



Fresnel diffraction with small apertures
by Andrew David Struckhoff

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in
Physics

Montana State University

© Copyright by Andrew David Struckhoff (1991)

Abstract:

Scalar diffraction theory has been shown to produce accurate results in both the far-field (Fraunhofer) and near-field (Fresnel) regions.^{7,8,11} The derivation of the theory leads to three possible results of what is known as the obliquity factor. Light from a laser was allowed to shine through a small aperture (circular hole). By investigating the appropriate region of the resulting patterns, it was hoped that the proper obliquity factor could be discerned. Through this study, experimental results have indicated that when the aperture becomes small enough, scalar diffraction theory fails to accurately predict the location of the Fresnel patterns occurring in the (very) near-field region. Rather than verifying the correct obliquity factor, this thesis provides an indication of the region where scalar diffraction theory is no longer valid.

FRESNEL DIFFRACTION WITH SMALL APERTURES

by

Andrew David Struckhoff

**A thesis submitted in partial fulfillment
of the requirements for the degree**

of

Master of Science

in

Physics

**MONTANA STATE UNIVERSITY
Bozeman, Montana**

December 1991

11378
St 893

APPROVAL

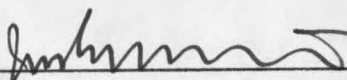
of a thesis submitted by

Andrew David Struckhoff

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

11-19-91

Date

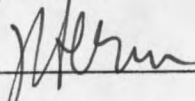


Chairperson, Graduate Committee

Approved for the Major Department

11-19-91

Date

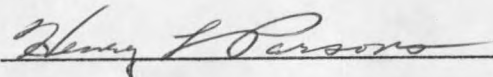


Head, Major Department

Approved for the College of Graduate Studies

11/27/91

Date



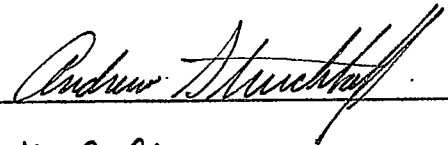
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under the rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Dean of Libraries when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature



Date

11-19-91

ACKNOWLEDGEMENTS

I would like to thank several people for their assistance throughout this project. first, Dr. Hal Kraus at Idaho National Laboratory, without whose inspiration and support this project would not have been possible. Next, I am deeply indebted to Dr. John Carlsten for his assistance and support throughout this project. His guidance has prompted me to want to learn about physics, not just to learn as needed. Through his experience, I have learned more about the process of experimentation than can be learned in any class. Thank you John. Also, the graduate physics students who provided questions which enabled me to think about the problems which I faced throughout the past two and one half years. I would also like to thank the secretarial staff at the MSU Physics Department for their assistance with the word processor and their invaluable patience. This author would also like to thank the other committee members, Dr. Rufus Cone and Dr. Hugo Schmidt, for their comments and questions regarding the improvement of this thesis.

Final thanks go to my family. To my wife, Ginger, who stood by my side, took care of me, provided a voice of reason in times of turmoil, and simply was present in times of frustration, I am deeply gracious. To my parents, for their support and encouragement. And, a warm thanks to my father, who not only provided a sense of scholarship and importance in education, but whose fight with cancer has taught me more about appreciating life than can be learned out of any textbook.

TABLE OF CONTENTS

	Page
1. INTRODUCTION.....	1
2. THEORY.....	4
Beam Propagation.....	5
Diffraction Integral.....	5
Obliquity Factor.....	6
Fresnel Equation.....	8
3. EXPERIMENT.....	9
Initial Concepts.....	10
Need for Magnification.....	11
Imaging the Aperture.....	15
Explaining the Edge Vibration.....	16
Refined Experiment.....	22
4. RESULTS.....	27
Diffraction Equation Comparisons.....	29
Fresnel Equation Comparisons.....	38
5. DISCUSSION.....	44
APPENDICES.....	47
Appendix A - Derivation of Scalar Diffraction Theory.....	48
Appendix B - Study in Spherical Aberrations/Computer Simulation.....	67
Appendix C - Study in Spherical Aberrations Observations using a Twyman-Green Interferometer.....	72
Appendix D - The Convolution Integral: Calculating the Response Function.....	77
Appendix E - Source Code for Convolution Integral.....	87
Appendix F - Computing the Beam's Radius of Curvature.....	93
REFERENCES CITED.....	100

LIST OF FIGURES

Figure	Page
1. Huygens's reconstruction for propagating a wave.....	5
2. Geometry for Diffraction by a Plane Screen.....	6
3. Geometry for the Derivation of the Fresnel Equation.....	7
4. A Fresnel Diffraction Experiment.....	10
5. Cross Section of a Gaussian beam.....	11
6. Cross Section of a Fresnel Ring Pattern.....	12
7. Experimental Apparatus for Imaging the Fresnel Patterns.....	13
8. Image of a Magnified Fresnel Pattern.....	14
9. Expected Image of Aperture.....	15
10. Actual Image of the 250 Micron Aperture.....	16
11. Fourier Construction of a Square Wave.....	17
12. Simulated Edge Reflection of Light.....	19
13. Details of the Aperture Mask.....	21
14. Revised Apparatus.....	22
15. Final Image of the Aperture.....	24
16. Cutting the Casing off of the Microscope Objective.....	26
17. Equation 2:1 Comparison, Fresnel Number 6.....	29
18. Equation 2:1. Comparison, Fresnel Number 8.....	30

LIST OF FIGURES (continued)

Figure	Page
19. Equation 2:1 Comparison, Fresnel Number 10.....	31
20. Equation 2:1 Comparison, Fresnel Number 12.....	32
21. Equation 2:1 Comparison, Fresnel Number 14.....	33
22. Equation 2:1 Comparison, Fresnel Number 16.....	34
23. Equation 2:1 Comparison, Fresnel Number 18.....	35
24. Equation 2:1 Comparison, Fresnel Number 20.....	36
25. Equation 2:4 Comparison, 250 micron Aperture.....	38
26. Equation 2:4 Comparison, 250 micron Aperture (expanded).....	39
27. Equation 2:4 Comparison, 100 micron Aperture.....	40
28. Equation 2:4 Comparison, 50 micron Aperture.....	41
29. Equation 2:4 Comparison, 25 micron Aperture.....	42
30. Equation 2:4 Comparison, All Apertures.....	43
31. Geometry used in Green's Theorem.....	51
32. Kirchhoff Formulation for Diffraction.....	54
33. Point Source Illumination of a Plane Screen.....	57
34. Effects of Obliquity Factor.....	59
35. Sommerfeld's Dual Green's Function.....	61
36. Geometry for Derivation of the Fresnel Equation.....	65

LIST OF FIGURES (continued)

Figure	Page
37. Spherical Aberration.....	69
38. Comparison Between Spot Size and Gaussian Waist.....	71
39. Twyman-Green Interferometer.....	74
40. Image of Spherical Aberration Fringes.....	76
41. One Dimensional Convolution.....	79
42. Geometry for Calculating the Response Function.....	80
43. Plot of Response Function vs Entrance F-Number.....	84
44. One Test Target Line Set.....	85
45. Computer Simulated Convolution of Test Pattern.....	85
46. Experimental Image of Test Pattern.....	86
47. Source Code for Computing the Convolution Integral.....	88
48. Geometrical Considerations for Altering a Beam.....	95
49. Propagation of a Ray by Matrices.....	97
50. Matrices Applied to a Gaussian Beam.....	98

ABSTRACT

Scalar diffraction theory has been shown to produce accurate results in both the far-field (Fraunhofer) and near-field (Fresnel) regions.^{7,8,11} The derivation of the theory leads to three possible results of what is known as the obliquity factor. Light from a laser was allowed to shine through a small aperture (circular hole). By investigating the appropriate region of the resulting patterns, it was hoped that the proper obliquity factor could be discerned. Through this study, experimental results have indicated that when the aperture becomes small enough, scalar diffraction theory fails to accurately predict the location of the Fresnel patterns occurring in the (very) near-field region. Rather than verifying the correct obliquity factor, this thesis provides an indication of the region where scalar diffraction theory is no longer valid.

