WAVE-FRONT ABERRATIONS: THEIR EFFECTS IN WHITE-LIGHT TRANSMISSION HOLOGRAPHY

Stephen A. Benton

The Media Laboratory, Massachusetts Institute of Technology
Room E15-422, 20 Ames Street, Cambridge, MA 02139
and
Research Laboratories, Polaroid Corporation
750 Main Street-LA, Cambridge, MA 02139

ABSTRACT-

Practical real-image projection produces wavefronts that are not perfect spheres converging toward ideal image points, but are instead twisted or warped surfaces. The wavefront warps are termed "aberrations," and cause light rays to miss the points predicted by simple holographic theories, often by significant amounts. Here we will discuss astigmatism, coma, and spherical aberration, and the image distortions they produce in white-light transmission "rainbow" holograms. The effects are different in the "transfer" and "viewing" stages, and can sometimes be controlled by proper choices of recording geometry.

1. INTRODUCTION-

Real-image projection has come to be one of the most important optical techniques in modern display holography, being used in transmission and reflection holography, for both full-aperture and slit-aperture transfers. The technique is based on optical phase conjugation, perhaps more vividly described as "reverse ray tracing," a concept pioneered by holographers, and recently taken up by explorers of non-linear optics. The principle is that if a normally recorded Leith-Upatnieks virtual-image hologram is illuminated through its back by a beam that is headed in a direction opposite to the original reference beam, then the diffracted light will also head in the direction opposite to the original object beam, and so converge to each point of the object, creating a real image focused in space. If the illumination is an exact "phase conjugate" or "time-reverse" of the reference beam, then the diffracted (or "projected") beam will be an exact "phase conjugate" or "time reverse" of the object beam. That is, all the rays will trace backwards upon themselves, and each point of the object will be accurately focused at its original position in space. Thus an un-magnified, undistorted, interferometrically accurate image is projected into space (an aerial image, specifically a "real" image), for viewing or further holographic recording.

The problem is that accurate phase-conjugate illumination is often difficult to provide. It may require large (and expensive) converging lenses, for example. Thus, in practice, diverging beams are often used for reference and illumination. Also the wavelength may have to be changed if, for example, He-Ne (633 nm) laser light is used to record a master that is subsequently transferred into photo-resist with a He-Cd (442 nm) laser. As a rule, a simple prism/lens (or rather, diffraction grating & Fresnel zone plate) model of the hologram is adequate to predict the changes of projected real image location, size, and even simple distortions (Benton, 1982).

However, the fact is that as soon as imperfect conjugates are used for the reference and illumination beams, the diffracted wavefronts are no longer perfectly spherical, and degraded images of object points are formed. Both the warped wavefront and the spread-out image point are said to be "aberrated." Rather than discuss the aberrations of wavefronts, which are the most direct result of a coherent-optical analysis, we will be describing the fates of mis-directed "rays," by which we mean the local perpendiculars to the wavefront, which describe (at least approximately) the straight-line trajectories of the wavefront's energy. Practicing holographers find "rays" much easier than wavefronts to trace out and understand, and similarly we will discuss the aberrated point image in terms of "spot diagrams" of ray bundles that don't share a common center, and so miss the intended target.

Optical aberrations have generally been discussed in the context of diffused images, such as formed on the surfaces of detectors (film and TV cameras, etc.), and have hardly been discussed at all in terms of their effects on aerial images (Bazargan & Forshaw, 1980). Because of the relatively high tolerance of the human eye to focus errors, aberrated aerial images seem mostly to be distorted, to swim about, or simply to hurt the viewer's eyes. The situation is doubly complex in holography because real image projection is generally used twice, in a transfer step, and then during viewing, and the effect of any particular aberration will be quite different in each case. The study of aberrations is itself quite complex (Meier, 1965), so that we are compelled to make as many simplifications as possible in this brief paper. We will restrict the discussion to white- light transmission "rainbow" holography, partly because of the author's interest in that field, but mainly because the strip aperture makes the ray aberrations somewhat easier to visualize and understand. No derivation of the aberration formulae will be attempted, but the mathematical results will be displayed in the hope of learning how to mitigate the worst of these common imaging errors.

