clearance allowance), and SH is the sagittal height of the tool. For a tool which subtends 180° , $\text{SH} = R$ and this reduces to

$$\text{No. per tool} = \frac{6R^2}{d^2} - \frac{1}{2} = \frac{1.5}{(\sin B)^2} - \frac{1}{2} \quad (15.2b)$$

rounded downward to the nearest integer, where B is the half-angle subtended by the lens diameter (plus spacing allowance) from the center of curvature of the surface, as indicated in Fig. 15.2.

Where there are only a few lenses per tool, the following tabulation for 180° tools is convenient and more accurate.

No. per tool	Maximum d/D_t	Maximum $\sin B$	$2B$
2	0.500	0.707	90°
3	0.462	0.655	81.79°
4	0.412	0.577	70.53°
5	0.372	0.507	60.89°
6	—	0.500	60°
7	0.332	0.447	53.13°
8	0.301	0.398	46.91°
9	0.276	0.383	45°
10	—	0.369	43.24°
11	0.253	0.358	41.88°
12	0.243	0.346	40.24°

Grinding. The surface of the element is further refined by a series of grinding operations, performed with loose abrasive in a water slurry and cast iron grinding tools. If the elements have not been generated, the grinding process begins with a coarse, fast-cutting emery. Otherwise, it begins with a medium grade and proceeds to a very fine grade which imparts a smooth velvety surface to the glass.

The grinding (and polishing) of a spherical surface is accomplished to a high degree of precision with relatively crude equipment by taking advantage of a unique property of a spherical surface, namely, that

a concave sphere and a convex sphere of the same radius will contact each other intimately regardless of their relative orientations. Thus, if two mating surfaces which are approximately spherical are contacted (with abrasive between them) and randomly moved with respect to each other, the general tendency is for both surfaces to wear away at their high spots and to approach a true spherical surface as they wear. (For a detailed analytical treatment of the subject of relative wear in optical processing, the reader is strongly urged to consult the reference by Deve, listed at the end of this chapter.)

Usually the convex piece (either blocker or grinding tool) is mounted in a power-driven spindle and the concave piece is placed on top as shown in Fig. 15.3. The upper tool is constrained only by a ball pin-and-socket arrangement and is free to rotate as driven by its sliding contact with the lower piece; it tends to assume the same angular rate of rotation as the lower piece. The pin is oscillated back and forth so that the relationship between the two tools is continuously varied. By adjusting the offset and amplitude of the motion of the pin, the optician can modify the pattern of wear on the glass and thus effect minute corrections to the value and uniformity of the radius generated by the process.

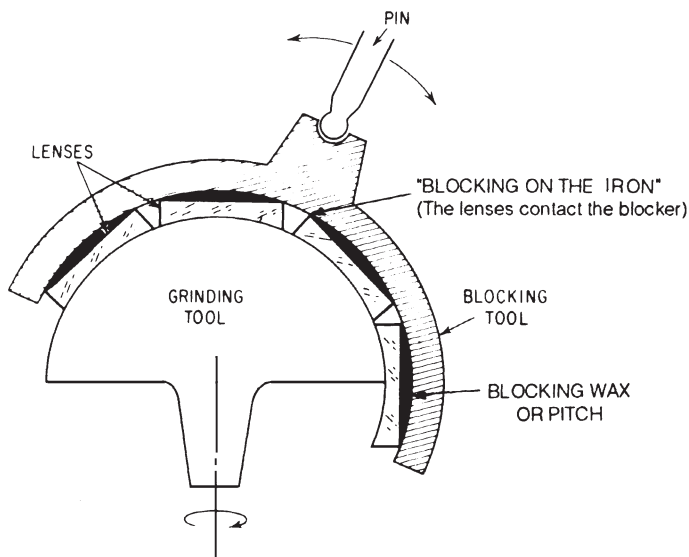


Figure 15.3 In grinding (or polishing), a semirandom scrubbing action is set up by the rotation of the lower (convex) tool about its axis and the back-and-forth oscillation of the upper (concave) tool. Note that the upper tool is free to rotate about the ball end of the driving pin and takes on a rotation induced by the lower tool.

Each successively finer grade of emery is used until the grinding pits left by the preceding operation are ground out.

Polishing. The mechanics of the polishing process are quite analogous to the grinding process. However, the polishing tool is lined with a layer of pitch and the polishing compound is a slurry of water and rouge (iron oxide) or cerium oxide. The polishing pitch will cold flow and thus take on the shape of the work in a very short time.

The polishing process is a peculiar one that is still incompletely understood. It appears that the surface of the glass is hydrolyzed by the polishing slurry and the resulting gel layer is scraped away by the particles of polishing compound embedded in the polishing pitch. This analysis explains many of the phenomena associated with polishing, such as scratches and cracks which are “flowed” shut by polishing, but which later open up when heated or exposed to the atmosphere. But when one considers that historically, polishing tools have been made from materials as diverse as felt, lead, taffeta, leather, wood, copper, and cork, and that polishing compounds other than rouge have been successfully used, and that many optical materials (e.g., silicon, germanium, aluminum, nickel, and crystals) have a different chemistry than glass, it would seem that a variety of polishing mechanisms is quite likely. Some polishing agents are actually etchants of the material that they polish; some materials can be polished dry.

Polishing is continued until the surface is free of any grinding pits or scratches. The accuracy of the radius is checked by the use of a test plate (or test glass). This is a very precisely made master gage which has been polished to an exact radius and which is a true sphere to within a tiny fraction of a wavelength. The test plate is placed in contact with the work, and the difference in shape is determined by the appearance of the interference fringes (Newton’s rings) formed between the two. The relative curvatures of the two surfaces can be determined by noting whether the gage contacts the work at the edge or the center. If the number of rings is counted, the difference between the two radii can be closely approximated from the formula

$$\Delta R \approx N\lambda \left(\frac{2R}{d} \right)^2 \quad (15.3)$$

where ΔR is the radius difference, N is the number of fringes, λ is the wavelength of the illumination, R is the radius of the test plate, and d is the diameter over which the measurement is made. One fringe indicates a change of one-half wavelength in the spacing between the two surfaces. A noncircular fringe pattern is an indication of an aspheric surface. An elliptical fringe indicates a toroidal surface.

Small corrections are made either by adjustment of the stroke of the polishing machine or by scraping away portions of the polishing tool so that the wear is concentrated on the portion of the work which is too high.

Centering. After both surfaces of an element are polished, the lens is centered. This is done by grinding the rim of the lens so that the mechanical axis (defined by the ground edge of the lens) coincides with the optical axis, which is the line between the centers of curvature of the two surfaces. In *visual* centering the element is fastened (with wax or pitch) to an accurately trued tubular tool on a rotating spindle. When the lens is pressed on the tool, the surface against the tool is automatically aligned with the tool and hence with the axis of rotation. While the pitch is still soft, the operator slides the lens laterally until the outer surface also runs true. If the lens is rotated slowly, any decentration of either surface is detectable as a movement of the reflected image (of a nearby target) formed by that surface, as indicated in Fig. 15.4. For high-precision work, the images may be viewed with a telescope or microscope to increase the operator's sensitivity to the image motion. The periphery of the lens is then ground to the desired diameter with a diamond-charged wheel. Bevels or protective chamfers are usually ground at this time.

For economical production of moderately precise optics, a mechanical centering process is used. In this method, called "cup" or "bell" centering, the lens element is gripped between two accurately trued tubular tools. The pressure of the tools causes the lens to slip sideways until the distance between the tools is at a minimum, thus centering the lens. The lens is then rotated against a diamond wheel to grind the diameter to size.

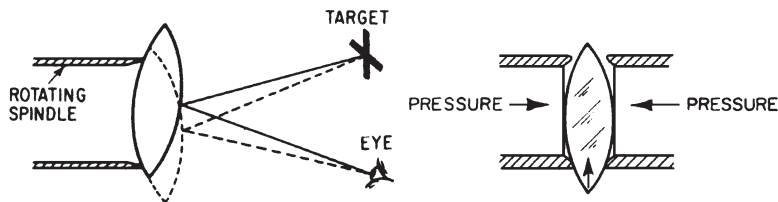


Figure 15.4 Left: In visual centering the lens is shifted laterally until no motion of the image of a target reflected from the lens surface can be detected as the lens is rotated. Right: In mechanical centering the lens is pressed between hollow cylinders. It slides laterally until its axis coincides with the common axis of the two tools.

The manufacture of the lens is completed by low-reflection coating the surfaces as required and by cementing, if the element is part of a compound component; these processes are outlined in Chap. 7.

Modifications of the standard processing techniques are sometimes required for unusual materials. Brittle materials (e.g., calcium fluoride) must be treated gently, especially in grinding. A finer, softer abrasive is required; sometimes soap is added to the abrasive and soft brass grinding tools are used in place of cast iron. At the other extreme, sapphire (Al_2O_3) cannot be processed with ordinary materials because of its extreme hardness, and diamond powder is used for both grinding and polishing.

Materials which are subject to attack by the grinding or polishing slurry are sometimes processed using a saturated solution of the optical material in the liquid of the slurry. For example, if a glass is attacked by water, one could make up the slurry with water in which a powder of the glass has been boiled or soaked for several days. Alternately a slurry of kerosene or oil sometimes works well. Other liquids which have been used in slurries include ethylene glycol, glycerol, and triacetate.

High-speed processing. For optics where the surface accuracy requirements are not high, the processes described above can be materially accelerated. Ordinary grinding usually takes tens of minutes. Polishing may take from 1 or 2 hours up to 8 or 10 hours in difficult cases. These operations can be speeded up by increasing both the speed of the spindle rotation and the pressure between work and tool. Tool wear and deformation are then a problem, so tools which are very resistant to change are used. Grinding is accomplished using tools faced with pellets or pads of diamond particles sintered in a metal matrix; loose abrasive is not used. This is called pellet grinding or pel-grinding. Polishing is done with a metal (typically aluminum) tool faced with a thin (0.01 to 0.02 in) layer of plastic (e.g., polyurethane). Processing times are to the order of minutes; a surface may be generated, ground, and polished in 5 or 10 minutes. Since the tools are not compliant, it is necessary that the radius from the generator have an exact relationship to the radius of the diamond pellet grinding tool, and that the ground radius match (to within a few fringes, à la Eq. 15.3) the radius required by the hard plastic polishing tool. This process is widely used for sunglasses, filters, inexpensive camera lenses, and the like. The surface geometry tends to be marginal as regards accuracy of figure, and the fixed-abrasive grinding does cause some subsurface fracturing, but the process is fast and economical. The tooling required and the fine-tuning adjustments of the steps of the process limit its application to large-quantity production.

Nonspherical surfaces. Aspherics, cylinders, and toroids do not share the universality of the spherical surface, and their manufacture is difficult. While a sphere is readily generated by a random grinding and polishing (because any line through the center is an axis), optical aspherics have only one axis of symmetry. Thus the simple principle of random scrubbing which generates a sphere must be replaced by other means. An ordinary spherical optical surface is a true sphere to within a few millionths of an inch. For aspherics this precision can only be obtained by a combination of exacting measurement and skilled “hand correction” or its equivalent.

Cylindrical surfaces of moderate radius can be generated by working the piece between centers (i.e., on a lathe). However, any irregularity in the process tends to produce grooves or rings in the surface. This can be counteracted by increasing the rate of working *along* the axis relative to the rate of rotation *about* the axis. It is difficult to avoid a small amount of taper (i.e., a conical surface) in working cylinders. Large-radius cylinders are difficult to swing between centers and are usually handled with an x - y rocking mechanism which constrains the axes of work and tool to parallelism so as to avoid a saddle surface.

Aspherics of rotation, such as paraboloids, ellipsoids, and the like, can be made in modest production quantities if the precision required of the surface is of a relatively low order, as, for example, in an eyepiece. The usual technique is to use a cam-guided grinding rig (with a diamond wheel) to generate the surface as precisely as possible. The problem is then to fine-grind and polish the surface without destroying its basic shape. The difficulty is that any random motion which works the surface uniformly tends to change the surface contour toward a spherical form. Extremely flexible tools which can follow the surface contour are required; however, their very flexibility tends to defeat their purpose, which is to smooth or average out small local irregularities left in the surface by the generating process. Pneumatic (i.e., air-filled, elastic) or spongy tools have proved quite successful for this purpose.

Where precise aspherics are required, “hand” or “differential” correction is practically a necessity. The surface is ground and polished as accurately as possible and is then measured. The measurement technique must be precise enough to detect and quantify the errors. For high-quality work, this means that the measurements must be able to indicate surface distortions of a fraction of a wavelength. The Foucault knife edge test and the Ronchi grating tests are widely used for this purpose; these tests can usually be applied directly to the aspheric surface, although there are many aspheric applications (e.g., the Schmidt corrector plate) where the test must be applied to the complete system to determine the errors in the aspheric.

When the surface is close to the required figure, it can be tested with an interferometer, just as a spherical surface on a lens is tested with a test plate (which is of course a simple interferometer). However, for a nonspherical surface some sort of arrangement is necessary to reshape the wave front reflected from the aspheric so that it matches the reference wave front of the interferometer. For a conic surface, auxiliary mirrors can be arranged so that the conic is imaging from focal point to focal point, and a perfect conic will then produce a perfectly spherical wave front. A more generally applicable approach is the use of a *null lens*, which is designed and very carefully constructed to distort the reflected wave front into an exactly spherical shape. For a paraboloid tested at its center of curvature, the null lens can be as simple as one or two plano-convex lenses whose undercorrected spherical cancels the overcorrection of the paraboloid. For general aspherics, the null lens may need to be quite complex.