CHAPTER 1

INTRODUCTION

The propagation of an electromagnetic wave past an opaque body casts an intricate shadow made up of bright and dark regions quite unlike anything one might expect from the tenets of geometrical optics.¹ This phenomenon of diffraction contradicted the extremely popular belief that light propagated as a ray in a single direction. The first documentation of diffraction was presented by Grimaldi, in a book published two years after his death (1663).^{1,2,4} Over the next 100 years, substantial evidence was found supporting Grimaldi's discoveries, however no real explanation of the observed phenomena was advanced.⁴ Huygens, the first proponent of the wave theory, seems to have been unaware of Grimaldi's findings, as he would have relied on them to support his ideas.² It was not until 1818, when Fresnel applied both Huygens' construction for propagating a wave and the principle of interference to arrive at a possible explanation for diffraction.² This combination was given a firm mathematical basis by Kirchhoff in 1882, and ever since, the subject has been investigated by many writers.¹⁻¹¹

The phenomenon of diffraction, and the mathematics which support it, has been

tested extensively over the past century.^{7,8,11} As with most practical problems, though, typical solutions still involve some assumptions and approximations. The resulting mathematics depends solely on what approach one takes to solve the problem and the validity of the assumptions. Kirchhoff's classical theory has been shown to work quite well in describing the diffraction field when relative dimensions of the experiment are much larger than the incident wavelength.^{7,8} However, the theory is often criticized, as the solutions to his theory do not recover the assumed boundary conditions at the diffracting aperture.⁶ Since Kirchhoff's initial formulation in 1882, there have been two other formulations for describing diffraction theory which do not encounter mathematical difficulties. These are the formulations by Rayleigh and Sommerfeld, involving slight alterations from Kirchhoff's theory. For a majority of the diffraction problems encountered, the Kirchhoff formulation seems to be appropriate, as the mathematical inconsistencies do not contribute to the final solution.^{6,7,8}

Chapter 2 gives a summary of the mathematical derivation which Kirchhoff presented, and the subsequent additions given by Rayleigh and Sommerfeld. While the phase relationships of the propagating wave remain the same in all of the derivations, the obliquity factors differ. Recently, Kraus^{9,10} has studied the differences in the obliquity factors in order to test the validity of each derivation. This experiment was conceived in an effort to determine which factor is indeed correct. The resulting area of investigation provided interesting results.

Chapter 3 describes the process of the experimental studies. As is often the case, after the initial conceptualization, many refinements had to be made in order to obtain

the required data. Throughout the process of development, other related areas of optics were studied. These studies are presented as separate topics and can be found in the appendices.

Chapter 4 presents the results of the experiment. The results go beyond the initial goal of the project. Aside from testing the diffraction integral, an interesting area of diffraction theory was exposed. The experiment was found to probe the region where scalar diffraction theory itself is seemingly no longer valid. This occurs in areas close to the diffracting aperture. In this region, vector diffraction theory is probably needed. Several references appear to indicate that scalar theory will start to fail in the region being probed; however, no definite calculations have been done. The results presented here may give a point of reference as to where scalar diffraction theory will begin to fail.

CHAPTER 2

THEORY

After the discovery of diffraction, the evolution of a quantitative theory did not start until Huygens' principle was combined with wave interference. Christian Huygens hypothesized that as a wave propagates, each point along the wavefront acts as an emitter of secondary sourcelets.⁶ The sum of these secondary sourcelets represents the motion of the wave. Figure 1 shows a typical Huygens' construction. By taking relative phases into account, Fresnel was able to accurately predict light distributions using this wave construction method.

The problem of diffraction is that of light impinging on a screen with a specified aperture. On the "shadow" side of the screen, geometrical optics would predict a cone of light, governed by the properties of the light rays which enter the aperture. However, for a circular aperture, physical observations yield a pattern of concentric rings. Kirchhoff's derivation in 1882 was the first true mathematical explanation for these rings. The derivation can be found in almost any basic optics book.^{1-3,6} Appendix A reviews the complete derivation of scalar diffraction theory, while the results are simply presented in this chapter. Figure 2 shows the geometry to be used for the problem.

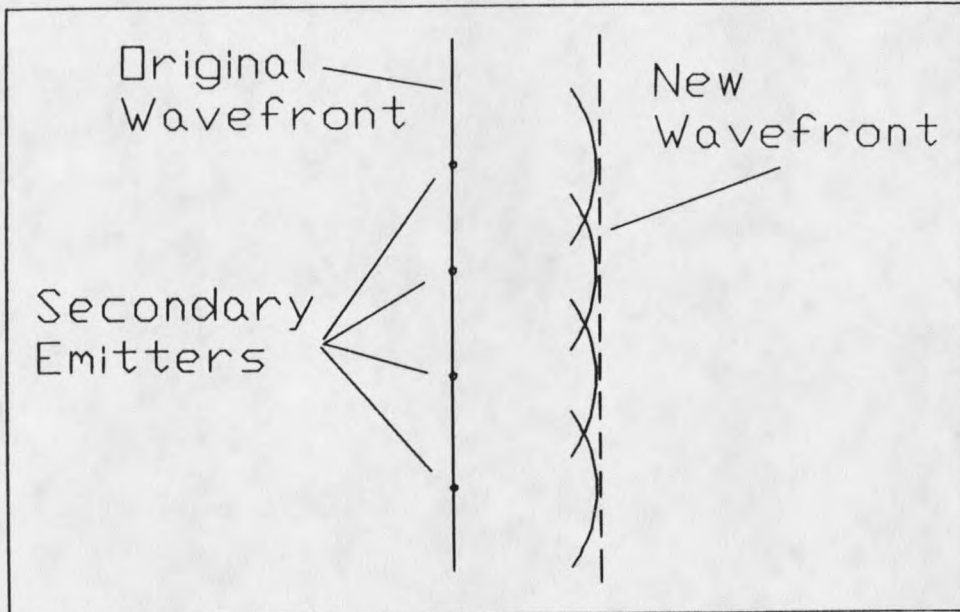


Figure 1 Huygens' construction for propagating a wave

Consider a point source of light located to the left of a screen with an aperture. An accurate description of the light distribution on the right side of the screen is needed. The observation point is denoted as P_o on the right of the screen.

Using Huygens' construction, the sourcelets are placed in the plane of the aperture. The total distribution at the point P_o can be attained by adding all of the relative amplitudes and phases of the sourcelets oscillating in the aperture plane. The equation which describes the field at the point P_o is called the scalar diffraction formula and is written as

$$U(P_o) = \frac{1}{i\lambda} \iint_{\Sigma} \frac{\exp[ik(r_{21} + r_{o1})]}{r_{21}r_{o1}} \phi \, ds \quad (2:1)$$

The integral occurs over the opening of the aperture Σ , and, r_{o1} (r_{21}) is the distance from

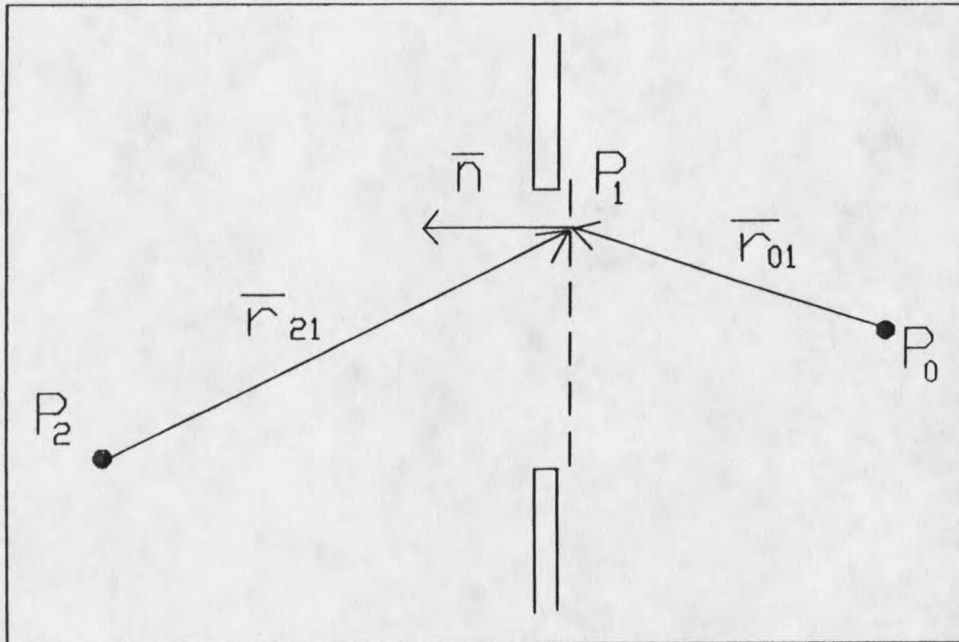


Figure 2 Geometry for diffraction by a plane screen

the observation (source) point to a point in the aperture plane. In essence, the integral adds all the amplitude and phase contributions from a wavelet which originated at the point P_2 and ends up at the point P_0 . The difference in pathlength now depends on which point, P_1 , in the aperture, is used as the intermediate point. The term ϕ is referred to as the *obliquity factor*. The obliquity factor depends upon how equation (2:1) is obtained. In appendix A, it is shown how three possible obliquity factors exist. They are