2. ABERRATIONS OF REAL-IMAGE PROJECTIONS-

The optical/holographic geometry is defined as shown in Figs. 1a and 1b. These definitions have no basis in optical logic or consistency, but come close to how the directions and distances might actually be laid out on a holography table. To get an un-aberrated (or "stigmatic") projection, the illumination would be of the same wavelength as the reference beam (ie. lambda-2 equals lambda-1), and come in directly opposite to it (theta-ill equals theta-ref plus 180 degrees). Also, if the reference beam is diverging from a point, the illumination beam would have to be converging back toward that point (d-ill equals negative d-ref).

Figures 2 & 3 show how successive real-image projections are used in white-light transmission holography. First in the transfer step, where a converging wavefront from the H1 will produce a point focus of any one object point. If the focus is out of the plane of the H2, a "streak" is formed that is often quite visible if the beam ratio is too low. Upon viewing, a point along this "streak" lights up, and moves from side to side to present visual parallax as the viewer moves about. Aberrations in the transfer will distort this streak, and the point's motion within it. Also during viewing, a real image of the H1 slit is formed in space, which serves as the viewing zone. Aberrations of the viewing zone are more subtle in their effect, but can also distort the three-dimensional image. Often, both the reference and illumination beams are collimated in the case of H1, but usually not for H2, and a wavelength change may be introduced, especially for the H2 in embossed rainbow holography, or in color-shifted reflection holography.

Lenses have always been imperfect, but the analytical study of aberrations goes back only to L. Seidel in 1856, who categorized the simplest of them into four (now five) types: spherical aberration (which limits the on-axis performance of simple telescopes and collimators), coma (which produces comet-like streaks in off-axis images), astigmatism (the dreaded nemesis of contact lens wearers), curvature of field (which requires spherical film plates) and distortion ("pincushion" and "barrel" distortion are the usual types) (Born & Wolf, 1962). We have all heard of at least some of these optical denizens, and here we will discuss only the first three, with an emphasis on astigmatism. In practice, each has its most apparent effect in only one of the holographic steps, and the magnitude of each effect depends on the details of the choice of table geometry. Indeed, the various combinations of effects can be so distinctive as to serve as virtual "signatures" of the holographers. Perhaps art historians of the future will make the most detailed study of aberration theory, as a way of verifying work from these early crucial years!

Again, this will be only a drastically simplified overview of each of the first three aberrations in turn. Much work remains to be done in unravelling the details, especially in more general setups.

3. SPHERICAL ABERRATION-

Spherical aberration is so-named because perfectly spherical lens surfaces do not generally produce a perfectly spherical wavefront. Instead, the edges often curl in too tightly, so that "marginal" rays are focused closer to the lens than central or "paraxial" rays. In holography, its effects are most apparent in the transfer step, where it causes the familiar "rolling barrel" effect, in which a flat surface that is calculated to come out in the plane of the H2 is found to be curved backward, like a cylinder, and only the part directly in front of the viewer approaches the image plane. As the viewer moves, that cylindrical "barrel" seems to "roll" to stay directly in front of him or her. It is an effect every beginning rainbow holographer has seen, and its worth a modest struggle to understand. Unfortunately, spherical aberration is a function only of the beam distances, and not angles, so that accurate collimation is the only way to control it.