When the surface errors have been measured and located, the surface is corrected by polishing away the areas which are too high. This can be accomplished (with a full-size polisher and a very short stroke) by scraping away those areas of the polisher which correspond to the low areas of the surface. In making a paraboloid of low aperture, such as used in a small astronomical telescope, the surface is close enough to a sphere that the correction can often be effected simply by modifying the stroke of the polisher. However, for large work and for difficult aspherics, it is usually better to use small or ring (annular) tools and to wear down the high zones by a direct attack. A certain amount of delicacy and finesse is required for this approach; if the process is continued for a minute or so longer than required, the result is a depressed ring which then requires that the entire balance of the surface be worn down to match this new low point.

A few companies have developed equipment which more or less automates this process. In one technique, a *computer-controlled polisher* uses a small polishing tool (or a tool consisting of three small tools which are driven to spin about their centroid) which is directed to dwell on the regions of the work which are high and need to be polished down. The location and dwell time are determined from interferograms of the surface, plus a knowledge of the wear pattern which the polishing tool produces. The use of a small, driven polisher means that the device is not limited to polishing annular zones on the work, and thus unsymmetrical surface errors can be efficiently corrected.

Another computer-controlled process is called *magnetorheologic* polishing. Here the polishing slurry includes a magnetic iron compound. The slurry is moved past the rotating lens, and at the lens a magnetic field causes the slurry to become stiff. This produces a localized polishing (or wearing) action on the surface. By rocking, spinning, and

advancing the lens into the moving slurry under computer control, the surface can be locally polished to achieve the desired surface figure. Again, an unsymmetrical figure error can be corrected by synchronizing the localized polishing action with the position of the lens.

Single-point diamond turning. Extremely accurate, numerically controlled lathes and milling machines are now available which are capable of generating both the finish and the precise geometry required for an optical surface. The cutting tool used is a single-crystal diamond, and the optic is machined as in a lathe or as with a fly-cutter in a mill. A single-point machining operation leaves tool marks—the finished surface is scalloped, and in some respects resembles a diffraction grating. This is one limitation of the process, and finished surfaces are often lightly “postpolished” to smooth out the turning marks. The more severe limitation is that only a few materials are suitable for machining, and unfortunately, optical glass is not one of them. However, several useful materials are turnable, including copper, nickel, aluminum, silicon, germanium, zinc selenide and sulfide, and, of course, plastics. Thus mirrors and infrared optics can be fabricated this way. Infrared optics do not require the same level of precision as do visual-wavelength optics, simply because a quarter-wave is almost 20 times larger at 10 μm than in the visible wavelengths. With this process an aspheric surface is just about as easy to make as a spherical surface. It has found significant acceptance in the infrared and military applications.

15.2 Optical Specifications and Tolerances

Many otherwise fully competent optical workers come to grief when it is necessary for them to send their designs to the shop for fabrication. The two most common difficulties are underspecification, in the sense of incompletely describing what is required, and overspecification, wherein tolerances are established which are much more severe than necessary.

Optical manufacture is an unusual process. If enough time and money are available, almost any degree of precision (that can be measured) can be attained. Thus, specifications must be determined on a dual basis: (1) the limits which are determined by the performance requirements of the system, and (2) the expenditure of time and money which is justified by the application. Note well that optical tolerances which represent an equal amount of difficulty to maintain may vary widely in magnitude. For example, it is not difficult to control the sphericity of a surface to one-tenth of a micrometer; the comparable (in terms of difficulty) tolerance for thickness is about 100 μm (0.1 mm),

three orders of magnitude larger. For this reason it is rare to find “box” tolerances in optical work; each dimension, or at least each class of dimension, is individually toleranced.

Every essential characteristic of an optical part should be spelled out in a clear and unambiguous way. Optical shops are accustomed to this, and if a specification is incomplete, either time must be wasted in questioning the specification to determine what the requirements are, or the shop must arbitrarily establish a tolerance. Either procedure is undesirable.

The following paragraphs are an attempt to provide a general guide to the specification of optics. The discussion will include the basis for the establishment of tolerances, the conventional methods of specifying desired characteristics, and an indication of what tolerances a typical shop may be expected to deliver.

The intelligent choice of specifications and tolerances for optical fabrications is an extremely profitable endeavor. The guiding philosophy in establishing tolerances should be to allow as large a tolerance as the requirement for satisfactory performance of the optical system will permit. Designs should be established with the aim of minimizing the effect produced by production variations of dimensions. Frequently, simple changes in mounting arrangements can be made which will materially reduce fabrication costs without detriment to the performance of a system. One should also be certain that the tightly specified dimensions of a system are the truly critical dimensions, so that time and money are not wasted in adhering to meaningless demands for accuracy.

Surface quality. The two major characteristics of an optical surface are its quality and its accuracy. *Accuracy* refers to the dimensional characteristics of a surface, i.e., the value and uniformity of the radius. *Quality* refers to the finish of the surface, and includes such defects as pits, scratches, incomplete or “gray” polish, stains, and the like. Quality is usually extended to similar defects within the element, such as bubbles or inclusions. In general (with the exception of incomplete polish which is almost never acceptable) these factors are merely cosmetic or “beauty defects” and may be treated as such. The percentage of light absorbed or scattered by such defects is usually a completely negligible fraction of the total radiation passing through the system. However, if the surface is in or near a focal plane, then the size of the defect must be considered relative to the size of the detail it may obscure in the image. Also, if a system is *especially* sensitive to stray radiation, such defects may assume a functional importance. In any case, one may evaluate the effect of a defect by comparing its area with that of the system clear aperture at the surface in question.

The standards of military specification MIL-O-13830 (now formally obsolete) are widely utilized in industry. The surface quality is specified by a number such as 80-50, in which the first two digits relate to the *apparent* width of a tolerable scratch and the second two indicate the diameter of a permissible dig, pit, or bubble in hundredths of a millimeter. Thus, a surface specification of 80-50 would permit a scratch of an *apparent* width which matched (by visual comparison) a #80 standard scratch and a pit of 0.5-mm diameter. The total length of all scratches and the number of pits are also limited by the specification. In practice, the size of a defect is judged by a visual comparison with a set of graded standard defects. Digs and pits can, of course, be readily measured with a microscope; unfortunately the apparent width of a scratch is not directly related to its physical size, and this portion of the specification is not as well founded as one might desire. The concept of a visual comparison with a standard is a good and efficient one.

McLeod and Sherwood, who originated this method of specifying surface quality, in their article describing it said that the number of a scratch was equal to the measured width in microns (micrometers) of a scratch made by a certain technique. Recently the government has used a relationship which indicates that the width in micrometers is only one-tenth of the scratch number. There is a widespread (and not unreasonable) suspicion that the widths of the standard scratches (which are maintained on physical pieces of glass) have become smaller in the decades since the 1940s (when the system originated.)

Surface qualities of 80-50 or coarser (i.e., larger) are relatively easily fabricated. Qualities of 60-40 and 40-30 command a small premium in cost. Surfaces with quality specifications of 40-20, 20-10, 10-5, or similar combinations require extremely careful processing, and the more critical are considerably more expensive to fabricate. Such specifications are usually reserved for field lenses, reticle blanks, or laser optics.

Surface accuracy. Surface accuracy is usually specified in terms of the wavelength of light from a sodium lamp (0.0005893 mm) or HeNe laser (0.0006328 mm). It is determined by an interferometric comparison of the surface with a test plate gage, by counting the number of (Newton's) rings or "fringes" and examining the regularity of the rings. As previously mentioned, the space between the surface of the work and the test plate changes one-half wavelength for each fringe. The accuracy of the fit between work and gage is described in terms of the number of fringes seen when the gage is placed in contact with the work.

Test plates are made truly flat or truly spherical to an accuracy of a small fraction of a fringe. Spherical test plates, however, have radii which are known to an accuracy only as good as the optical-mechanical means which are used to measure them. Thus the radius of a test

plate is frequently known only to an accuracy of about one part in a thousand or one part in ten thousand. Further, test plates are expensive (several hundred dollars per set) and are available as “stock tooling” only in discrete steps. Thus it frequently pays to enquire what radii the optical shop has as standard tooling.

The usual shop specification for surface accuracy is thus with respect to a *specific* test plate, and it takes the form of requiring that the piece must fit the gage within a certain number of rings and must be spherical (or flat in the case of plane surfaces) within a number of rings. A fit of from five to ten rings, with a sphericity (or “regularity”) of from one-half to one ring is not a difficult tolerance. Fits of from one to three rings with correspondingly better regularity can be achieved in large-scale production at a very modest increase in cost. Note that an irregularity of a small fraction of a ring is difficult to detect when the fit is poor. Thus, little is saved by specifying a ten-ring fit and a quarter-ring sphericity, since the fit must be considerably better than ten rings to be certain that the irregularity is less than one-quarter ring. The usual ratio is to have a fit of no worse than four or five times the maximum allowable irregularity. The change in radius due to a poor fit is frequently negligible in effect. For example, the radius difference between two (approximately) 50-mm radii at a 30-mm diameter which corresponds to five rings is (by Eq. 15.3) only about 33 μm .

The surface figure can be measured easily with an interferometer. While it is more difficult to control radius value with an interferometer than with a test plate, the interferometer is far superior when it comes to testing for sphericity or regularity. This is because the effective radius of the comparison wave front can be adjusted to match that of the surface under test, and also because the viewpoint of the interferometer is always normal to the surface and is thus not subject to the obliquity errors which afflict test plate readings.

If possible, one should avoid specifying accurate surfaces on pieces whose thickness-to-diameter ratio is low. Such elements tend to spring and warp in processing, and extreme precautions are necessary to hold an accurate surface figure. A common rule of thumb is to make the axial thickness at least one-tenth of the diameter for negative elements; where there is a good edge thickness, one-twentieth or one-thirtieth of the diameter is sometimes acceptable. For extremely precise work, especially on plane surfaces, the optician might prefer a thickness of one-fifth to one-third of the diameter.

The performance effects of errors in radius values (i.e., departures from the nominal design radii) are usually not too severe. In fact, it is the practice of some purchasers of optics *not* to indicate a tolerance on the specified radii, but to specify final performance in terms of focal length and resolution. It is usually possible for a well-tooled optical

shop to select judiciously (from its tooling list) nearby radii which produce a result equivalent to the nominal design. If tolerances are specified on radius values, one should bear in mind the fact that most effects produced by a radius variation are not proportional to ΔR , but to ΔC (or $\Delta R/R^2$). To take a simple example, we can differentiate the thin-lens focal-length equation

$$\phi = \frac{1}{f} = (n - 1) (C_1 - C_2) = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

with respect to the first surface to get the following:

$$\begin{aligned} d\phi &= (n - 1) dC_1 \\ df &= f^2 (n - 1) dC_1 = f^2 (n - 1) \frac{dR_1}{R_1^2} \end{aligned}$$

In a more complex system, the change in focal length resulting from a change in the i th curvature is approximated by

$$\begin{aligned} df &\approx \left(\frac{y_i}{y_1} \right) f^2 (n'_i - n_i) dc_i \\ df &\approx \left(\frac{y_i}{y_1} \right) f^2 (n'_i - n_i) \frac{dR_i}{R_i^2} \end{aligned}$$

The point is that if a uniform tolerance is to be established for all radii in a system, the uniform tolerance should be on curvature, *not* on radius. Therefore, radius tolerances should be proportional to the square of the radius. For example, given a lens with a radius of 1 in on one side and a radius of 10 in on the other, if we vary the 1-in radius by 0.001 in, the effect on the focal length is the same as a change of 0.100 in on the 10-in radius. If the second surface had a radius of 100 in, then the equivalent change would be about 10 in.

The preceding is, of course, based on focal-length considerations only. With regard to aberrations, it is difficult to generalize, since one surface of a system may be very effective in changing a given aberration while another may be totally ineffective. The relative sensitivity is determined by the heights of the axial and principal rays at the surface, the index break across the surface, and the angles of incidence at the surface. A good estimate of the effect that any tolerance has on the aberrations of a system can be determined by use of the third-order surface contribution equations of Chap. 10.

The effect of surface irregularity is more readily determined. Consider the case where the Newton's rings are not circular; this is an indication of axial astigmatism, since the power in one meridian is stronger than in the other. Here it is convenient to call on the Rayleigh

quarter-wave criterion. The OPD produced by a “bump” of height H on a surface is equal to $H(n' - n)$, or, expressing it in terms of interference rings (remembering that each fringe represents one-half wavelength change in surface contour),

$$\text{OPD} = \frac{1}{2} (\#FR) (n' - n) \text{ wavelengths}$$

where (#FR) is the number of fringes of irregularity.

Thus, to stay within the Rayleigh criterion, the total OPD, summed over the whole system, should not exceed one-fourth wavelength; this is expressed by the following inequality:

$$\Sigma (\#FR) (n' - n) < 0.5$$

Thus, a single element of index 1.5 could have one-half fringe of astigmatism (or any other surface irregularity) on both surfaces before the Rayleigh criterion was exceeded (assuming that the nominal correction was perfect and that the irregularities were additive).