$$\begin{aligned}
 \phi &= 1/2[\cos(\bar{n}, \bar{r}_{01}) - \cos(\bar{n}, \bar{r}_{21})] \quad (\text{Kirchhoff}) \\
 \phi &= \cos(\bar{n}, \bar{r}_{01}) \quad (\text{Rayleigh-Sommerfeld I}) \\
 \phi &= \cos(\bar{n}, \bar{r}_{21}) \quad (\text{Rayleigh-Sommerfeld II})
 \end{aligned}
 \tag{2:2}$$

where the term $\cos(\bar{n}, \bar{r}_{01})$ ($\cos(\bar{n}, \bar{r}_{21})$) refers to the angle between a) the vector pointing

from the observation (source) point to a point in the aperture plane and b) the normal vector of the aperture plane. At this point, an appropriate question would be; "Which obliquity factor is correct?". When the source and observation points are far away from the aperture, the cosine terms in each obliquity factor appear as

$$\cos(\vec{n}, \vec{r}_{01}) \approx 1 \quad \cos(\vec{n}, \vec{r}_{21}) \approx 1 \quad (2:3)$$

and all three factors will equal 1. Kraus, however, observed that as the angles increase, the different obliquity factors will produce noticeable differences in the diffraction fields.

Another interesting equation can be derived by confining the points P_0 and P_2 to the axis of symmetry for the problem of a circular aperture. The point P_0 will essentially "see" two different distances to: a) the center of the wave along the axis, and b) the edge of the aperture. The geometry for this situation is depicted in Figure 3.

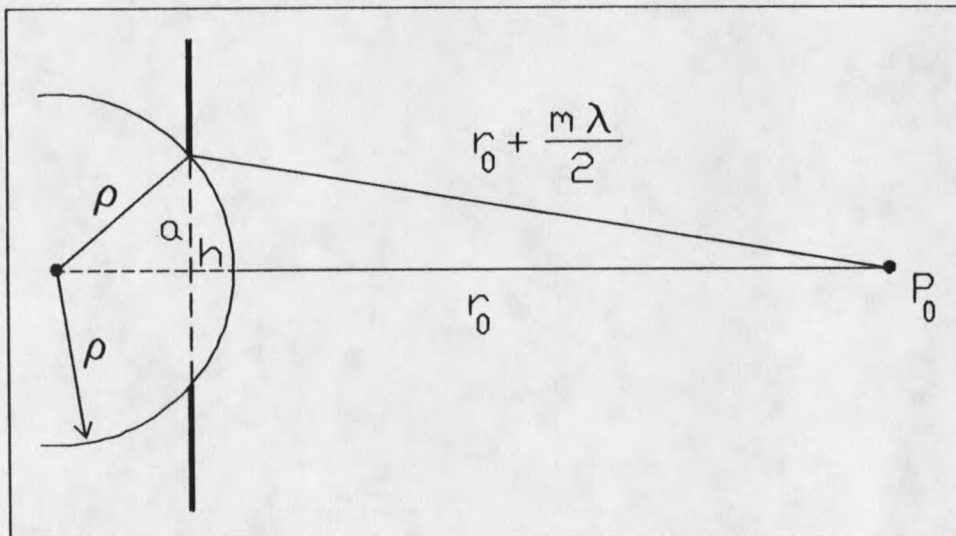


Figure 3 Geometry for the derivation of the Fresnel equation

Notice that the difference in pathlength is what matters here. When the net difference

in pathlength is an even number of half-wavelengths, all of the phases across the wavefront cancel out and there is a dark spot along the axis. One can obtain a black spot by shining light through a hole! When the net difference in pathlength is an odd number of half-wavelengths, there will be a bright spot. This effect has been confirmed in several experiments.^{1-4,7,8,11} When one of these distinctive patterns occurs, it is known as a Fresnel pattern. The patterns are numbered according to the integer number m in Figure 3. Utilizing the geometry of the problem one arrives at an expression for the location of the Fresnel patterns as a function of the experiment's geometry;

$$r_o(m) = \frac{\lambda^2 m^2 / 4 - a^2}{a^2 / \rho - m\lambda} \quad (2:4)$$

This is known as the Fresnel equation.

Chapter 3 describes the experiment which was done in an effort to test equations (2:1) and (2:3). Following Kraus' predictions, the appropriate geometry (aperture radius a , beam radius ρ , and observation point r_o) was investigated to find out which obliquity factor in equation (2:1) was indeed correct. Along the way, equation (2:3) was utilized as the first test that scalar diffraction theory was indeed working. Comparing the complicated intensity patterns produced with theory (equation (2:1)) was done at Idaho National Engineering Laboratory. The comparison of the position data was much less involved in terms of computer time, so equation (2:3) served as the initial requirement for testing the theory.

CHAPTER 3

EXPERIMENT

In typical experiments, one usually begins with an initial concept of how the experiment should work, and then the apparatus is refined until the required precision is attained. This experiment progressed in precisely that fashion. Once it was realized what sort of precision was needed, the proper instrumentation was procured and the data was taken. Several additional studies were done in an effort to understand what factors affected the precision of the experiment. These studies included; Spherical Aberrations (Appendices B and C) and Image Formation (Appendices D and E). As the studies were completed, the experiment was altered to accommodate these new considerations.

The basic diffraction experiment is pictured in Figure 4. The nature of the Fresnel diffraction required a coherent light source. A 5 mW helium-neon laser was employed for this purpose. The beam was filtered to remove the unwanted spatial "noise". The beam, prior to the spatial filter is scattered many times from dust particles inside of the laser cavity and in the air itself. These dust particles cause a distinct non-uniformity to occur on the beam. The spatial filter focuses the beam through a small pinhole. The event of focusing separates the desired beam from the unwanted noise.

The pinhole simply selects the beam and eliminates the noise. After the spatial filtering, the beam impinged on a circular aperture, 250 microns in diameter. In accordance with the geometry in Figure 4, the distribution of the light on the right side of the aperture was the area of interest.

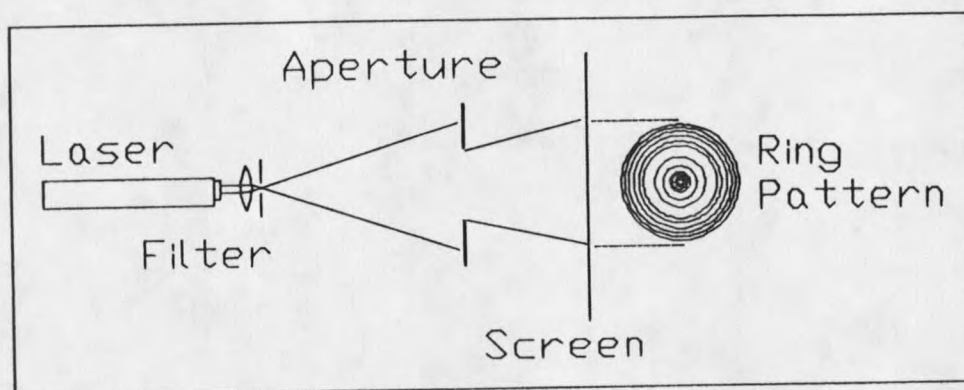


Figure 4 A Fresnel diffraction experiment

In order to record the diffraction fields, some sort of detection device was needed. A Reticon photodiode array was used for detecting the intensity distribution. The array consisted of 1024 photodiodes, with 15 micron widths, placed on 25 micron centers. Support electronics assisted in retrieving information from the diodes. Photodiodes are light sensitive, and each diode converts the amount of light it absorbs into a scaled voltage. Then, over a short period of time, the electronics outputs the voltages from the diodes as a voltage string. This voltage string was then sent to a Tektronix 2230 storage oscilloscope. The oscilloscope has the capability of storing the information which appears on the screen. The oscilloscope was also equipped with a General Purpose Interface Bus (GPIB). Using the GPIB, the images from the oscilloscope were transported to an AT&T PC 6300 PLUS personal computer and stored