The basic effect is that, upon projection, rays from near the ends of the H1 "miss" the focus by an amount that increases with the cube of the distance from the H1 center, and cross the axis further behind (this is opposite to the usual lens case). The horizontal "miss angle" is expressed as:

$$\Delta\theta y_{OUT} = \frac{y^3}{2\cos\theta y_{OUT}}\left[\frac{\lambda 2}{\lambda 1}\left(\frac{1}{d^3_{OBJ}} - \frac{1}{d^3_{REF}}\right) - \frac{1}{d^3_{ILL}} - \frac{1}{d^3_{HORIZ\ FOCUS}}\right] \qquad \text{(Eq. 1)}$$

If such a "splayed" ray bundle is transferred into an H2, and then viewed, the rays will seem to emanate not from a single point, but from a changing location. That location will seem to be at the intersection of nearby rays from the appropriate part of the H1, and will also be behind the central focus, by an amount that increases with the square of the distance from the viewing zone center. Because each point of a scene is closest when the viewer is directly in front of it, and moves back as the viewer moves to the side, flat objects appear curved into a cylinder that seems to roll as the viewer moves from side to side. This "rolling barrel" distortion is well-known to rainbow holographers who lack collimators for their master plates.

4. COMA-

When a lens is used to form an image of a point well away from the optical axis, the result is often a comet-shaped blob, with the bright "head" coming from the center of the lens, and the smeared-out "tail" from near the rim of the lens. An off-axis hologram produces a similar aberration when imperfectly illuminated. The results appear in viewing a rainbow hologram as a curving of the output spectrum. But a more uncomfortable result is produced during the transfer, when the "streak" for each image point becomes bowed, most often upward. During viewing, the resulting image appears to move up and down as the viewer moves from side to side, and so the eyes are aimed at slightly different vertical positions. Our eyes have a very small tolerance for such vertical disparities, and they quickly cause discomfort, to the point where a viewer will exclaim "this makes my eyes hurt!" Thus while the apparent spatial distortions may be small, we may be making holograms that only flexible-eyed holographers can truly enjoy!

Analytically, coma during transfer is described by a vertical "miss angle" that varies as the square of the distance from the center of the plate. The angle is expressed as:

(Eq. 2)

$$\Delta\theta x_{OUT} = \frac{-y^2}{2\cos\theta x_{OUT}}\left[\frac{\lambda 2}{\lambda 1}\left(\frac{\sin\theta_{OBJ}}{d^2_{OBJ}} - \frac{\sin\theta_{REF}}{d^2_{REF}}\right) - \frac{\sin\theta_{ILL}}{d^2_{ILL}} - \frac{\sin\theta_{OUT}}{d^2_{HORIZ\ FOCUS}}\right]$$

The vertical miss angle is seen to be a function of the sines of the angles over the cubes of the related distances. Thus the amount of coma-induced "bowing" of the image streaks can be minimized by adjusting the reference and object beam angles, which reduces to choosing the "tip angle" of the master plate with respect to the object. Depending on the exact wavelength changes and source distances involved, the desired "tip angle" is usually very close to zero; that is, the plate nearly faces the object. Holographers usually prefer to "split" the object and reference beam directions (theta-obj equals-theta-ref, or theta-tip = (theta-ref-theta-obj)/2), because the interference fringes are perpendicular to the emulsion in that case, and so shrinkage during processing need not be carefully controlled. Minimization of coma therefore also requires more careful processing for best results.

5. ASTIGMATISM-

In some ways, astigmatism is the most elemental of the aberrations, although its effects may be fairly subtle. It will receive the most discussion here because of the important modifications to the usual mathematical lens model that it requires. Astigmatism is the well-known nemesis of contact lens wearers, because it requires that the lens be worn in a specific orientation; otherwise, horizontal and vertical edges cannot be focused at the same time. In holography, its effects are more obscure: astigmatism during a rainbow transfer may actually improve the apparent depth-of-field of an image (Leith & Chen, 1978), or equivalently minimize the dispersion in a full-aperture transfer (Bazargan & Forshaw, 1980). Here, we will emphasize the effects of astigmatism in viewing rainbow holograms, where it separates those distances at which images are either seen in a single color or seen without distortion.

An astigmatic wavefront has different curvatures in the vertical and horizontal directions. Thus instead of converging to a single focal point, the rays cross to form a horizontal line, and then perhaps further away, a vertical line. This situation is sketched in Fig. 4 to represent the focus of an H2 rainbow hologram when imperfectly illuminated. It is forming an aberrated image of the H1 slit instead of a point, but we shall consider a small area at the center of the slit as a "nearly-point" object.