Note that the expressions above do not take into account the fact that the system will probably be refocused to minimize the effects of any surface irregularity. See the discussion of OPD and spherical aberration in Sec. 11.3, for example. For astigmatism, refocusing reduces the OPD by a factor of 2.

Thickness. The effects of thickness and spacing variation on the performance of a system are readily analyzed, either by raytracing or by a third-order aberration analysis. The importance of thickness variation differs greatly from system to system. In the negative doublets of a Biotar (double-Gauss) objective, the thickness is extremely critical, especially as regards spherical aberration; for this reason the crown and flint elements are usually selected so that their *combined* thickness is very close to the design nominal. At the other extreme, the thickness variation of a planoconvex eyepiece element may be almost totally ignored, since it ordinarily has little or no effect on anything.

In general, thicknesses and spacings may be expected to be critical where the slope of the marginal axial ray is large. Anastigmats in general, and meniscus anastigmats in particular, are prone to this sensitivity. High-speed lenses, large-NA microscope objectives, and the like are usually sensitive.

Unfortunately the thickness of an optical element is not as readily controlled as some of the other characteristics. In production procedures where many elements are processed on the same block, the maintenance of a uniform nominal thickness requires precise blocking and tooling. The grinding operation, while precise enough in terms of radius, is difficult to control in terms of its extent. For close thickness

control, the generating operation must be accurate and each subsequent grinding stage must be exactly timed so that the proper finish, radius, and thickness are arrived at simultaneously.

A reasonable thickness tolerance for precise work is ± 0.1 mm (± 0.004 in). This can cause a shop some difficulty on certain lens shapes and on larger lenses; where a relaxation is possible, a tolerance of ± 0.15 or ± 0.2 mm is more economical. It is possible to hold ± 0.05 mm in large-scale production by taking care throughout the fabrication procedure. The rejection rate at this tolerance can become disastrous if the smallest mischance occurs. Of course it is possible, by hand-working and selection, to produce pieces to any desired tolerance level; the author has seen ± 0.01 mm held in moderate production quantities (although at rather immoderate cost).

Centering. The tolerances in centering are (1) on the diameter of the piece, and (2) on the accuracy of the centering of the optical axis with the mechanical axis. If the piece is to be centered (i.e., as a separate operation), the diameter can be held to a tolerance of plus nothing, minus 0.03 mm by ordinary techniques, and this is the standard tolerance in most shops. A small economy is effected by a more liberal tolerance. Tighter tolerances are possible, but are not often necessary for ordinary work.

The concentricity of an element is most conveniently specified by its *deviation*. This is the angle by which an element deviates an axial ray of light directed toward the mechanical center of the lens. The deviation angle is an especially useful measure of decentration, since the deviation of a group of elements is simply the (vector) sum of the deviations of the individual elements. Figure 15.5 is an exaggerated sketch of a decentered element. The optical and mechanical axes are shown separated by an amount Δ (the decentration). Since a ray parallel to the optical axis must pass through the focal point, the angular deviation δ in radians of the ray aimed along the mechanical axis is given by the decentration divided by the focal length.

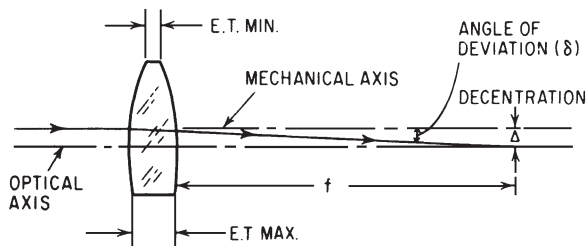


Figure 15.5 Showing the relationships between the optical and mechanical axes, and the decentration and angle of deviation in a decentered lens.

$$\delta = \frac{\Delta}{f} \quad (15.4)$$

Note that a decentered element may be regarded as a centered element plus a thin wedge of glass. The angle of the wedge W is given by the difference between the maximum and minimum edge thicknesses divided by the diameter of the element

$$W = \frac{E_{\max} - E_{\min}}{d} \text{ radians} \quad (15.5)$$

Since the deviation of a thin prism is given by $D = (n - 1)A$, we can similarly relate the wedge angle of an element to its deviation by

$$\delta = (n - 1) W \text{ radians} \quad (15.6)$$

If an element is centered on a high-production mechanical (clamping) centering machine, the limit on the accuracy of the concentricity obtained is determined by the residual difference in edge thickness which the cylindrical clamping tools cannot “squeeze out.” On most machines, this is to the order of 0.0005 in when residual tooling and spindle errors are also taken into account. Thus the residual wedge angle for a lens with a diameter d is given by

$$W = \frac{0.0005 \text{ in}}{d}$$

and the resulting deviation is

$$\delta = \frac{0.0005 \text{ in} (n - 1)}{d}$$

Thus, for ordinary lenses ($n = 1.5$ to 1.6) a reasonable estimate of the deviation is given by

$$\delta = \frac{1}{d} \text{ minutes} \quad (15.7)$$

where d is in inches, and the centering is done mechanically.

If the centering is accomplished visually (as indicated in the left-hand sketch of Fig. 15.4) then the ability of the eye to detect motion is the limiting factor. If we assume that the eye can detect an angular motion of 6 or 7×10^{-5} radians, then the deviation will be approximately

$$\delta = (n - 1) \left(\frac{1}{R} \pm 0.06 \right) \pm (\text{contact and spindle errors}) \quad (15.8)$$

where δ is in minutes and R is the radius of curvature of the outer surface in inches.*

*Equation 15.8 assumes that the image reflected from the outer radius is viewed at 10 in. This is obviously impossible if R is a convex surface with a radius longer than 20 in, and it is impractical if R is a long concave radius. Thus for $|R| > 20$ in, one should substitute 0.05 for the $1/R$ term.

The term $(n - 1)/R$ is from the visually undetected “wobble” of the outer radius and the $0.06(n - 1)$ term is due to the tilt in the tool which the eye could not detect in the truing of the tool (this is tested by pressing a flat glass plate against the rotating tool and observing any motion in the reflected image). The eye can, of course, be aided by means of a telescope or microscope which will further reduce the amount of decentration which can be detected by a factor equal to the magnification.

Occasionally lenses are not put through a separate centering operation. When this is the case, the concentricity of the finished lens is determined by the wedge angle which is left in by the grinding operations. If the blocking tooling is carefully worked out, it is possible to produce elements with a wedge (i.e., the difference between the edge thickness of opposite edges) to the order of 0.1 or 0.2 mm. Centering is often omitted on inexpensive camera lenses, condensers, magnifiers, or almost any single element of a simple optical system. Simple elements made from rounded circles of window glass are often left uncentered.

Prism dimensions and angles. The linear dimensions of prisms can be held to tolerances approximating those of an ordinary machined part, although the fabrication requirements of a prism are more difficult because of the finish and accuracy requirements of an optical surface. Thus tolerances of 0.1 or 0.2 mm are usually reasonable and tighter tolerances are possible.

Prism angles can be held to within 5 or 10 minutes of nominal by the use of reasonably good blocking forms. Indeed it is possible, although exceedingly difficult, to make angles accurate to a few percent of these tolerances if one takes exquisite pains with the design, fabrication, correction, and use of the blocking tools. Usually angles which must be held to tolerances of a few seconds (such as roof angles) are “hand corrected.” Such angles are checked with an autocollimator, either by comparison with a standard or by using the internal reflections to make the piece a retrodirector. Angles of 90° and 45° among others can be self-checked in this way since their internal reflections form constant-deviation systems of 180° deviation (as discussed in Chap. 4).

Prism size tolerances are usually based on the necessity to limit the image displacement errors (lateral or longitudinal) which they produce. Angular tolerances are usually established to control angular deviation errors. One can usually find one or two angles in a prism system which are more critical than the others; these can be tightly controlled and the other angles allowed to vary. For example, with respect to the deviation of a pentaprism, an angular error in the 45° angle between the reflecting faces is six times as critical as an error in the 90° angle between the entrance and exit faces, and the other two

angles have no effect on the deviation. On occasion, prism tolerances are based on aberration effects. Since a prism is equivalent to a plane parallel plate and introduces overcorrected spherical and chromatic; an increase in prism thickness in a nominally corrected system will overcorrect these aberrations. Some prism angle errors are equivalent to the introduction of thin-wedge prisms into the system. The angular spectral dispersion of a thin wedge is $(n - 1)W/V$ (where W is the wedge angle and V is the Abbe V -value of the glass) and the resultant axial lateral color may limit the allowable angular tolerances.

Materials. The characteristics of the refractive materials used in optical work which are of primary concern are index, dispersion, and transmission. For ordinary optical glass procured from a reputable source, visual transmission is rarely a problem. Occasionally, where a thick piece of dense glass is used in a critical application, transmission limits or color must be specified. Similarly, the dispersion, or V -value, is seldom a problem, except in special cases. For apochromatic systems where the partial dispersion ratio is exceedingly critical, very special precautions are required.

The index of refraction is usually of prime concern in optical glass. As indicated in Chap. 7, the standard index tolerance is ± 0.001 or ± 0.0015 , depending on the glass type. The glass supplier can hold the index more closely than this by selection or by extra care in the processing; either increases the cost somewhat. In practice the glass supplier will ordinarily use up only a fraction of this tolerance, since the index within a single melt or batch of glass is remarkably consistent. Thus, within a single lot of glass the index may vary only one in the fourth place. However, bear in mind that this variation *may* be centered about a value which is 0.001 or 0.0015 from the nominal index. It is sometimes economical to accept the standard tolerance and to adjust a design to compensate for the variation of a lot of glass in cases where the index is critical.

Transmission and spectral characteristics are often poorly specified. For filters and coatings, ambiguity can usually be avoided by specifying spectral reflection (or transmission) *graphically*, i.e., by indicating the area of the reflection (or transmission) versus wavelength plot within which the characteristics of the part must lie. One should also indicate whether or not the spectral characteristics outside the specified region are of importance. For example, in a bandpass filter, it is important to indicate how far into the long- and short-wavelength regions the blocking action of the filter must extend.

Figure 15.6 is a table of typical tolerances and may be used as a guide. Bear in mind, however, that the values given are *typical* and that there are many special cases that this sort of tabulation cannot cover.

	Surface Quality	Diameter, mm	Deviation (concentricity), min	Thickness, mm	Radius	Regularity (asphericity)	Linear Dimension, mm	Angles
Low cost	120-80	± 0.2	> 10	± 0.5	Gage	Gage	± 0.5	Degrees
Commercial	80-50	± 0.07	3-10	± 0.25	10 Fr	3 Fr	± 0.25	± 15'
Precision	60-40	± 0.02	1-3	± 0.1	5 Fr	1 Fr	± 0.1	± 5'-10'
Extraprecise	60-40	± 0.01	< 1	± 0.05	1 Fr	½ Fr	As req'd.	Seconds
Plastic	80-50		1	± 0.02	10 Fr	5 Fr	0.02	minutes

Figure 15.6 Tabulation of typical optical fabrication tolerances.

Additive tolerances. In analyzing an optical system to determine the tolerances to be applied to specific dimensions, one can readily calculate the partials of the system characteristics with respect to the dimensions under consideration. Thus, one obtains the value of the partial derivative of the focal length (for example) with respect to each thickness, spacing, curvature, and index; likewise for the other characteristics, which may include back focus, magnification, field coverage, as well as the aberrations or wave-front deformations. Then each dimensional tolerance, multiplied by the appropriate derivative, indicates the contribution of that tolerance to the variation of the characteristic. Now if it were necessary to be *absolutely* certain that (for example) the focal length did not vary more than a certain amount, one would be forced to establish the parameter tolerances so that the sum of the absolute values of the derivative-tolerance products did not exceed the allowable variance. Although this “worst-case” approach is occasionally necessary, one can frequently allow much larger tolerances by taking advantage of the laws of probability and statistical combination.

As a simple example, let us consider a stack of disks, each 0.1 in thick. We will assume that each disk is made to a tolerance of ± 0.005 in and that the probability of the thickness of the disk being any given value between 0.095 and 0.105 in is the same as the probability of its being any other value in this range. This situation is represented by the rectangular frequency distribution curve of Fig. 15.7a. Thus, for example, there is 1 in 10 chance that any given disk will have a thickness between 0.095 and 0.096 in. Now if we stack two disks, we know that it is *possible* for their combined thicknesses to range from 0.190 to 0.210 in. However, the *probability* of the combination having either of these extreme thickness values is quite low. Since the probability of either of the disks having a thickness between 0.095 and 0.096 is 1 in 10, if we randomly select two disks, the probability of *both* falling in this range is 1 in 100. Thus, the probability of a pair of disks having a thickness between 0.190 and 0.192 is 1 in 100; similarly for a combined thickness of 0.208 to 0.210 in. The probability of a combined thickness of 0.190 to 0.191 (or 0.209 to 0.210) is much less; 1 in 400.