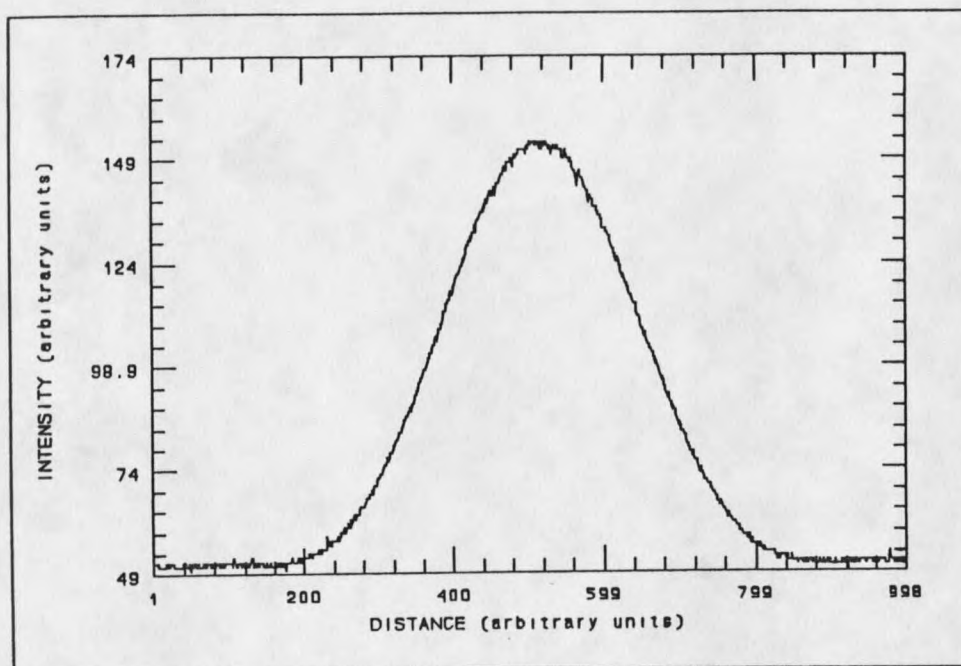


Figure 5 Cross-section of a Gaussian beam as obtained through the photodiode array

as data files. A resident plotting program was used to convert the data files into graphs. Figure 5 shows how a simple gaussian beam profile recorded by the electronics. This pattern is of a beam which exits the spatial filter only. There is no aperture being used in obtaining this image.

Considering the relative dimensions of the aperture and array, the detection of a diffraction pattern would result in the use of only 10-20 of the 1024 available photodiodes. Figure 6 shows an image of a pattern obtained by placing the diode array extremely close to the aperture. The image is that of a slice across the diameter of a Fresnel ring pattern. The resolution of the diode array is obviously not good enough on its own, as the detection of even more complex patterns was needed.

At this point, a lens was utilized to produce an enlarged intensity pattern across

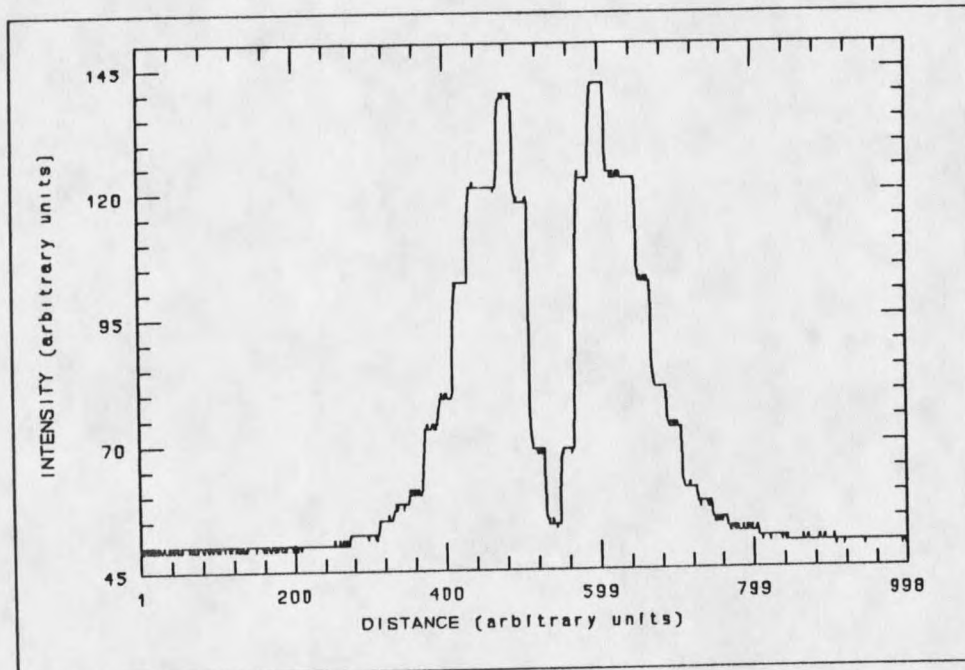


Figure 6 Image of a Fresnel ring pattern

the diode array. It has been shown that the use of a lens to image Fresnel patterns could be done successfully.¹¹ By enlarging the pattern, more diodes were used to detect the diffraction field, which greatly enhanced the performance of the system. In order to gather information about the different ring patterns, it should be realized that the imaging system would now have to act as a "scanner" to obtain the different Fresnel patterns, as each pattern occurs at a distinct distance away from the aperture. Figure 7 shows the experimental setup at this point. Both the imaging lens and diode array were placed on X-Y-Z translation stages to facilitate ease in system alignment. In order to keep the magnification of the patterns the same, the object and image distances had to be kept the same, in accordance with

$$M = \left| \frac{d_i}{d_o} \right| \quad (= \frac{\text{image distance}}{\text{object distance}}) \quad (3:1)$$

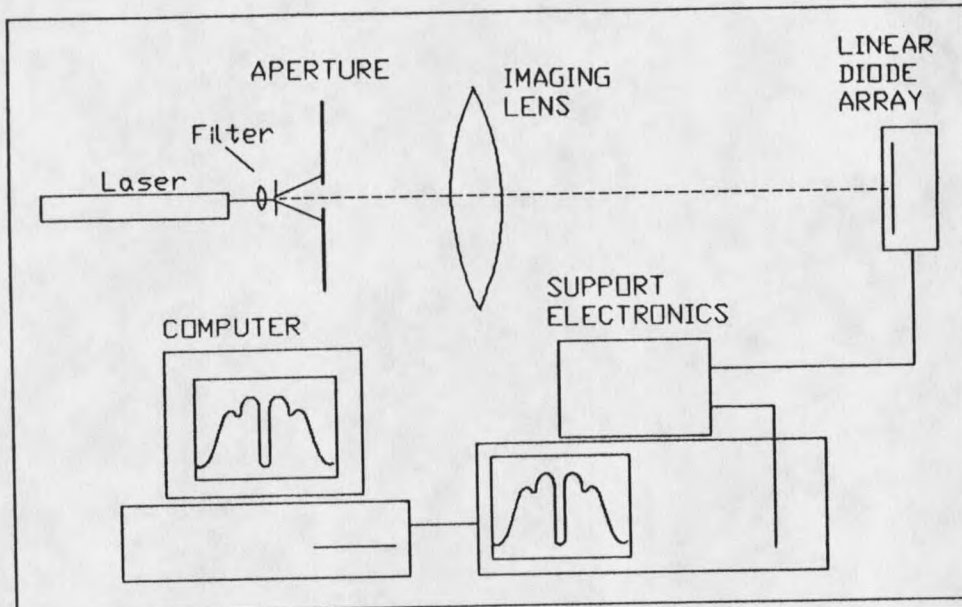


Figure 7 Experimental apparatus for imaging the Fresnel patterns

This was accomplished by fixing the relative distance between the lens and the diode array. By moving the two together, the object distance of the imaging system was kept constant according to

$$\frac{1}{d_o} = \frac{1}{f} - \frac{1}{d_i} \quad (\text{the thin lens equation}) \quad (3:2)$$

where f is the focal length of the lens. Since both the image and object distances were fixed, the magnification remained constant. Now, different "slices" of the diffraction

field could be selected. By "scanning" along the system's axis, one could obtain images of the various ring patterns. Figure 8 shows a pattern attained using a 5 cm focal length imaging lens configured for a magnification factor of 5. It should be noticed that the curve is now much smoother than that in Figure 6, obtained without the magnifying lens.