Any small area of the H1 can be considered as recording a particular "pin-hole camera" view of the subject, which is subsequently projected onto the H2 during transfer. Upon perfect illumination, an image of that H1 area is formed in space, and if an eye is placed in that image, it sees only the corresponding perspective view, without any distortion, and also in a single color over the entire height of the H2. If the viewer is at another distance, then out of H2-plane circles will appear elliptical, and the apparent color will vary from top to bottom. For example, at very large viewing distances, a circle's image will be higher than wide, and the top of the image will be reddish and the bottom bluish (assuming overhead illumination) (the reader should explore independently the physical bases of these viewing-distance effects).

If the illumination is not perfectly conjugate to the reference beam, perhaps with the wrong divergence, so that the image of the H1 area is astigmatic, then the distance to the horizontal or side-to-side focus will determine the distortion-free viewing zone, while the distance to the vertical or up-to-down focus will determine the viewing zone for single-color viewing. The two different distances are predicted by very similar equations of the lens-focusing law type; in fact the horizontal focus is predicted by the same formula used for the simple white-light transmission hologram model (Benton, 1982). It is only the vertical focus that requires a slight modification, replacing the "ones" in all the denominators by the squares of the cosines of the associated angles.

$$\frac{1}{d_{\substack{VERT\\FOCUS}}} = \frac{\lambda 2}{\lambda 1}\left(\frac{1}{d_{OBJ}} - \frac{1}{d_{REF}}\right) - \frac{1}{d_{ILL}} \qquad \text{(Eq. 3A)}$$

$$\frac{\cos^2\theta_{OUT}}{d_{\substack{HORIZ\\FOCUS}}} = \frac{\lambda 2}{\lambda 1}\left(\frac{\cos^2\theta_{OBJ}}{d_{OBJ}} - \frac{\cos^2\theta_{REF}}{d_{REF}}\right) - \frac{\cos^2\theta_{ILL}}{d_{ILL}} \qquad \text{(Eq. 3B)}$$

These equations are presented without proof, but it should suffice to say that the proof hinges on the observation that the distance from a source point (such as for the reference beam) to a recording point near the hologram center has a second derivative with respect to distance from the hologram center that depends on the direction of the recording point from the center. If the recording point is moving in the horizontal direction, it depends on the reciprocal of the source distance, but if it is moving in the vertical direction, it depends on the cosine-squared of the reference beam angle, divided by the source distance.

In practice, the vertical, or single-color, viewing distance is usually closer to the hologram than the horizontal, or distortionless, viewing distance, although this can be varied by changing the "tip angle" of the H2 (that is, the angle its perpendicular makes between the object and reference beam), as shown in Fig. 5. If the H2 "splits the angle" between the object and reference beams, then astigmatism is zeroed, and the two viewing distances coincide even for imperfect illumination.

However, holograms are often required to hang vertically, so that a choice among the two of intended viewing distances is required. Generally, if the image is very deep, the horizontal focus should be used. But if a carefully-controlled multi-color image is desired, then the more complex vertical focus calculation should be used.

6. CONCLUSIONS-

We all depend on the development of an intuitive grasp of holographic imaging in order to conceive of and produce effective holograms, and this practical sense of "what works" follows fairly quickly from studio experience. But aberrations are a very real part of every holographer's life, and when precision image placement is required, such as for the registration of multi-image components, it is important to double-check those guidelines with calculations based on mathematical optics. These mathematics can be peeled apart like cabbage leaves, to be dealt with one leaf at a time, but they are all part of a whole that has a more complex structure than we may wish to deal with, but whose real beauty lies in the integrity of that whole. Here we have held a few inner leaves up to the light, as aberrations that we should begin to feel comfortable in confronting, if not controlling entirely.