The frequency distribution curve representing this situation is shown in Fig. 15.7b as a triangular distribution. Figure 15.7c shows frequency distribution curves for 1-, 2-, 4-, 8-, and 16-element assemblies. These curves have been normalized so that the area under each is the same and the extreme variations have been equalized. The important point here is that the probability of an assembly taking on an extreme value is tremendously reduced when the number of elements making up the assembly is increased. For example, in a stack of 16 disks with a nominal total thickness of 1.6 in and a possible

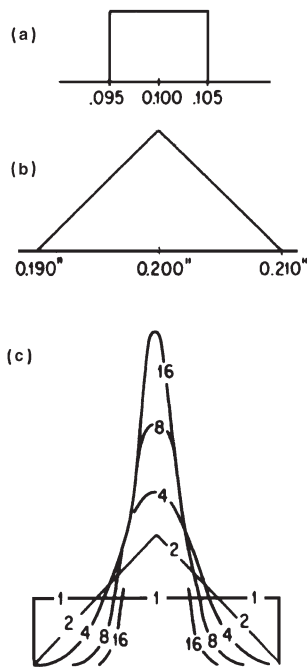


Figure 15.7 Showing the manner in which additive tolerances combine in assembly. Plot A shows a uniform probability in a dimension of a single piece. When two such pieces are combined, the resulting frequency distribution is shown in B. Normalized curves for assemblies of 1, 2, 4, 8, and 16 pieces are shown in C.

variation in thickness of ± 0.080 in, the probability of a random stack having a thickness less than 1.568 in or more than 1.632 in (i.e., ± 0.032 in) is less than 1 in 100.

The importance of this in setting tolerances is immediately apparent. In the stacked-disks example, if the range of thicknesses represented by 1.568 to 1.632 in for 16 disks were the greatest variation that could be tolerated, we could be absolutely sure of meeting this requirement *only* by tolerancing each individual disk at ± 0.002 in. However, if we were willing to accept a rejection rate of 1 percent in large-scale production, we could set the thickness tolerance at ± 0.005 in. If the cost of the pieces made to the tighter tolerance exceeded the cost of the pieces made to the looser tolerance by as little as 1 percent (plus one sixteen-hundredth of the assembly, processing, and final inspection costs), the looser tolerance would result in a less costly product.

In a frequency distribution curve such as those shown in Fig. 15.7 the area under the curve between two abscissa values represents the (relative) number of pieces which will fall between the two abscissa values. Thus the probability of a characteristic falling between two values is the area under the curve between the two abscissas divided by the total area under the curve.

The “peaking-up” characteristic of multiple assemblies can also be represented by the two plots shown in Fig. 15.8. The graph on the left shows the percentage of assemblies which fall within a given central fraction of the total tolerance range as a function of that fraction. The number of elements per assembly is indicated on each curve. These curves were derived from Fig. 15.7c. The graph on the right in Fig. 15.8 is simply another way of presenting the same data. If one were interested in an assembly of 10 elements, the intersection of the abscissa corresponding to 10 and the appropriate curve would indicate that all but 0.2 percent (using the 99.8 percent curve) of the assemblies would fall within 0.55 of the total tolerance range represented by the sum of all 10 tolerances, and that over one-half of the assemblies (using the 50 percent curve) would fall within 0.15 of the total possible range.

The preceding discussion has been based upon the unlikely assumptions that (1) each individual piece had a rectangular frequency distribution, and (2) each tolerance was equal in effect. This is rarely true in practice. The frequency distribution will, of course, depend on the techniques and controls used in fabricating the part, and the tolerance sizes may represent the partial derivative tolerance products from such diverse sources as tolerances on index, thickness, spacing, and curvature. Note, however, that in Fig. 15.7c the progression of curves may be started at any point. If, for example, the production methods produce a triangular distribution (such as that shown for an assembly of two elements), then the curve marked 4 (for “four elements”) will be the frequency distribution for two elements (of triangular distribution) and so on.

Note also that as more and more elements are included in the assembly, the curve becomes a closer and closer approximation to the normal distribution curve which is so useful in statistical analysis

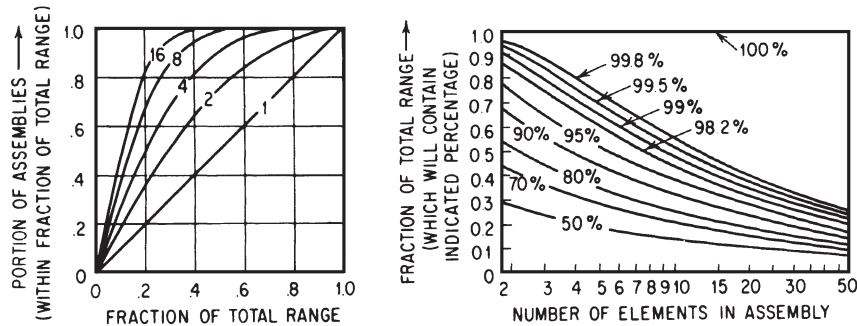


Figure 15.8 Probability distributions of additive tolerances in multiple assemblies. See text for details.

(except that the tolerance-type curves do not go to infinity as do normal curves). One useful property of the normal curve for an additive assembly is that its “peakedness” is proportional to the square root of the number of elements in assembly. Thus if 99 percent of the individual pieces are expected to fall within some given range, then for an assembly of 16 elements, 99 percent would be expected to fall within $\sqrt{1/16}$, or one-quarter of the total range. A brief examination will indicate that even the rectangular distribution assumed for Figs. 15.7 and 15.8 tends to follow this rule when there are more than a few elements in the assembly.

A rule of thumb frequently used to establish tolerances may be represented as follows:

$$T \approx \sqrt{\sum_{i=1}^n t_i^2} \quad (15.9)$$

This is frequently referred to as the RSS rule, shorthand for the square root of the sum of the squares. What the RSS rule means is this: If some percentage (say 99 percent) of the part tolerances produces effects less than t (and varies according to a normal, or gaussian, distribution), then the same percentage (i.e., 99 percent in our example) of the assemblies will show a total tolerance effect less than T .

While this section may seem to be a far cry from optical engineering, consider that a simple Cooke triplet has the following dimensions which affect its focal length and aberrations: six curvatures, three thicknesses, two spacings, three indices, and three V -values. These total fourteen for monochromatic characteristics and seventeen for chromatic aberrations. Such a system is eminently qualified for statistical treatment. *Note that the validity of this approach does not depend on a large production quantity; it depends on a random combination of a certain number of tolerance effects.*

There are two obvious features of the RSS rule which are well worth noting. One is the square root effect: If you have n tolerance effects of a size $\pm x$, then the RSS rule says that a random combination will produce an effect equal to $\pm x$ times the square root of n . For example, given 16 tolerance effects of ± 1 mm, we should expect a variation of only ± 4 mm, not ± 16 mm. The other feature is that the larger effects dominate the combination. As an example, consider a case with nine tolerances of ± 1 mm and one tolerance of ± 10 mm. If we use the RSS rule on this, we get an expected variation equal to the square root of 109, or ± 10.44 mm. Compare this with the fact that the single ± 10 -mm tolerance has an RSS of ± 10 mm. The addition of the nine ± 1 -mm tolerances changed the expected variation by only 4.4 percent.

One possible way to establish a tolerance budget using this principle is as follows:

1. Calculate the partial derivatives of the aberrations with respect to the fabrication tolerances (radius, asphericity, thickness and spacing, index, homogeneity, surface tilt, etc.). Express the aberrations as OPDs (wave-front deformation).
2. Select a preliminary tolerance budget. Figure 15.6 can be used as a guide to appropriate tolerance values.
3. Multiply the individual tolerances by the partial derivatives calculated in step 1.
4. Compute RSS for all the aberrations for each individual tolerance. This will indicate the relative sensitivity of each tolerance.
5. Compute RSS for all of the effects calculated in step 4 combined.
6. Compare the results of step 5 with the performance required of the system. This can be done by computing the RSS for the design OPD (as indicated by its MTF or whatever measure is convenient) combined with the tolerance budget OPD and using the material of Chap. 11 to determine the resulting MTF or Strehl ratio.
7. Adjust the tolerance budget so that the result of step 6 is equal to the required performance. Since the larger effects dominate the RSS, if you are tightening the tolerances (as is quite likely on the first go-round), you should tighten the most sensitive ones (and possibly loosen the least sensitive). Note that there is no economic gain if you loosen tolerances beyond the level at which costs or prices cease to go down. Conversely, one should be sure that the tolerances are not tightened beyond a level at which fabrication becomes impossible—since cost rises asymptotically toward infinity as this level is approached.
8. After one or two adjustments (steps 2 through 7) the tolerance budget should converge to one which is reasonable economically and which will produce an acceptable product.

If the tolerances necessary to get an acceptable performance are too tight to be fabricated economically, there are several ways which are commonly used to ease the situation:

1. A *test plate fit* is a redesign of the system using the measured values of the radii of existing test plates. This eliminates the radius tolerance (except for the variations due to the test glass “fit” in the shop, and any error in the measurement of the radius.)

2. A *melt fit* can effectively eliminate the effects of index and dispersion variation. Again, this is a redesign, using the measured index of the actual piece of glass to be used, instead of the catalog values.
3. A *thickness fit* uses the measured thicknesses of the actual fabricated elements to be assembled; this amounts to an adjustment of the airspaces during the assembly process.

The redesigns called for in all three “fitting” operations above, while hardly trivial, are not major undertakings when an automatic lens design program is used.

While the above may tend to induce a desirable relaxation in tolerances, one or two words of caution are in order. As previously mentioned, the index of refraction distribution within a melt or lot of glass may or may not be centered about the nominal value. When it is centered about a nonnominal value, the preceding analysis is valid only with respect to the central value, not the nominal value. Further, in some optical shops, there is a tendency to make lens elements to the high side of the thickness tolerance; this allows scratched surfaces to be reprocessed and will, of course, upset the theoretical probabilities. Another tendency is for polishers to try for a “hollow” test glass fit, i.e., one in which there is a convex air lens between the test plate and the work. This is done because a block of lenses which is polished “over” is difficult to bring back. Surprisingly, these nonnormal distributions have very little effect on Eq. 15.9 (if there are enough elements in the assembly).

Thus, the situation is seen to be a complex one, but nonetheless one in which a little careful thought in relaxing tolerances to the greatest allowable extent can pay handsome dividends. For those who wish to avoid the labor of a detailed analysis, the use of Eq. 15.9, or even the assumption that the tolerance buildup will not exceed one-half or one-third of the possible maximum variation, are fairly safe procedures in assemblies of more than a few elements. Above all, when cost is important, one should try to establish tolerances which are readily held by normal shop practices.

15.3 Optical Mounting Techniques

General. In optical systems, just as in precise mechanical devices, it is best to observe the basic principles of kinematics. A body in space has six degrees of freedom (or ways in which it may move). These are translation along the three rectangular coordinate axes and rotation about these three axes. A body is fully constrained when each of these possible movements is *singly* prevented from occurring. If one of

these motions is inhibited by more than one mechanism, then the body is overconstrained and one of two conditions occurs; either all but one of the (multiple) constraints are ineffective or the body (and/or the constraint) is deformed by the multiple constraint.

The laboratory mount indicated in Fig. 15.9 is a classical example of a kinematic mount. Here it is desired to uniquely locate the upper piece with respect to the lower plate. At A the ball-ended rod fits into a conical depression in the plate. This (in combination with gravity or a springlike pressure at D) constrains the piece from any lateral translations. The V -groove at B eliminates two rotations, that about a vertical axis at A and that about the axis AC . The contact between the ball end and the plate at C eliminates the final rotation (about axis AB). Note that there are no extra constraints and that there are no critical tolerances. The distances AB , BC , and CA can vary widely without introducing any binding effects. There is one unique position which will be taken by the piece; the piece may be removed and replaced and will always assume exactly the same position.

A perfectly kinematic system is frequently undesirable in practice and semikinematic methods are often used. These substitute small-area contacts for the point and line contacts of a pure kinematic mount. This is necessary for two reasons. Materials are often not rigid enough to withstand point contact without deformation, and the wear on a point contact soon reduces it to an area contact in any case.

Thus, in the design of an instrument, optical or otherwise, it is best to start by defining the degrees of freedom to be allowed and the degrees of constraint to be imposed. These can be outlined first by geometrical points and axes and then reduced to practical pads, bearings, and the like. This sort of approach results in a thorough and clear understanding of the effects of manufacturing tolerances on the function of the device and often indicates relatively inexpensive and simple methods by which a high order of precision can be maintained.

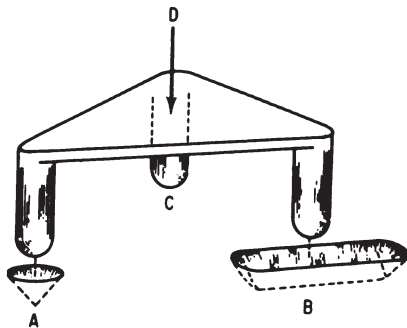


Figure 15.9 An example of a kinematic locating fixture. The three ball-ended legs of the stool rest in a conical hole at A , a V -groove (aligned with A) at B , and on a flat surface at C .

Lens mounts. Optical lens elements are almost always mounted in a close-fitting sleeve. A number of methods are used to retain the element in the mount; several are sketched in Fig. 15.10. In sketches (a) and (b) the lenses are retained by spring rings. In the left-hand mount (a), the spring catches in a V-groove, and if the mount is properly executed, the spring wire (which in its free state assumes a larger diameter) presses against the face of the element and the outer face of the groove. The lens is thus under a light pressure. The flat spring retainer (b) is less satisfactory, since the retainer will readily slip out unless the spring is strong or has sharp edges which bite into the mount. Other methods suitable for retaining low-precision elements include staking or upsetting ears of metal from the cell which clamp a thin metal washer over the lens element. Condenser systems are often mounted between three rods which are grooved as indicated in Fig. 15.10c. This provides a loose mount which leaves the condenser elements free to expand with the heat from the projection lamp without being constricted by the mount; it also allows cooling air to circulate freely. Both points are especially important in the mounting of a heat absorbing filter.