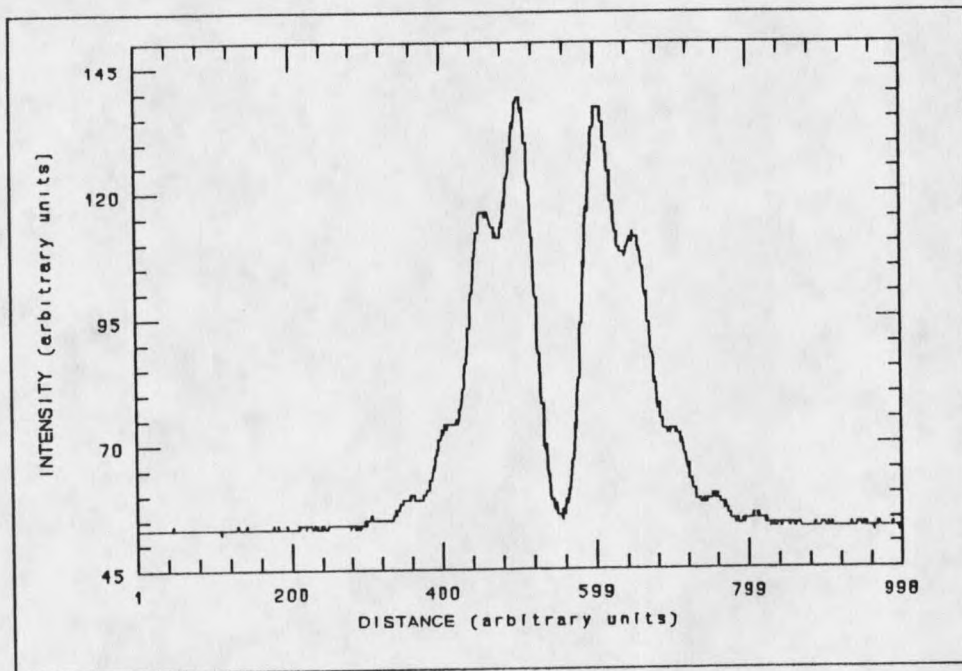


Figure 8 Image of a magnified Fresnel pattern

Inserting the lens into the system indeed showed a great improvement to the resolution of the system.

In order to test the equations given in chapter 2, it was necessary to know which "slice" of the diffraction field was being imaged. By first obtaining an image of the hard edge of the diffracting aperture, the imaging system could be "zeroed"; then by moving the lens-Reticon system away from the aperture, different slices of the field could be imaged. As it was virtually impossible to place the scanning apparatus in a perfect

position, two translation stages were utilized. Setting one translation stage to read zero, the second could be moved to the point where the aperture was being imaged, the second translation stage was then locked in that position. The imaging system was then moved away from the aperture using the first translation stage. Now, the reading on the first stage would be the distance away from the aperture at which the image on the diode array was occurring. When the time came to attempt the imaging of the actual aperture, it was found that several difficulties were encountered.

It was hoped that the image of the aperture should resemble a top-hat type pattern, as shown in Figure 9.

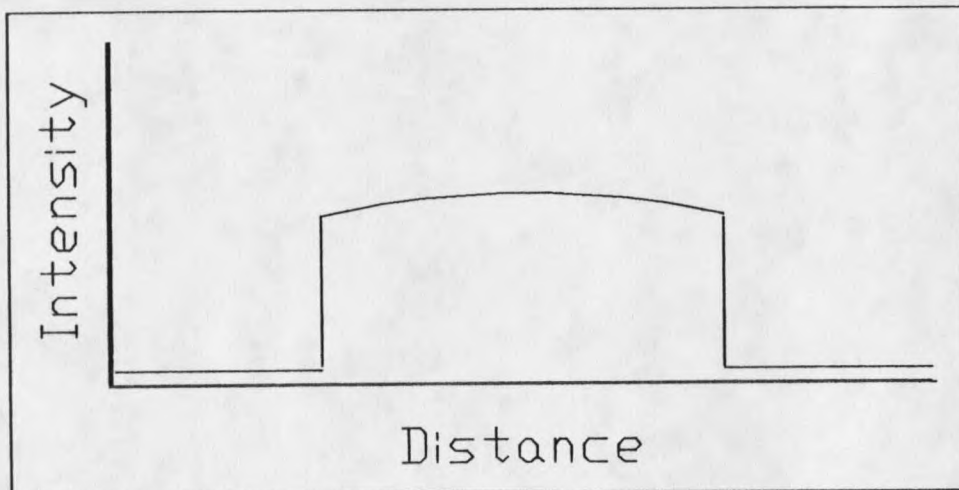


Figure 9 Expected image of aperture with slight gaussian profile

This should seem reasonable, as by imaging the plane of the aperture, intensity should exist inside the aperture, but none should exist outside of the aperture.

With the He-Ne beam being of a gaussian profile, a slight variation in intensity was expected. Other than that, it was expected that a fairly constant intensity pattern could be obtained without much trouble. Unfortunately, an image similar to Figure 10

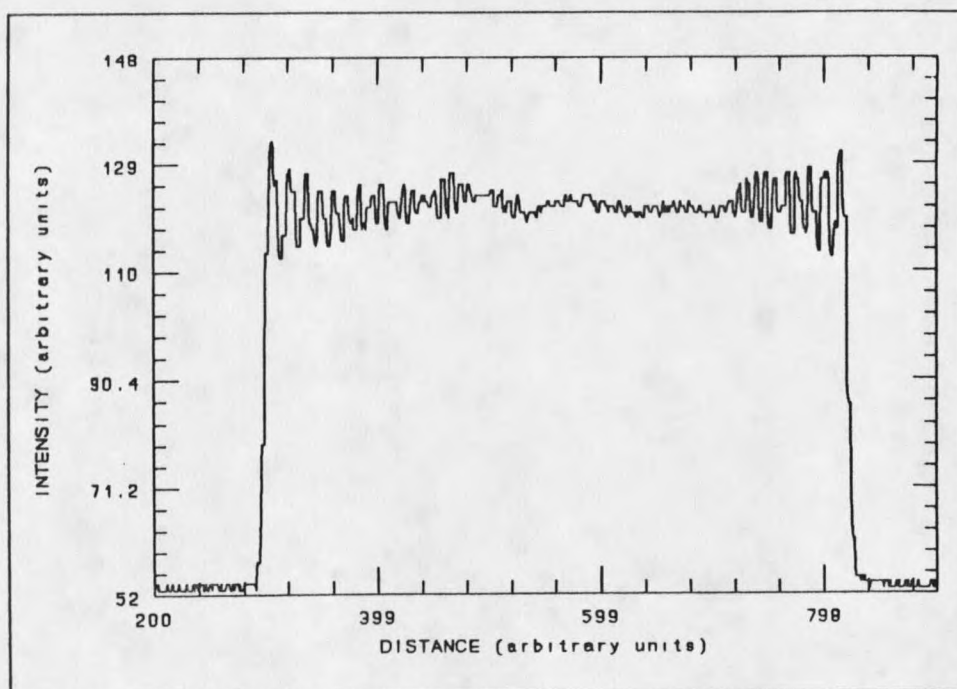


Figure 10 Actual image of the 250 micron aperture

was the best image which could be obtained. The image does have the expected sharp edges, however, there are some extra details at the edge of the image which are disturbing. A couple different possibilities exist as reasons for this "edge vibration".