Aberrations are complex, and daunting to behold, and their effects have long been considered as almost mysterious side-effects of holography as usually practiced. But the message here is that they are not impossible to analyze and discuss, and even to moderate. Their understanding can be based on the most fundamental principles of wavefront reconstruction. And as holographic imaging continues to become more and more sophisticated, it seems that more and more precision in dealing with optical physics will required.

7. ACKNOWLEDGEMENTS-

Experimental verification of these principles has been undertaken by Mr. Michael Teitel. The author is indebted to him and Ms. Julie Walker for useful discussions of the subject.

8. REFERENCES-

Bazargan, K. & M.R.B. Forshaw (1980), "Image-Plane Hologram with Non-Image- Plane Motion Parallax," Optics Comm'n 32(1), pp. 45-47 (1980)

Benton, S.A. (1982), "The Mathematical Optics of White-Light Transmission Holograms," in T.H. Jeong, Ed., Proc. Int'l Symp. on Display Holgraphy, Vol. 1, (Holography Workshops, Lake forest College, 1982). pp. 5-14

Born, M., and E. Wolf (1964), Principles of Optics, (Pergamon Press, Oxford, 2nd edn, 1964) Ch. V

Leith, E., and H. Chen (1978), "Deep Image Rainbow Holograms," Opt. Lett., Vol 2 (4), pp. 82-84 (April 1978)

Meier, R.W. (1965), "Magnification and Third-Order Aberrations in Holography," J.Opt.Soc.Amer. 55-8, pp. 987-992 (August 1965)