Where precision is required, the cell is fitted rather closely to the lens. For good-quality optics the lens diameter may be toleranced $+0.000$, -0.001 in and the inside cell diameter toleranced $+0.001$, -0.000 in with 0.001 - or 0.0005 -in clearance between the nominal diameters. For small lenses which demand high precision, these tolerances can be halved, at the expense of some difficulty in production.

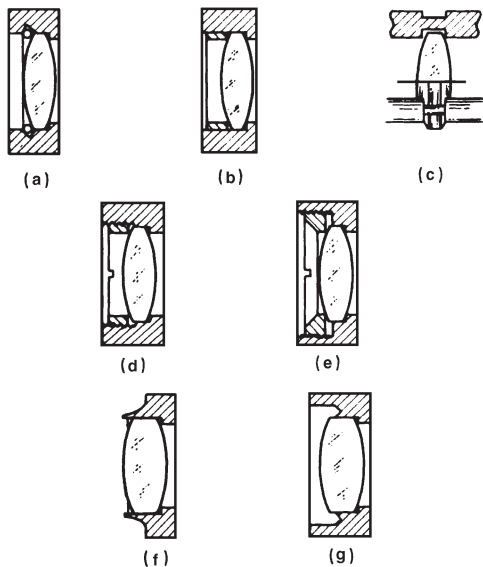


Figure 15.10 Several methods of retaining optical elements. (a) Wire spring ring in a V-groove; (b) flat spring ring; (c) three grooved rods at 120° ; (d) and (e) threaded lock ring; (f) spinning shoulder, before burnishing and (dotted) after; (g) cemented in place with trough for cement overflow.

Large-diameter optics are usually specified to somewhat looser tolerances. The lenses are most commonly retained by a threaded lock ring, as indicated in Fig. 15.10d or e. Sometimes the lock ring has an unthreaded pilot whose diameter is the same as the lens in order to be certain that the lens will ride on the bored seat and not on the threads. A separate spacer may be substituted for the pilot. The fit of the threaded parts should be loose so that the lens takes its orientation from the seat and shoulder, rather than from the threaded lock ring which frequently cocks.

A lens may be spun into the mount, as shown in Fig. 15.10f. In this method the mount is made with a thin spinning shoulder which protrudes past the edge of the lens (which is preferably beveled). This spinning shoulder is a few thousandths of an inch thick at the outside edge and has an included angle of 10 or 20°. The lens is inserted and the thin lip is turned over, usually by rotating the cell while the lip is bent over. Care and skill are required, but there are a number of advantages to this technique. The pressure of the spinning shoulder tends to center the lens in the mount. In assemblies requiring extreme precision, the seat can be bored to fit the lens diameter and the lens can be spun in place without removing the piece from the lathe; the result is concentricity of an order which is difficult to duplicate by any other means.

Another technique which results in both economy and precision is to cement the lens into its seat. The cement has a modest centering action, and with a good plastic cement the lens is securely retained. Care should be taken to provide an overflow groove (Fig. 15.10g) so that excess cement is kept away from the surface of the lens.

For optics which must withstand a difficult thermal and/or vibration environment, a useful form of mount is achieved by making the inside diameter of the cell oversize and cementing the element in place with a compliant, elastomeric RTV type of cement. The lens is trued in the mount before it is cemented in place. This technique is especially useful for large-diameter elements where the thermal expansion difference between the element and the mount is a serious problem; the layer of RTV between the element and the cell is made thick enough to take up the expansion difference.

In an assembly where several lenses and spacers are retained by a single lock ring, care must be taken that the thickness tolerances on the lenses and spacers are not allowed to build up to a point where the outside lens (1) extends beyond its seat and is not constrained by the seat diameter, or (2) is down into the mount so far that the lock ring cannot seat down on it. Another point to watch is that the mouth of a long inside-diameter bore is frequently bell-shaped, and a lens located near the mouth may have several thousandths of an inch more lateral

(diametral) freedom than intended. In critical assemblies, it frequently pays to locate the lens well inside the mouth of the bore.

When elements of different diameters are to be mounted together, the mount can be designed so that the lens seats can all be bored in one operation. This not only tends to reduce the cost of the mount but eliminates a possible source of decentration of each element with respect to the others, which can occur when the lens seats are bored in two or more separate operations.

The microscope style of element mounting shown in Fig. 15.11 illustrates a number of valuable devices. The lens seat and the outside support diameter of each cell can be turned in the same operation; indeed, in a critical system, the optical element may be spun in place without removing the piece from the lathe. (Cementing the lenses in place can be substituted for spinning.) All the cells are seated in the same bore of the main mount and they are isolated from the lock-ring threads (not shown) by a long spacer. All these techniques contribute to maintaining the exquisite concentricity necessary in a first-class microscope objective.

In mounting any type of optical element, it is important to avoid any warping or twisting. In the case of lens elements (which are in effect clamped between a shoulder and a lock ring, or their equivalents), this is not too difficult, since the pressure points are opposite each other and result in compression of the lens. More care is necessary in mounting mirrors and prisms, however, since it is quite easy to make the mistake of restraining a mirror in such a way that its surface is warped out of shape. One way to avoid this is to be sure that for each point at which pressure is exerted, there is a pad directly opposite so that no twisting moment is introduced.

Figure 15.12 serves as an indication of how few constraints are necessary to kinematically define the location of a piece. This illustration might apply to a piece of cubical shape. The three points in the XZ plane define a plane on which the lower face of the piece rests; these points take up one translational and two rotational degrees of freedom. The two points in the YZ plane take up one translation and one

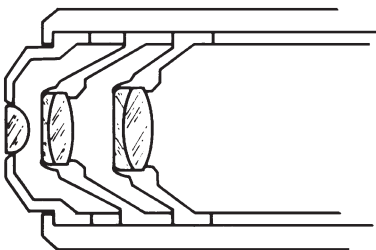


Figure 15.11 Mounting detail, microscope objective.

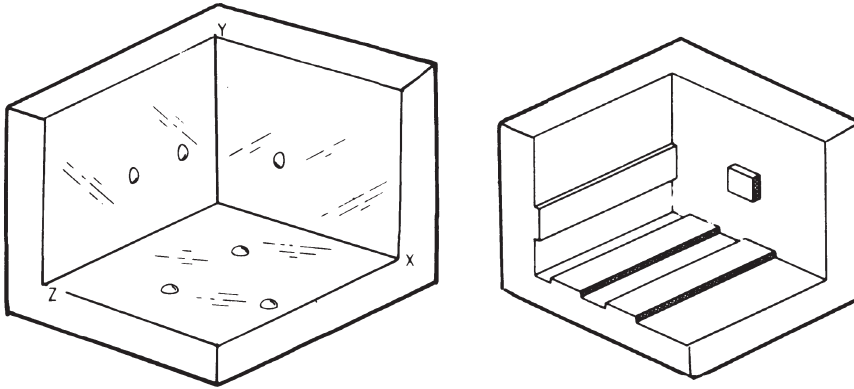


Figure 15.12 Kinematic and semikinematic position defining mount for a rectangular piece.

rotational freedom. Note that if there were three nonaligned points in this plane, they would then define an angle between the XZ and YZ faces of the piece; if the piece had a different angle, then there would be *two* ways in which the piece could be seated. The single point in the XY plane eliminates the last remaining of the six available degrees of freedom. A flexible pressure on the near corner of the piece will now uniquely locate the piece in this mount.

The sketch on the right illustrates one way of putting this type of mount into practice. The points are replaced by pads or rails. As shown, the two rails in the XZ plane must be carefully machined in the same operation to assure that they are exactly coplanar: this is not difficult, but if it were, the substitution of a short pad for one rail would eliminate any difficulty on this score.

Prisms and mirrors are usually clamped or bonded to their mounts. In clamp mounts the pressure is usually exerted by a screw on a metal pressure pad. A piece of cork or compressible composition material is placed between the glass and the metal pad to distribute the pressure evenly over the glass; this prevents the pressure from being exerted at a single point. There are a number of excellent cements available for bonding glass pieces to metal mounts. Some care is necessary in designing the mount when bonding a thin mirror, since the cement may warp the mirror (toward the shape of the mount) if the cemented area is large.

15.4 Optical Laboratory Practice

The lens bench. An optical bench or lens bench consists, in essence, of a collimator which produces an infinitely distant image of a test target,

a device for holding the optical system under test, a microscope for the examination of the image formed by the system, and a means for supporting these components. Each of the components may take various forms, depending on the usage for which it is primarily designed.

The collimator consists of a well-corrected objective and an illuminated target at the focus of the objective. For visual work, the objective is usually a well-corrected achromat; for infrared work, a paraboloidal mirror is used, usually in an “off-axis” or Herschel configuration. The target may be a simple pinhole (for star tests or energy distribution studies), a resolution target, or a calibrated scale if a “focal” collimator is desired.

The lens holder can range in complexity from a simple platform with wax to stick the lens in place to a T-bar nodal slide which generates a flat image surface. The microscope is usually equipped with at least one micrometer slide, and frequently with two or three orthogonal slides so that accurate measurements may be made.

In subsequent paragraphs, we will discuss some of the applications of the lens bench and will describe the components of the bench more fully in the context of their applications.

The measurement of focal length. There are two basic lens bench techniques for the routine measurement of effective focal length: the nodal slide method and the focal collimator. Both schemes are sketched in Fig. 15.13.

The *nodal slide* is a pivoted lens holder equipped with a slide which allows the lens to be shifted axially (i.e., longitudinally) with respect to the pivotal axis. Thus, by moving the lens forward or backward, the lens can be made to rotate about any desired point. Now note that, if the lens is pivoted about its second nodal point (as indicated in Fig.

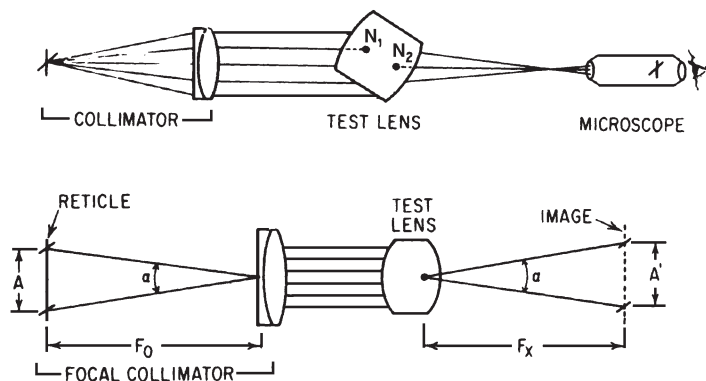


Figure 15.13 Illustrating the nodal slide (upper) and the focal collimator (lower) methods of measuring focal length on the optical bench.

15.13), the ray emerging from this point (which by definition emerges from the system parallel to its incoming direction) will coincide with the bench axis (through the nodal point). Thus there will be no lateral motion of the image when the lens is rotated about the second nodal point. Once the nodal point has been located in this manner, the lens is then realigned with the collimator axis and the location of the focal point is determined. Since the nodal points and principal points are coincident when a lens is in air, the distance from the nodal point to the focal point is the effective focal length.

This technique is basic and applicable to a wide variety of systems. Its limitations are primarily in the location of the nodal point. The operation of swinging the lens, shifting its position, swinging again, and so on, is tedious, and since it is discontinuous, it is difficult to make an exact setting. If the axis of the test lens is not accurately centered over the axis of rotation of the nodal slide, there will be no position at which the image stands still. Lastly, the measurement of the distance from the axis of rotation to the position of the aerial image is subject to error unless the equipment is carefully calibrated.

A *focal collimator* consists of an objective with a calibrated reticle at its focal point. The focal length of the objective and the size of the reticle must be accurately known. The test lens is set up and the size of the image formed by the lens is accurately measured with the measuring microscope. From Fig. 15.13 it is apparent that the focal length of the test lens is given by

$$F_x = A' \left(\frac{F_0}{A} \right) \quad (15.10)$$

where A' is the measured size of the image, A is the size of the reticle, and F_0 is the focal length of the collimator objective. Note that the focal collimator may be used to measure negative focal lengths as well as positive; one simply uses a microscope objective with a working distance longer than the (negative) back focus of the lens under test.

It is apparent from Eq. 15.10 that any inaccuracies in the values of A' , A , or F_0 are reflected directly in the resultant value of the focal length. Further, any error in setting the longitudinal position of the measuring microscope at the focus will be reflected in F_x . Note that both the nodal slide and focal collimator methods assume that the test lens is free of distortion. If an appreciable amount of distortion exists, the measurements must be made over a small angle; this, of course, will limit the accuracy possible.

In setting up a focal collimator, it is necessary to determine the collimator constant (F_0/A) to as high a degree of accuracy as possible. The value of A , the reticle spacing, can be readily measured with a measuring microscope. The focal length of the collimating lens can be deter-

mined to a high degree of accuracy by a finite conjugate version of the focal collimator technique. An accurate scale (or glass plate with a pair of lines) is set up 20 to 50 ft from the collimator lens, as shown in Fig. 15.14. The measuring microscope is used to measure the size of the image of the target accurately, and the distance from object to image is measured. The value of p , the distance between the principal points is estimated, either from the design data of the lens or by assuming it to be about one-third the lens (glass) thickness. (As long as p is small compared to D , the error introduced by an inaccurate value of p is small.) Now since $D = s + s' + p$ and $A:s = A':s'$, s and s' can be determined and substituted (with due regard for the sign convention) into

$$\frac{1}{s'} = \frac{1}{f} + \frac{1}{s} \quad (15.11)$$

and the value of the effective focal length determined. The necessity of estimating a value for p can be eliminated, if desired, by measuring the front focal length and applying the newtonian equation for magnification (Eq. 2.6) or, alternatively, by measuring front and back focal lengths (as outlined in the next paragraph), determining $p = \text{ffl} + \text{bfl} + t - 2f$, and repeating the original calculation; after a few iterations the calculation will converge to the exact p and f .