The first possibility was that of the omission of certain spatial frequencies in the reconstruction of the image. This interesting detail comes from the fact that in imaging the top-hat pattern, two separate Fourier transforms are done. The first Fourier transform involves the lens collecting the object for transmission, the second involves the lens creating the image.

When successive Fourier transforms are done, one expects to arrive at the same function. However, if certain terms are omitted from the Fourier reconstruction, the initial function will not be properly reconstructed. For example, a square wave can be

expressed as:

$$\sum_{n=1}^{\infty} \frac{\sin((2n+1)x)}{(2n+1)} \quad (3:3)$$

If, for some reason, the sum is not infinite, higher terms will be left off and the square wave will instead appear similar to Figure 11, where only 30 terms have been included.

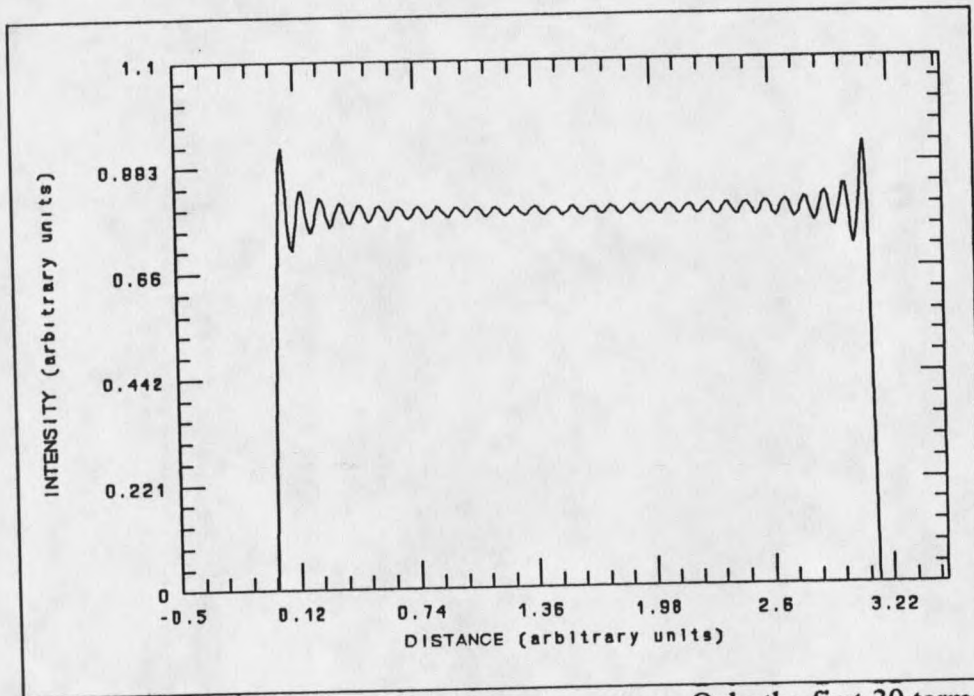


Figure 11 Fourier construction of a square wave. Only the first 30 terms are included

The vibrational pattern of this image looks similar to Figure 10. The solution to this problem was to obtain a lens which somehow "collected" more of the high spatial frequencies of the object in order to reconstruct a better image. It is known that the resolving power of an optical system depends strictly on the ratio of the lens' focal

length to its diameter.^{1,2,6} This ratio, f/d , is called the f-number of the system, and is often denoted as $f/\#$.

At this point, the details of a lens became extremely important. In order to study the effects of aberrations and $f/\#$ on the imaging system, a few side projects were done. Appendix B refers to a computer simulated study done on spherical aberrations. In this study, an optical ray tracing program, called Beam 3 was used to simulate the focusing of a beam. A measurement of aberration was made and this measurement was compared to what one would expect when focusing a real gaussian laser beam. Appendix C addresses the testing of lenses for spherical aberration. An appropriate interferometer was used as a simple device for observing spherical aberration. Appendix D discusses a study of the dependence of the system's resolution on its entrance $f/\#$. According to Fourier optics, the smaller the entrance $f/\#$ of a system, the better its resolution. Appendix E presents the computer program which was used to emulate an imaging system in accordance with Fourier optics. Through this program, comparisons between theory and experiment could be made to indicate the entrance $f/\#$ of the system.

Another possible explanation for the appearance of this edge vibration was that light was being scattered from the edge of the aperture. The light would then enter the aperture, where this scattered light interferes with the incident wave to form the observed pattern. The aperture used to obtain Figure 10 was simply a hole in a metal foil. The method for making this hole was simply to have a laser "punch" it. The process of making a hole in a foil takes a very concentrated amount of energy. Since the metal was melted in the process of creating a hole, it would seem appropriate that the edges of the

aperture may have also melted away to form a smooth curve. The aperture was placed under a microscope and indeed, the edge of the aperture had a curved surface. This curved surface was then programmed into the computer ray-tracing program. In this program, the curved surface was to act as a mirror in order to verify if it was possible for light rays to bounce off of the edge of the aperture and then enter it, causing an interference pattern. Figure 12 shows the results of this computer simulated reflection.

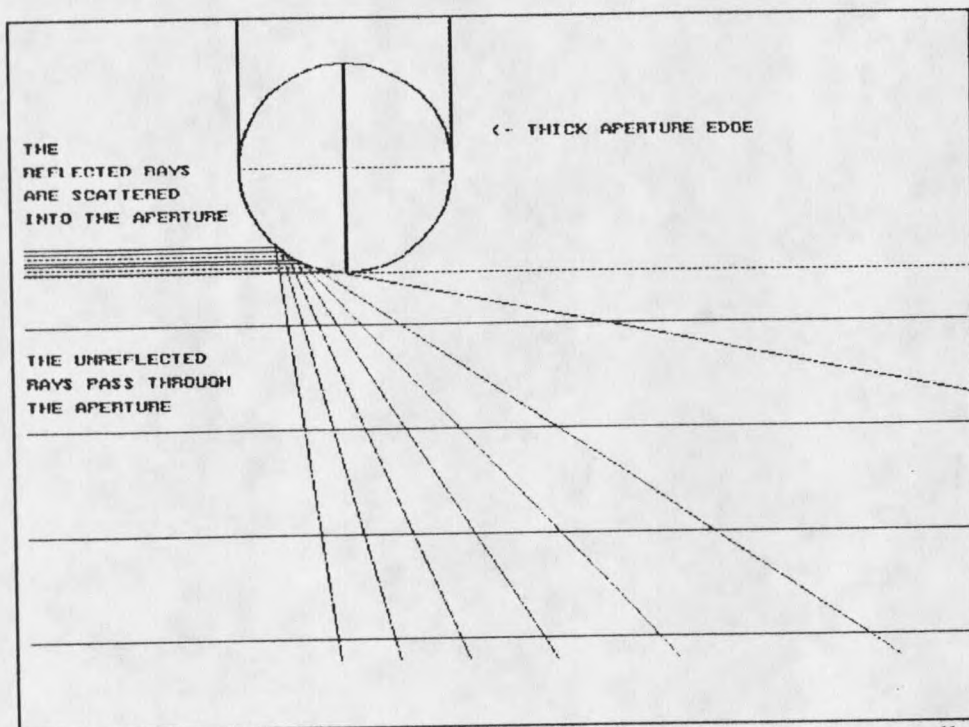


Figure 12 Simulated edge reflection of light which would not normally pass through the aperture

The aperture being used was about 20 wavelengths in thickness. It would seem possible that the edge vibration was caused by a reflection of a portion of the light into the aperture. To solve this problem, it was hoped that a "thin" aperture could be obtained. Several aperture companies were contacted, however, it seemed as though the thinnest

aperture to be obtained was almost 13 microns in thickness. When the wavelength being used in the experiment was .63 microns, this aperture would still be considered to be extremely thick. After a thorough investigation of foil apertures it was discovered that a process did exist which allowed a thin opaque substrate to be deposited on a glass plate.