FIGURE 1a. RECORDING AT λ1

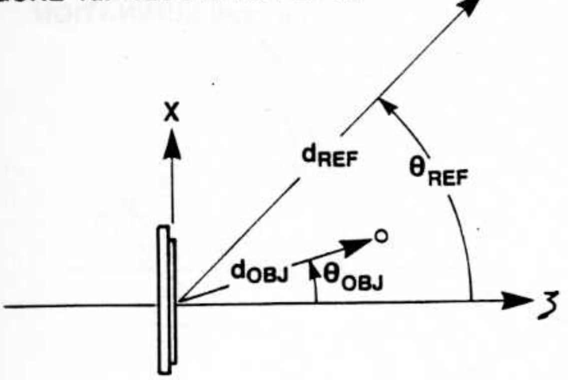

FIGURE 1b. RECONSTRUCTION AT λ2

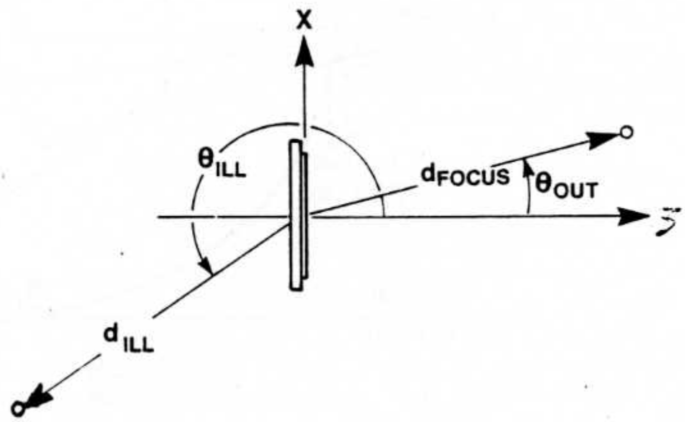

FIGURE 2. RAINBOW TRANSFER

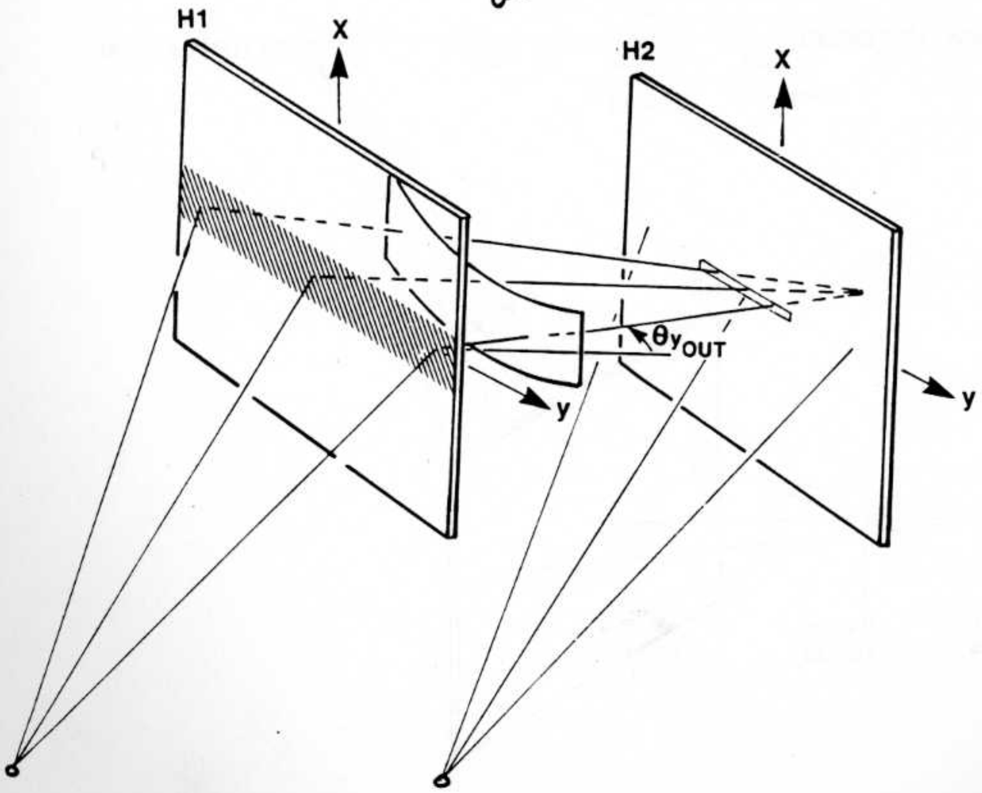