Collimation and the measurement of front and back focal lengths. A basic method of locating the focal points is by autocollimation. As indicated in Fig. 15.15, an illuminated target is placed near the focus of the lens

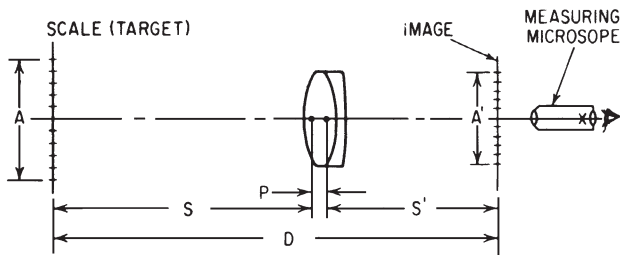


Figure 15.14 Setup for basic measurement of focal length.

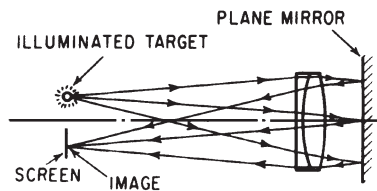


Figure 15.15 Autocollimation as a method of locating the focal points. When the object and reflected image are in the same plane (the focal plane), the system is autocollimated.

under test and a plane mirror is placed in front of the lens so as to reflect the light back into the lens. When the reflected image is focused on a screen in the same plane as the target, both screen and target lie in the focal plane. For accurate work an autocollimating microscope, shown in Fig. 15.16, produces excellent results. The lamp and condenser illuminate the reticle, which may consist of clear lines scribed through an aluminized mirror. The reticle is then imaged at the focus of the microscope objective. The eyepiece of the microscope is positioned so that its focal plane is exactly conjugate with the reticle. Thus when the microscope is focused on the focal plane of the test lens, the reticle image is autocollimated by the test lens-plane mirror combination and is seen in sharp focus at the eyepiece. The microscope is then moved in to focus on the rear surface of the test lens; the distance traveled by the microscope is equal to the back focus of the lens.

The lens bench collimator itself may be adjusted for exact collimation using this technique. When the collimator reticle and the reflected image of the microscope reticle are simultaneously in focus, then the collimator is in exact adjustment. Note that the mirror must be a precise plano surface if accurate results are expected.

For routine measurements of back focus the bench collimator is substituted for the plane mirror, and if no autocollimating microscope is available, a little powder or a grease pencil mark on the rear surface of the test lens can be used as an aid in focusing on the lens surface.

In the absence of many of the usual laboratory trappings, it is still possible to make reasonably accurate determinations of focal lengths and focal points. A lens may be collimated simply enough by focusing it on a distant object. The error in collimation can be determined by the newtonian equation $x' = -f^2/x$, where x is the object distance less one focal length and x' is the error in the determination of the focal position. A set of distant targets, such as building edges, smokestacks, and the like, whose angular separations are accurately known can often be substituted for a focal collimator in determining focal lengths.

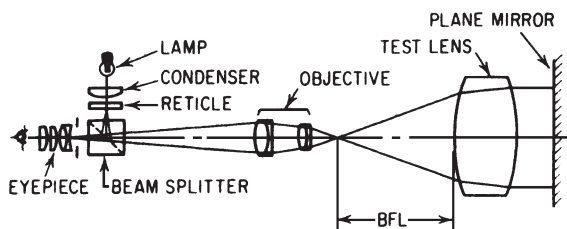


Figure 15.16 The autocollimating microscope is used to measure back focal length by focusing first on the surface of the test lens and then on the autocollimated image at the focal point.

Measurement of telescopic power. The power of a telescopic system can be measured in three different ways. If the focal lengths of the objective and eyepiece (including any erectors) can be measured, their quotient equals the magnification. The ratio of the diameters of the entrance and exit pupils will also yield the magnifying power. Occasionally the multiplicity of stops in a telescope will introduce some confusion as to whether the pupils measured are indeed conjugates; in this case the image of a transparent scale laid across the objective can be measured at (or near) the exit pupil to determine the ratio. When the field of view is sharply defined, the magnification can be determined by taking the ratio of the tangents of the half-field angles at the eyepiece and the objective. Note that the almost inevitable distortion in telescopic eyepieces will usually cause this measurement of power to differ from measurements made by focal lengths or pupil diameters. One should ascertain that the telescope is in afocal adjustment before measuring the power. One way of doing this is to use a low-power (3 to 5 \times) auxiliary telescope (or dioptometer) previously focused for infinity at the eyepiece; this reduces the effect of visual accommodation when the focus is adjusted.

The measurement of aberrations. In most instances the aberrations of a test lens can be readily measured on the lens bench by simulating a raytrace. For the measurement of spherical or chromatic aberration, a series of masks, each with a pair of small (to the order of a millimeter in diameter) holes, is useful. As indicated in Fig. 15.17, such a mask, centered over the test lens, simulates the passage of two "rays." When the image is examined with a microscope, a double image of the target is seen, except when the microscope is focused at the intersection of the two rays. By measuring the relative longitudinal position of the ray intersections for masks of various hole spacings, the spherical aberration can be determined. If the measurements are made in red and blue light, the data will yield the chromatic and spherochromatic aberration of the lens.

Figure 15.18 indicates how a similar three-hole mask can be used to measure the tangential coma of a test lens. A multiple hole mask can also be used to measure and plot a ray intercept curve, if desired. The

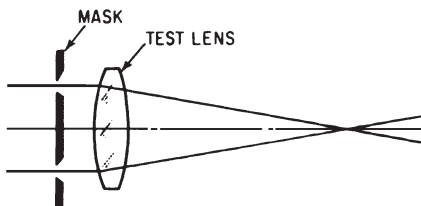


Figure 15.17 A two-hole mask can be used to locate the focus of a particular zone of a lens to determine the spherical aberration.

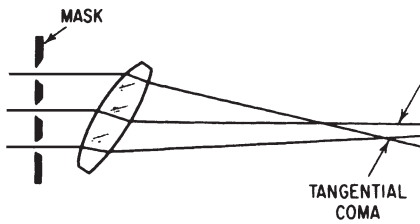


Figure 15.18 A three-hole mask can be used to measure the coma of a test lens.

technique for measurement of field curvature is indicated in Fig. 15.19. The bench collimator is equipped with a reticle consisting of horizontal and vertical lines. The focal length of the test lens is measured. The lens is then adjusted so that its second nodal point is over the center of rotation of the nodal slide and the position of the focal point (with the lens axis parallel to the bench axis) is noted. The lens is then rotated through some angle θ . From Fig. 15.19 it is apparent that the intersection of the (flat) focal plane of the lens with the bench axis will shift away from the lens by an amount equal to

$$efl \left(\frac{1}{\cos \theta} - 1 \right)$$

as the lens is pivoted through an angle θ . The bench microscope is used to measure D , the amount by which the focus shifts along the axis. Two measurements are necessary, one for the sagittal focus and one for the tangential focus; this is the reason for the orthogonal line pattern of the reticle. Now the departure (along the *bench* axis) of the image surface from a flat plane is equal to

$$D - efl \left(\frac{1}{\cos \theta} - 1 \right)$$

and the curvature of field (parallel to the *lens* axis) is given by

$$x = \cos \theta \left[D - efl \left(\frac{1}{\cos \theta} - 1 \right) \right]$$

Much of the numerical work in determining the field curvature by this method can be eliminated by the use of a T-bar attachment to the nodal slide. The cross bar of the T acts as a guide for the bench microscope, causing it to focus on the flat field position as the lens is pivoted. Thus one may measure the value of $x/\cos \theta$ directly; the use of the T-bar eliminates several sources of potential errors inherent in the method described above, although it does complicate the construction of the nodal slide. Measure at $\pm\theta$ to detect a tilted field.

Distortion is a difficult aberration to measure. The nodal slide may be used. The lens is adjusted on the slide so that no lateral image shift

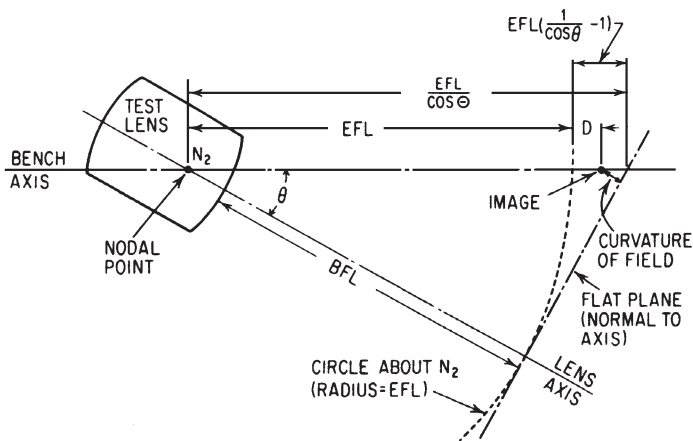


Figure 15.19 Geometry of the measurement of field curvature using the lens bench nodal slide.

is produced by a *small* rotation of the lens. Then as the lens is pivoted through larger angles, any lateral displacement of the image is a measure of the distortion. An alternate method is to use the lens to project a rectilinear target and to measure the sag or curvature of the lines in the image, or to measure the magnification of targets of several different angular sizes. The difficulty with *any* method of measuring distortion is that one invariably winds up basing the work on measurements of magnification (or whatever) vanishingly close to the axis, and the accuracy of such small measurements is usually quite low.

The star test. If the object imaged by a lens is effectively a “point,” i.e., if its nominal image size is smaller than the Airy disk, then the image will be a very close approximation to the diffraction pattern. A microscopic examination of such a “star” image can indicate a great deal about the lens to the experienced observer. One should be sure that the microscope NA is larger than that of the lens being tested. On the axis, the star image of a perfectly symmetrical (about the axis) system obviously must be a symmetrical pattern. Therefore, any asymmetry in the on-axis pattern is an indication of a lack of symmetry in the system. A flared or coma-shaped pattern on axis generally indicates a decentered or tilted element in the system. If the axial pattern is cruciform or shows indications of a dual focus, the cause may be axial astigmatism due either to a toroidal surface, a tilted or decentered element, or an index inhomogeneity.

The axial pattern may also be used to determine the state of correction of spherical and chromatic aberration. The outer rings in the diffraction pattern of a well-corrected lens are relatively inconspicuous,

and the pattern, when defocused, looks the same both inside and outside the best focus point. In the presence of undercorrected spherical, the pattern will show rings inside the focus and will be blurred outside the focus; the reverse is true of overcorrected spherical. When the spherical aberration is a zonal residual, the ring pattern tends to be heavier and more pronounced than that from simple under- or overcorrected spherical.

In the case of undercorrected chromatic, the pattern inside the focus will have a blue center and a red or orange outer flare. As the microscope focus is moved away from the lens, the center of the pattern may turn green, yellow, orange, and will finally become red with a blue halo. The reverse sequence will result from overcorrected chromatic. A chromatically “corrected” lens with a residual secondary spectrum usually shows a pattern with a characteristic yellow-green (apple green) center surrounded by a blue or purple halo.

Off-axis star patterns are subject to a much wider range of variations. The classical comet-shaped coma pattern is easily recognized, as is the cross- or onion-shaped pattern due to astigmatism. However, it is rare to find a system with a “pure” pattern off-axis, and it is much more common to encounter a complex mixture of all the aberrations, which are difficult, if not impossible, to sort out.

The star test is a very useful *diagnostic* tool requiring only minimal equipment, and, in skilled hands, it can be highly effective. The novice should be warned, however, that reliable judgments of relative quality are difficult, and a considerable amount of experience is necessary before one can safely depend on a star check for even simple comparative evaluations. It should *not* be used for quality control acceptance tests.

The Foucault test. The Foucault, or knife-edge, test is performed by moving a knife (or razor-blade) edge laterally into the image of a small point (or line) source. The eye, or a camera, is placed immediately behind the knife, and the exit pupil of the system is observed. The arrangement of the Foucault test is shown in Fig. 15.20. If the lens is perfect and the knife is slightly ahead of the focus, a straight shadow will move across the exit pupil in the same direction as the knife. When the knife is behind the focus, the direction of the shadow movement is the reverse of the knife direction. When the knife passes exactly through the focus, the entire pupil (of a perfect lens) is seen to darken uniformly.