After some additional investigation, a company was found which could provide a custom made aperture plate, or "mask". This mask would consist of a chromium deposit on a substrate of glass. The chromium deposit was to be 400 - 700 angstroms in thickness and have 5 apertures in its design. For reference, a 600 angstrom thick deposit would have a thickness of $\lambda/10$. Figure 13 shows the essential details of the mask. In order to prevent unwanted reflections from the surfaces of the glass, both sides of the mask received an anti-reflection coating. This coating was to have a reflectivity of less than 0.1% per side of glass at normal incidence.

As the experiment developed, both the quality of the lens and the thickness of the aperture were found to influence the results which could be obtained. A series of lenses were investigated in order to solve the problem of the entrance f/# of the system. Each lens which was investigated had two basic tests to pass. The first test was to consider its entrance f/#. The second test was to observe how much aberration occurred over the area of the lens. If there were aberrations over part of the lens, that part of the lens could not be considered as contributing to the overall resolution of the system. Thus, the advertised f/# was not necessarily the diffraction limited entrance f/#. Another consideration of the lenses to be investigated was the construction. At one

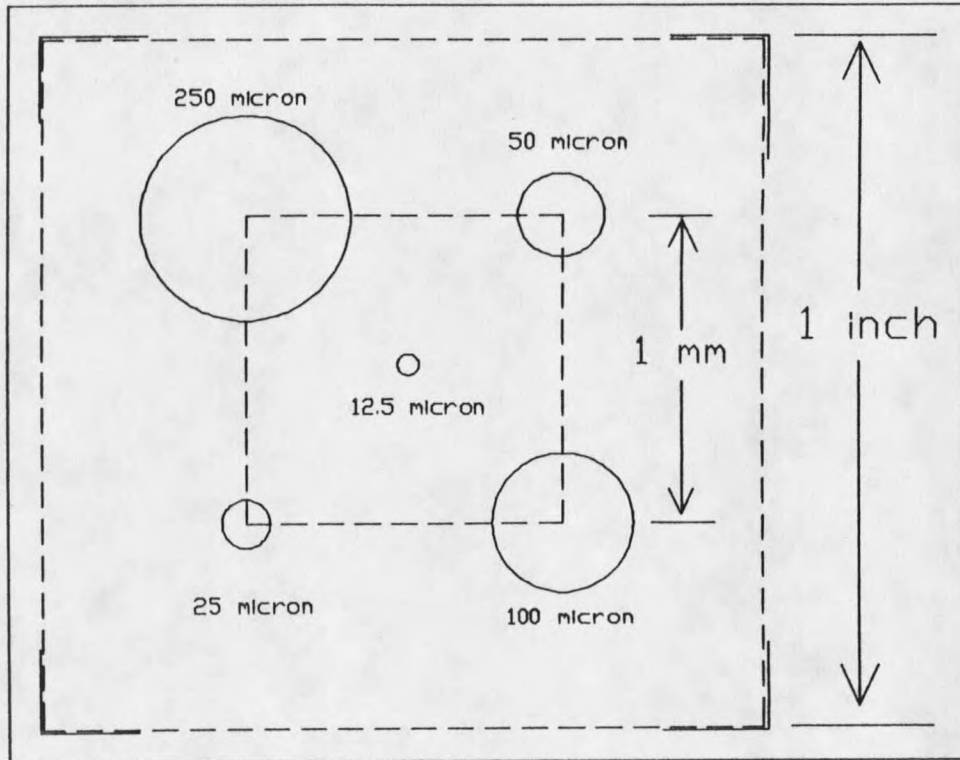


Figure 13 Details of the aperture "mask"

point, video camera lenses were considered, however, these lenses contained several separate elements and dust on each element became a problem. A lens with little or no disassembly for cleaning was preferred. The mask was the item which solved the problem of the thick aperture.

As the resolving power of the system was increased, it was found that dust became more of a problem. A dust-free environment was needed. In order to combat the ambient room dust getting into the system, a laminar downflow system was obtained. This downflow system or "hood" simply acts as a large dust filter which was suspended above the optical system. By enclosing the system, it was found that the dust in the lab was prevented from settling on the optics.

At this point, the experiment was fairly established. The lens which was to be used was a laser achromat. Its diameter was .5 cm and its focal length was 1 cm, leading to an entrance f/# of 2. The lens was coated on both sides with a High Efficiency Broad Band Anti-Reflection (HEBBAR) coating. The reflectivity of the two surfaces was about .5 % per surface. This caused no problems in the system. Figure 14 shows the system at this point. Two lenses were found to help increase the resolution of the system. The laser achromat was used to collimate the image and a 1 meter lens was used to re-focus the image. The magnification of the system was now the ratio of the focal lengths of the two elements, or 100.

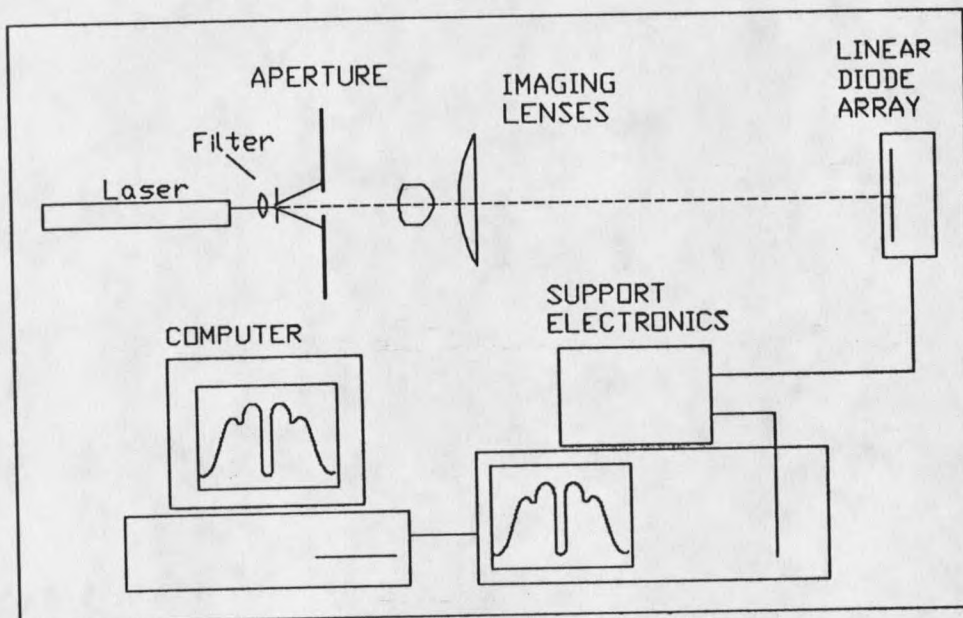


Figure 14 Revised apparatus with new imaging lenses

The mask had now been obtained. There were additional effects introduced by the glass substrate on which the mask was deposited, but these effects were not damaging to the experiment. Since the index of refraction of the glass was 1.515, it was

assumed at the time that the radius of curvature of the beam would be slightly altered by the propagation through the glass. This effect had to be accounted for, as the position of each Fresnel pattern depends upon the radius of curvature of the incident beam. In addition, a process for the measurement of the radius of curvature was needed. This was solved by the translation of the imaging system. It is known that the beam originates from essentially a point source located within the spatial filter of the system. By knowing the distance of separation between the spatial filter and the aperture, a measurement of the radius of curvature could be made. This measurement was made simply by imaging the spatial filter, then imaging the aperture and recording the separation distance. Extracting the radius of curvature from this information was done in two separate ways. These two ways are by simple geometry and by propagation matrices. These methods are discussed in Appendix F.

The mask helped to cure a problem which was mentioned earlier and caused an additional problem to be dealt with. There had been an edge vibration which was only partially explained. Part of the solution was the substitution of a new lens. But, the mask was able to improve the system even more. Figure 15 shows the top-hat pattern in its final state. As the aperture thickness was essentially zero, the small edge vibration can be attributed to the size of the lens. As the lens is not infinite, it will not capture all of the needed spatial frequencies, and the image will be left incomplete. However, this image does show an improvement from the initial thick aperture.