FIGURE 3. WHITE-LIGHT VIEWING

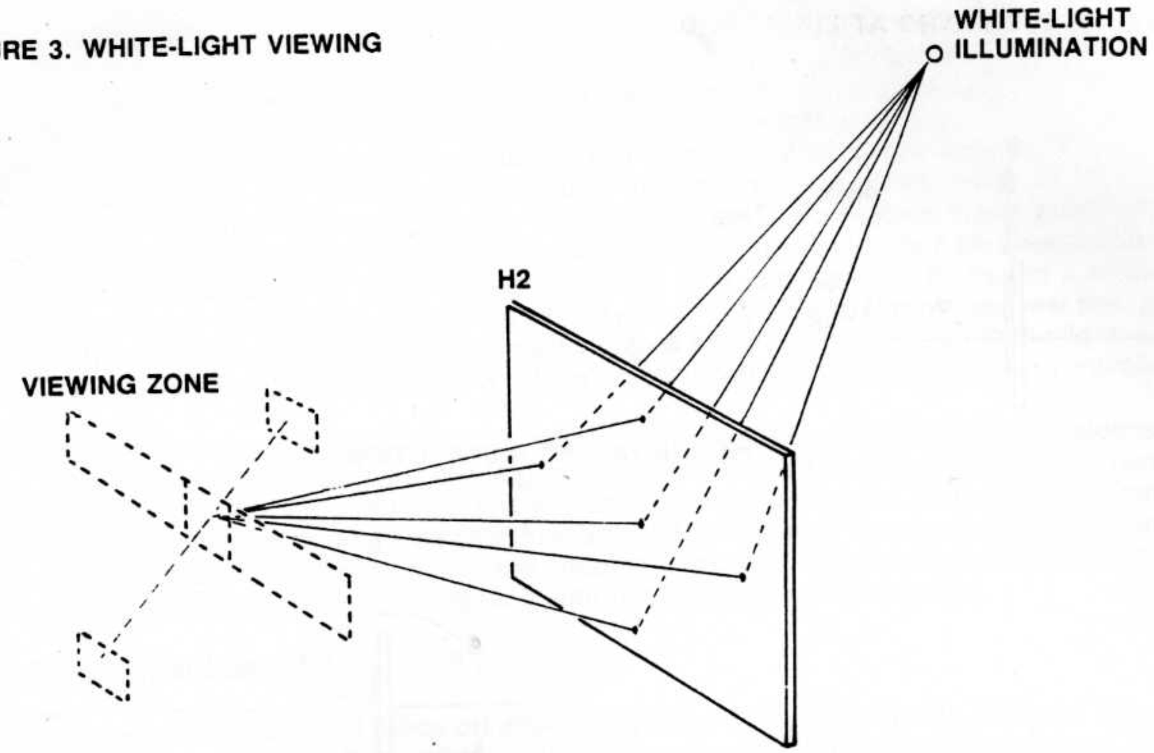

FIGURE 4. ASTIGMATIC FOCUS

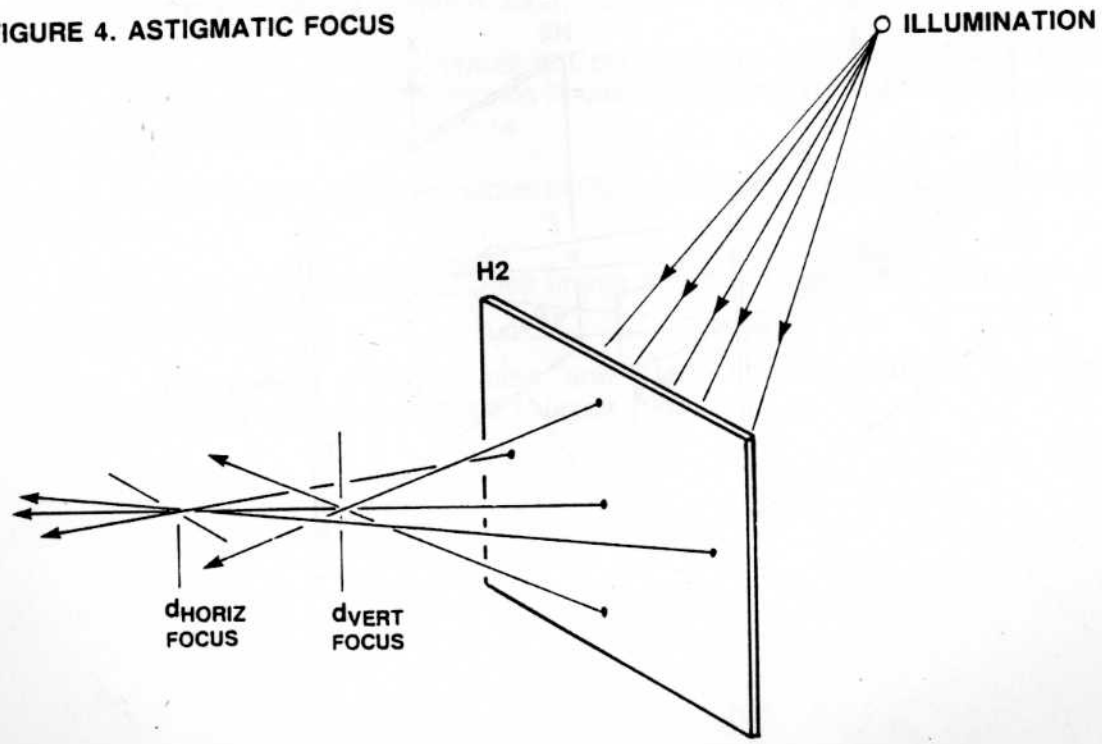

FIGURE 5. EFFECT OF PLATE TIP ANGLE

λ = 633nm