The same type of analysis can be applied to *zones* of the pupil. If a zone or ring of the pupil darkens suddenly and uniformly as the knife is advanced into the beam, then the knife is cutting the axis at the focus of that particular zone. This is the basis of most of the

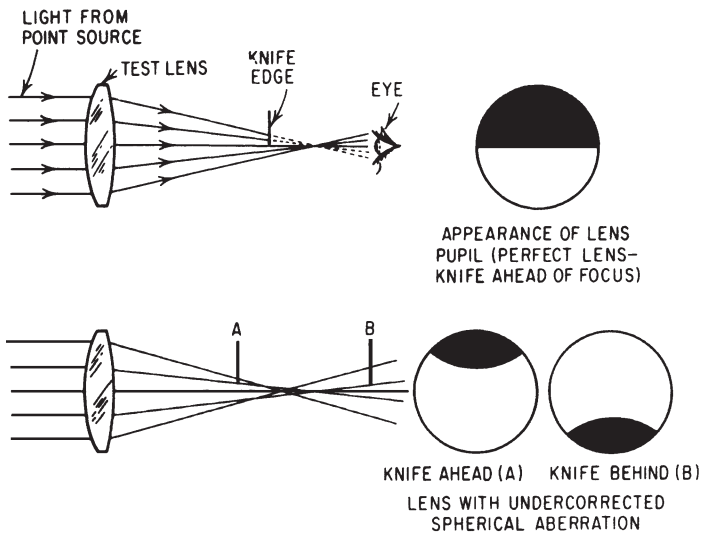


Figure 15.20 The Foucault knife-edge test. Upper: On a perfect lens the knife shadow has a straight edge. Lower: The shadow has a curved edge in the presence of spherical aberration. When the knife cuts through the focus, the pupil (or the zone of the focus) darkens uniformly.

quantitative measurements made with the Foucault test. The technique generally used is to place a mask over the lens with two symmetrically located apertures to define the zone to be measured. The knife is shifted longitudinally until it cuts off the light through both apertures simultaneously. It is then located at the focus for the zone defined by the mask. The process is repeated for other zones, and the measured positions of the knife are compared with the desired positions.

This test is extremely useful in the manufacture of large concave mirrors, which can be tested either at their focus or at their center of curvature. For the center-of-curvature test, the source is a pinhole closely adjacent to the knife (Fig. 15.21), and a minimum of space and equipment is required. Obviously if the mirror is a sphere, all zones will have the same focus, and a perfect sphere will darken uniformly as the knife passes through the focus. When the surface to be tested is an aspheric, the desired foci for the various zones are computed from the design data and the measurements are compared with the calculated values. It is a relatively simple matter to convert these focus differences into errors in the surface contour; in this way the optician can determine which zones of the lens or mirror require further polishing to lower the surface.

If the aspheric surface equation is expressed in the form

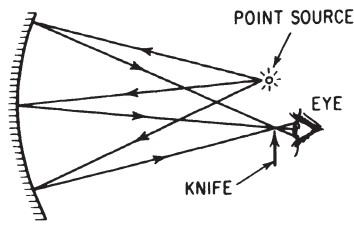


Figure 15.21 The knife-edge test applied to a concave mirror by placing both knife and source at the center of curvature.

$$x = f(y)$$

then the equation of the normal to the surface at point (x_1, y_1) is

$$y = y_1 + f(y_1) f'(y_1) - x f'(y_1)$$

[where $f'(\cdot) = dx/dy$], and the intersection of the normal with the (optical) axis is then

$$x_0 = x_1 + \frac{y_1}{f'(y_1)}$$

As an example, for a paraboloid represented by

$$x = \frac{y^2}{4f}$$

$$f'(y) = \frac{dx}{dy} = \frac{y}{2f}$$

and the axial intersection of the normal through the point (x_1, y_1) is

$$x_0 = x_1 + \frac{y_1}{(y_1/2f)} = x_1 + 2f = \frac{y_1^2}{4f} + 2f$$

This last equation gives the longitudinal position at which the knife edge should uniformly darken a ring of semidiameter y_1 , when a parabola is tested at the center of curvature (as in Fig. 15.21) and knife and source are simultaneously moved along the axis.

In practice, the knife edge is adjusted longitudinally until the central zone of the mirror darkens uniformly. The distance from the knife to mirror is then equal to $2f$. Then a series of measurements is made using masks with half-spacings of y_1, y_2, y_3 , etc., each measurement yielding an error e_1, e_2, e_3 , etc., where e is the longitudinal distance from the "desired" position for the knife to the actual position.

These data may be readily converted into the difference between the actual slope of the surface and the desired slope by reference to Fig. 15.22. When e is small, we can (to a very good approximation) write for our parabolic example

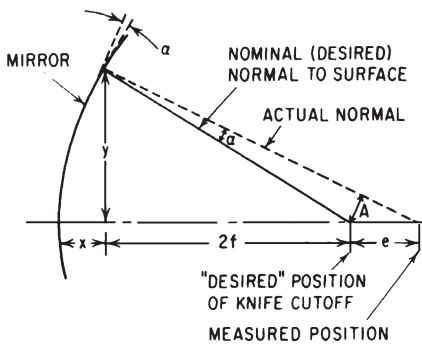


Figure 15.22 Geometry of knife-edge test used to determine the surface contour of a concave (paraboloidal) mirror.

$$\frac{A}{e} = \frac{y}{\sqrt{4f^2 + y^2}}$$

where the term in the right-hand denominator is the distance from the surface to the axis taken along the normal. Now the angle α between the actual normal and the desired normal is equal to

$$\alpha = \frac{A}{\sqrt{4f^2 + y^2}}$$

and substituting for A from the previous expression, we get

$$\alpha = \frac{ye}{4f^2 + y^2}$$

Note that α is also the amount by which the slope of the surface is in error; we can determine the actual departure of the surface from its desired shape by reference to Fig. 15.23. Taking the surface error at the axis as zero, the departure from the desired curve at y_1 is given by

$$d_1 = \frac{-y_1 \alpha_1}{2}$$

At y_2 it is

$$d_2 = d_1 - \frac{1}{2} (y_2 - y_1) (\alpha_1 + \alpha_2)$$

At y_3 it is

$$d_3 = d_2 - \frac{1}{2} (y_3 - y_2) (\alpha_2 + \alpha_3)$$

In general we can write

$$d_n = \frac{1}{2} \sum_{i=1}^{i=n} (y_{i-1} - y_i) (\alpha_{i-1} + \alpha_i)$$

where y_0 and α_0 are assumed zero, and the sign of d is positive if the actual surface is above (to the right in Figs. 15.22 and 15.23) the desired surface.

The method outlined above can be readily applied to any concave aspheric. Since it checks the aspheric only at discrete intervals, it must, of course, be supplemented with an overall knife-edge check to be certain that the surface contour is smooth and free from ridges or grooves. The testing of convex surfaces is more difficult; they are usually checked in conjunction with another mirror chosen so that the combination has an accessible "center focus." The computation of the normal is more involved in this case, but the principles involved are exactly the same.

The Schlieren test. The Schlieren test is actually a modification of the Foucault test in which the knife blade is replaced by a small pinhole. Thus any ray which misses the pinhole causes a darkened region in the aperture of the optical system. The Schlieren test is especially useful in detecting small variations in index of refraction, either in the optical system or in the medium (air) surrounding it. In wind-tunnel applications, the tunnel is set up between a collimating optical system and a matching system which focuses the image on the pinhole. When the test is recorded photographically, it is possible to derive quantitative data on the airflow from density measurements on the film.

Resolution tests. Resolution is usually measured by examining the image of a pattern of alternating bright and dark lines or bars. Conventionally, the bright and dark bars are of equal width. A target consisting of several sets of bar patterns of graded spacing is used, and

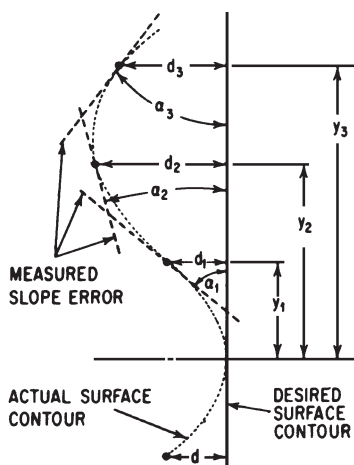


Figure 15.23 Conversion of measured errors of surface slope (α) into the departure (d) of the actual surface from the desired surface.

the finest pattern in which the bars can be distinguished (and in which the number of bars in the image is equal to the number in the object) is taken as the limiting resolution of the system under test.

The resolution patterns in use vary in two details of (relatively minor) significance: the number of lines or bars per pattern and the length of the lines relative to their width. The most common practice is to use three bars (and two spaces) per pattern, with a length of five, or more, bar widths. The USAF 1951 target is of this type and the patterns are graded in frequency with a ratio of the sixth root of 2 between patterns. The National Bureau of Standards Circular No. 533 includes both high- (25:1) and low- (1.6:1) contrast three-bar patterns which are approximately 1-in long and range in frequency from about one-third line per millimeter to about three lines per millimeter in steps of the fourth root of 2. A number of transparent (on film or glass) targets are commercially available; these are, for the most part, based on the USAF target.

Figure 15.24 shows two types of resolution test targets. The USAF 1951 target is probably the most widely used and accepted resolution target. The radial target is interesting since it nicely demonstrates the 180° phase shift of the optical transfer function. This produces the “spurious resolution” which is illustrated in Fig. 15.24c. See also Figs. 11.16 and 11.17.

In evaluating the resolution of a system it is important to adopt a rational criterion for deciding when a pattern is “resolved.” The following is *strongly* recommended: A pattern is resolved when the lines can be discerned, and when all coarser (lower-frequency) patterns also meet this requirement. This implicitly requires that the number of lines in the image be the same as in the target, and also rules out spurious resolution. Do *not* allow any consideration of “sharpness,” “definition,” “crispness,” “clearly resolved,” “contrast,” or the like to enter the evaluation; these are all subjective and involve individual interpretation. They lead to interminable arguments. The only consideration that should be used is “Can you discern the lines?”

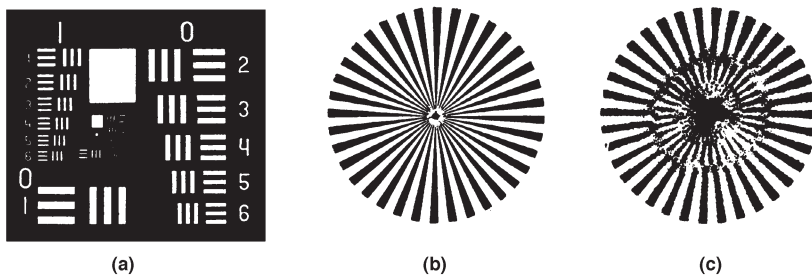


Figure 15.24 (a) USAF1951 resolution chart; (b) Siemens star resolution chart; (c) defocusing a well-corrected lens can cause a 180° phase shift which reverses the contrast of the pattern, causing areas which should be dark to be light and vice versa.

The resolution of a photographic system is tested by photographing a suitable target and examining the film under a microscope. In order to obtain optimum results, the photographic processes must be carried out with extreme care, especially with regard to the selection of the best focus, exposure and development, and the elimination of any vibration in the system. If the microscope used in examination of the test film has a power approximately equal to the number of lines per millimeter in the pattern, the visual image will have a frequency equal to one line per millimeter and will be easy to view.

Objective lenses can be tested on an optical bench with a resolution target in the collimator. For lenses with an appreciable angular coverage, an accurate T-bar nodal slide is practically a necessity if reliable off-axis results are to be obtained. Projection of a resolution target is a very convenient means of checking the resolution of lenses designed to cover areas less than a few inches in size. Care must be taken to ensure that the illumination system of the projector completely fills the aperture of the lens under test; otherwise, the results may be misleading. In all resolution tests, the alignment of the lens axis perpendicular to the target and film planes is a critical factor. The resolution of telescopic systems can be checked by visual observation of a suitably distant or collimated target. Since the limiting resolution of a telescope is frequently (by design) close to the limiting resolution of the eye, a common practice is to view the image through a low-power auxiliary telescope. Such a telescope serves a dual purpose in that it reduces the effect of the observer's visual acuity on the measurement and also reduces the effect that involuntary accommodation (focusing) can have.

The classical criterion for resolution, namely, the ability of a system to separate two point sources of equal intensity, is seldom used (except in astronomy). This is largely because a test using line objects is much easier to make.

Measurement of the modulation transfer function. The measurement of the MTF (frequency response) is, in principle, quite straightforward. The basic elements of the equipment are shown in Fig. 15.25. The test pattern is one in which the brightness varies as a sinusoidal function of one dimension. Such a target is not an easy thing to prepare; fortunately the errors introduced by a target which is not truly sinusoidal are unimportant for most purposes. Some instruments utilize "square-wave" targets. The target pattern is imaged by the test lens on a narrow slit whose direction is exactly parallel to the target pattern. The light passing through the slit is measured by a photodetector.

As the target *or the slit* is shifted laterally, the amount of light falling on the detector will vary, and the image modulation is given by

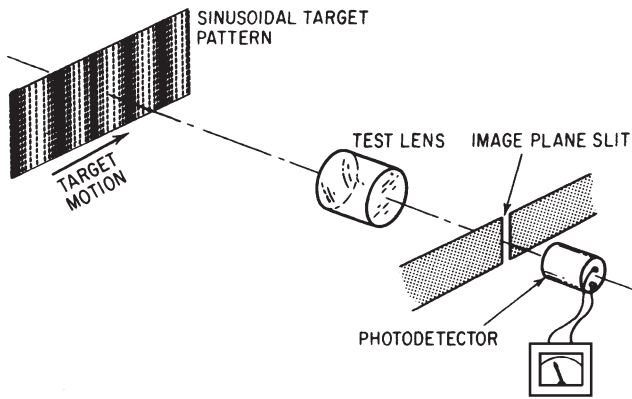


Figure 15.25 The basic elements of modulation transfer (frequency response) measurement equipment. The motion of the target scans its image across the narrow slit, where the maximum and minimum illumination levels are measured. By using targets of different spatial frequency, a plot of the modulation transfer function (vs. frequency) can be obtained.

$$M_i = \frac{\text{max} - \text{min}}{\text{max} + \text{min}}$$

where max and min represent the maximum and minimum illumination on the photodetector. The object modulation M_0 is similarly derived from the maximum and minimum brightness levels of the target. The MTF (or frequency response, or sine-wave response, or contrast transfer) is then the ratio $M_i : M_0$.

A provision is usually made to vary the spatial frequency of the target pattern so that the response may be plotted against frequency. The target portion of the system may be as simple as a set of interchangeable targets which are slowly traversed by hand, or it may be a fully automatic device which translates the target and scans a range of frequencies simultaneously.

The image-plane slit is almost never just a slit, since the manufacture of a slit of the required narrow dimensions can be fairly difficult. Instead, the image is magnified by a first-class microscope objective; this allows the use of a wider slit.

Obviously any real slit width will have some effect on the measurements, and a slit as narrow as the sensitivity of the photodetector will allow should be used. The effect of the slit width on the response may be readily calculated, since it simply represents a line spread function of rectangular cross section, and the data can be adjusted accordingly where necessary.

The source of illumination and the spectral response of the photodetector must, of course, be matched to the application for which the system under test is to be used. Otherwise, serious errors in measurement will result from the unwanted radiation outside the spectral band for which the system has been designed. Usually a set of filters can be found which will provide the proper response.

Another technique which is much more widely used than that described above is based on a *knife-edge scan*. A knife edge is passed through the image of a point (or slit) and the light passing by the edge is measured. If the measured light I is plotted against the lateral position of the knife edge y , the slope of the curve (dI/dy) is exactly equal to the line spread function of the lens. The MTF can be calculated from the line spread function using the methods outlined in Secs. 11.8 and 11.9. Most commercial MTF equipment is set up so that the knife-edge scan data are read directly into a computer which processes the data to calculate the MTF at whatever frequency is desired. Note that this technique does not require a sinusoidal target, nor does it require a separate target for each frequency. As in any MTF measurement, the spectral distribution of the source and the response of the light-measurement sensor must match that of the application.

The wave-front shape as measured by an interferometer can also be used to determine the MTF. The fringe pattern is scanned to digitize the data, and it is computer-processed to calculate the MTF at any desired frequency as in the knife-edge scan. This is entirely adequate for mirror systems or systems which operate at the laser wavelength. For systems which utilize a finite-width spectral band or a different wavelength, the results are not correct.

The analysis of “unknown” optics. It is frequently necessary to determine the constructional parameters of an existing optical system. An example might be the analysis of a sample system to determine the reason for its failure to perform to the designer’s expectations. Another example might be the analysis of an existing lens so that its design data can be used as the starting point for a new design. For the most part this amounts to the measurement of the radii, thicknesses, spacings, and indices of the system components.

Since the measurements to be made are frequently of a precision barely adequate for the purpose, it is best to provide as many interdependent checks on the process as possible. Thus the first steps should include accurate measurements of effective, back, and front focal lengths, as well as the aberrations, so that when all of the measured system data are collected, a calculation of the complete (measured) system can provide a final comparison check on the overall accuracy of the analysis.

The thicknesses and spacings of a system are readily measured. For small systems a micrometer (equipped with ball tips for concave surfaces) is sufficient. A depth gage or an oversize plunger caliper (Nonius gage) is useful for larger systems. If a dimension can be deduced from two different measurements (as a check), the extra time involved is usually a worthwhile investment.

The radius of an optical surface can be measured in many ways. The simplest is probably by use of a thin templet, or “brass gage,” cut to a known radius and pressed into contact with the surface. Differences between the gage and glass of a few ten-thousandths of an inch can easily be detected this way, but such a gage is not useful unless it very nearly matches the surface.

The classical instrument for radius measurement is the spherometer, the basic principles of which are outlined in Fig. 15.26. The spherometer measures the sagittal height of the surface over a known diameter; the radius is determined from the formula

$$R = \frac{Y^2 + S^2}{2S}$$

where Y is the semidiameter of the spherometer ring and S is the measured sagittal height. Since the sagittal height is a rather small dimension and thus subject to relatively large measurement errors, the accuracy of a spherometer leaves something to be desired even when extreme precautions are taken. One of the best ways to use a spherometer is as a comparison device, by measuring both the unknown radius and a (nearly equal) carefully calibrated standard radius (e.g., a test glass).

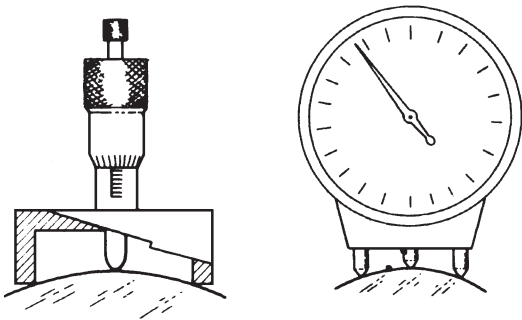


Figure 15.26 Left: Simple ring spherometer determines the radius of a surface through a measurement of the sagittal height. Right: The diopter gage or lens measure is a spherometer calibrated to read surface curvature in diopters.

The diopter gage, or lens measure, or Geneva lens gage is a handy tool which can provide a quick approximate measure of the surface curvature. As shown in Fig. 15.26, it consists of a dial gage with its plunger between two fixed points. The dial of a diopter gage is calibrated in diopters; the readings may be converted to radii by the formula

$$R = \frac{525}{D} \text{ millimeters}$$

where the 525 is the constant representing $1000(N - 1)$ for an “average” ophthalmic glass. The accuracy of a typical diopter gage is to the order of 0.1 diopter.

Probably the best way to measure a concave radius is by use of an autocollimating microscope. The microscope is first focused on the surface and is then focused at the center of curvature (where the microscope reticle image is imaged back on itself by reflection from the surface). The distance traveled by the microscope between these two positions is equal to the radius. The precision of this method can be to the order of micrometers; the accuracy is obviously dependent on the accuracy of the measurement method used. If the microscope used is of fairly high power (say $150\times$ with $NA = 0.3$), the quality of the reflected image at the center of curvature is an excellent indication of the sphericity of the surface. Convex surfaces can be measured in this way provided that the working distance of the microscope objective is longer than the radius. A series of long-focal-length objectives is useful in this regard, although the precision of the method drops as the NA of the objective is lowered (long-focal-length objectives usually have a small NA) due to the increased depth of focus. If a precise determination of a long convex radius is necessary, a mating concave surface can be made so that it fits perfectly (as tested by interference rings) and the measurement is made on the concave glass. Master test plates are measured by this technique. Note that Eq. 15.3 can be used to calculate small radius differences from interference fringe readings.

If a separate piece of glass from which the lens under analysis was made is available, the measurement of its index can be made with considerable precision. The minimum deviation of a test prism may be measured on a laboratory spectrometer, and the prism equations of Chap. 4 used to find the index. Alternatively, a Pulfrich refractometer measurement can be made. Either method will readily yield the index value accurate to the fourth decimal place. When one is constrained to measure the lens element itself, without destroying it, the problem is more difficult. A crude determination of the index can be made for normal glasses (i.e., not the newer “light” glasses) by measuring the density of the element. A plot of the catalog values of the index against density is then used to determine (very approximately) the corre-

sponding index. The relationship between index (n) and density (D) is very approximately $n = (11 + D)/9$.

A somewhat more general method is to measure the axial thickness of the element and then to measure the apparent optical thickness by focusing a measuring autoreflecting microscope first on one surface and then the other. A simple paraxial calculation, taking into account the refractive properties of the surface radius through which the second surface is viewed, will yield a value for the index. Depending on the thickness of the element, the index value achieved will probably be almost completely unreliable in the third place, due to the large relative inaccuracy in the measurement of the apparent thickness and to the spherical aberration introduced by the thickness of the glass.

If one measures the radii carefully and makes a good determination of the paraxial focal length of the element, the thick-lens formula for focal length can be solved to determine the index of refraction. Although this method requires skilled laboratory technique, it is capable of producing results which are accurate to one or two digits in the third place. Note that if care is not taken to eliminate the effects of spherical aberration from the focal-length measurement, the resulting index value will tend to err on the high side. Another nondestructive technique involves immersion in index-matching liquids, then measuring the index of the liquid.

Bibliography

Note: Titles preceded by an asterisk (*) are out of print.

Baird, K., and G. Hanes, in Kingslake (ed.), *Applied Optics and Optical Engineering*, vol. 4, New York, Academic, 1967 (interferometers).

DG-G-451, Flat and Corrugated Glass.

*Deve, C., *Optical Workshop Principles*, London, Hilger, 1945.

Habell, K., and A. Cox, *Engineering Optics*, London, Pitman, 1948.

Hopkins, R., in Shannon and Wyant (eds.), *Applied Optics and Optical Engineering*, vol. 8, New York, Academic, 1980 (lens mounting).

Ingalls, G., *Amateur Telescope Making*, books 1, 2, and 3, *Scientific American*, 1935, 1937, 1953.

JAN-P-246 Slide Projectors.

Malacara, D., "Optical Testing," in *Handbook of Optics*, vol. 2, New York, McGraw-Hill, 1995, Chap. 30.

Malacara, D., and Z. Malacara, "Optical Metrology," in *Handbook of Optics*, vol. 2, New York, McGraw-Hill, 1995, Chap. 29.

MIL-A-003920 Thermosetting Optical Cement.

MIL-C-48497 Scratch and Dig for Opaque Coatings.

MIL-C-675 Antireflection Coatings.

- MIL-G-1366 Aerial Photography Window Glass.
 MIL-G-16592 Plate Glass.
 MIL-L-19427 Anamorphic Projection Lenses.
 MIL-M-13508 Front Surface Aluminized Mirrors.
 MIL-O-13830 Scratch and Dig Specifications.
 MIL-O-16898 Packaging Optical Elements.
 MIL-P-47160 Optical Black Paint.
 MIL-P-49 16-mm Projectors.
 MIL-R-6771 Glass Reflectors, Gunsight.
 MIL-STD-1241 Optical Terms and Definitions.
 MIL-STD-150 Photographic Lenses.
 MIL-STD-34 Drawings for Optical Elements and Systems.
 MIL-STD-810 Interference Filters.
 McLeod and Sherwood, *J. Opt. Soc. Am.*, vol. 35, 1945, pp. 136–138
 (origin of the scratch and dig standards).
 Offner, A., *Applied Optics*, vol. 2, 1963, pp. 153–155 (null lens for
 parabola).
 Parks, R., “Optical Fabrication,” in *Handbook of Optics*, vol. 1, New
 York, McGraw-Hill, 1995, Chap. 40.
 Parks, R., in Shannon and Wyant (eds.), *Applied Optics and Optical
 Engineering*, vol. 10, San Diego, Academic, 1987 (fabrication).
Photonics Buyers Guide, Optical Industry Directory, annually, Laurin
 Publishing Co., Pittsfield, Mass.
 Rhorer and Evans, “Fabrication of Optics by Diamond Turning,” in
Handbook of Optics, vol. 1, New York, McGraw-Hill, 1995, Chap. 41.
 Sanger, G., in Shannon and Wyant (eds.), *Applied Optics and Optical
 Engineering*, vol. 10, San Diego, Academic, 1987 (fabrication, dia-
 mond turning).
 Scott, R., in Kingslake (ed.), *Applied Optics and Optical Engineering*,
 vol. 3, New York, Academic, 1965 (optical manufacturing).
 Shannon, R. R., “Optical Specifications,” in *Handbook of Optics*, vol. 1,
 New York, McGraw-Hill, 1995, Chap. 35.
 Shannon, R. R., “Tolerancing Techniques,” in *Handbook of Optics*, vol.
 1, New York, McGraw-Hill, 1995, Chap. 36.
 Shannon, R., in Kingslake (ed.), *Applied Optics and Optical
 Engineering*, vol. 3, New York, Academic, 1965 (testing).
 Shannon, R., in Shannon and Wyant (eds.), *Applied Optics and Optical
 Engineering*, vol. 8, San Diego, Academic, 1980 (aspherics).
 *Strong, J., *Procedures in Experimental Physics*, Englewood Cliffs,
 N.J., Prentice-Hall, 1938.
 Strong, J., *Procedures in Applied Optics*, New York, Dekker, 1989.
The Optical Industry Directory, Pittsfield, Mass., Photonics Spectra
 (published annually).
 Twyman, F., *Prism and Lens Making*, London, Hilger, 1988.

- Yoder, P. R., *Mounting Lenses in Optical Systems*, S.P.I.E., vol. TT21, 1995.
- Yoder, P. R., "Mounting Optical Components," in *Handbook of Optics*, vol. 1, New York, McGraw-Hill, 1995, Chap. 37.
- Yoder, P., *Opto-Mechanical System Design*, New York, Dekker, 1986.
- Young, A., in Kingslake (ed.), *Applied Optics and Optical Engineering*, vol. 4, New York, Academic, 1967 (optical shop instruments).
- Zschommler, W., *Precision Optical Glassworking*, New York, Macmillan/S.P.I.E., 1984.