The additional problem which was caused by the mask arose from the chromium substrate. Essentially the light which does not propagate through the aperture is reflected

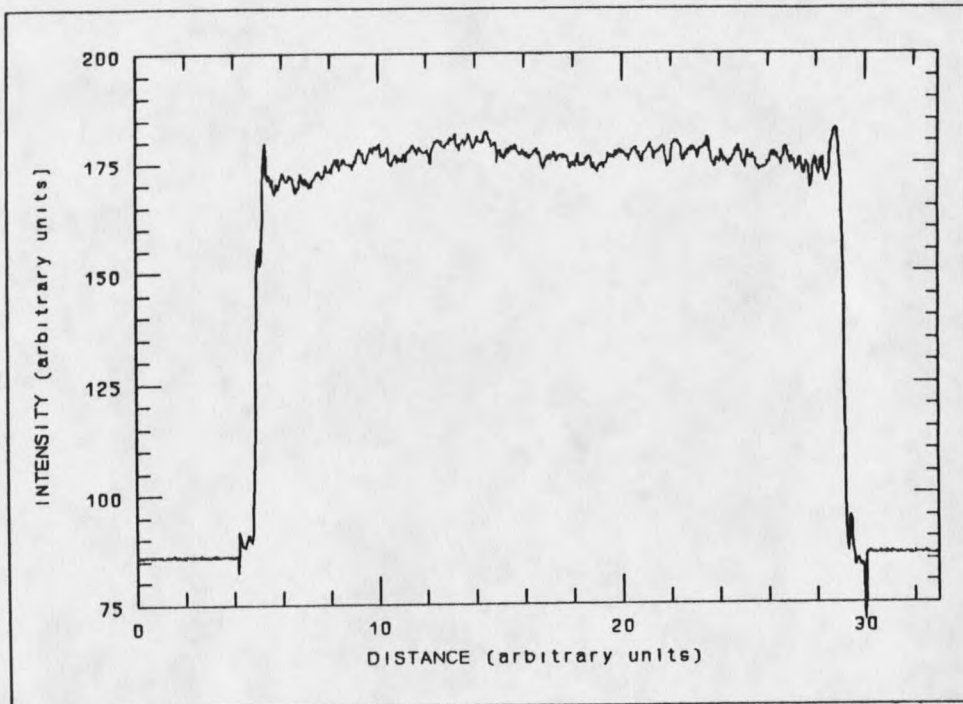


Figure 15 Final image of the aperture. Note the disappearance of the edge vibrations.

back towards the spatial filter. The filter itself is a small hole in a reflective foil. The combination of these two reflective surfaces caused a resonating cavity to occur. This caused additional interference patterns in the system. The solution of this problem was to cause a slight vibration in the system. If the system could be vibrated on a sub-wavelength scale, it would be enough to shift the fringes a very small amount. A high frequency vibration would essentially "wash out" these fringes, as the electronics used to obtain the intensity patterns could be programmed to "ignore" the high frequency alterations to the system. The sub-wavelength vibration was accomplished with the aid of a piezo-electric crystal mounted in a translation stage micrometer barrel. This piece of equipment was obtained through Burleigh Instruments. The crystal was rated at 35

microns of travel at 150 volts applied. A high-frequency, 2 volt pulsed signal was sufficient to eliminate the unwanted fringes.

Finally, useful experimental measurements could be made. The 250 micron aperture was employed for this purpose. The system was zeroed and the separation between the spatial filter and mask was measured. The optical system was scanned from a distance of 5.128 mm away from the aperture up to a distance of 1.176 mm away from the aperture. Along the way, the distinctive even number Fresnel patterns were observed and recorded. The detection of the Fresnel patterns occurred when the central dark spot produced the smallest voltage as measured by the oscilloscope. The even Fresnel patterns numbered 6 through 20 were observed and recorded. The field patterns were sent to Idaho National Engineering Laboratory for comparison to the field equation. The positions of the Fresnel patterns were also recorded and compared to the predictions of the Fresnel equation.

In order to investigate the 100 micron, 50 micron and 25 micron apertures, the system had to be further refined. Instead of using the laser achromat as the initial lens in the system, a microscope objective was used. The objective had a much smaller $f/\#$ than the achromat, thus it gave a better resolution. The optics inside the objective were covered with a protective casing. In order to improve the imaging quality, this casing was removed to expose more of the lens. Figure 16 shows the result of cutting away the casing. The entrance $f/\#$ of the lens was now .65, a factor of 4 improvement from the laser achromat, as the achromat was found to be diffraction limited only to an $f/\#$ of 3. The improvement was confirmed both by the interferometer (Appendix B) and by test

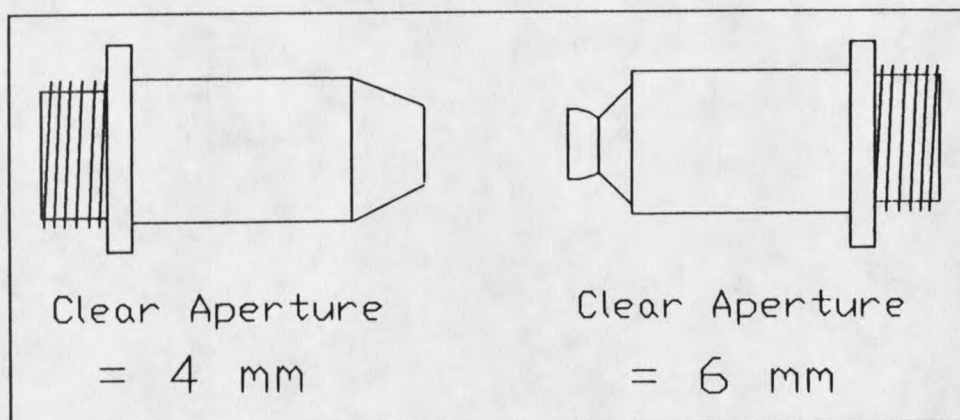


Figure 16 Cutting the casing off of the microscope objective to increase the area of the lens to be used

target imaging (Appendix D). The focal length of the objective was given as 3.8 mm, which led to a magnification of 256.

Unfortunately, the anti-reflection coating on the objective was not as efficient as the coating on the achromat. The unwanted reflectivity caused additional interference patterns to occur, thus the field distribution could not be accurately measured. However, the locations of the Fresnel patterns could be measured with a very high accuracy. To obtain the necessary accuracy in position measurement, differential micrometers were used. The smallest increment on the micrometer was .5 micron, however only 400 microns of travel at that accuracy was attainable. The limitation of micrometer travel did not hamper the measurements. The 250 micron aperture was also investigated at this resolution, so that a region closer to the aperture could be probed.

Chapter 4 presents the results of these measurements. There are some deviations from the standard diffraction theory. The deviation is the occurrence of a shift in the location of the Fresnel patterns which can be observed in the smaller apertures.

CHAPTER 4

RESULTS

The results of the experiment will reflect comparisons between theory and experiment based on the two equations given in Chapter 2. The initial purpose of the experiment was to confirm which of the obliquity factors in the field equation is the correct obliquity factor. Kraus provided the comparisons for the field equation, as the necessary computer programs for generating the field patterns were not available to the author. The Fresnel equation comparisons were done at Montana State University with the assistance of a resident plotting program, as the equation itself is fairly simple.

In order to generate the proper curves for comparison, the true size of each aperture had to be known. The mask specifications called for 250 micron, 125 micron, 50 micron, 25 micron and 12.5 micron diameter apertures. However, initial data/theory comparisons for the fresnel equation seemed to indicate otherwise. The apertures were apparently not the requested size. The sizes had to be confirmed to within a few percent in order for the data to be valid.

Several methods were employed to measure the size of each aperture. The first method was by utilizing a travelling microscope. However, only the 250 micron aperture

could be measured in this fashion, as parallax and resolution problems greatly limited the effort. The next instrument to measure the apertures was a Confocal Scanning Laser Microscope. This microscope advertised a much better resolution in addition to being able to measure extremely small details. Again, though, there were details which could not guarantee a distinct measurement. The final method to be tried, and the method to work, was to utilize the USAF Test Target (Appendix D). Documentation of the line-pair separations accompanied the target. By magnifying several line pairs, an accurate value of the magnification could be obtained. Then by imaging each aperture on the mask, the actual size of the aperture could be deduced from the magnification and the image size. The actual diameters of the apertures were found to be: 250 microns $\pm$ 5 microns, 98.5 microns $\pm$ 1.97 microns, 48.7 microns $\pm$.97 microns and 22.6 microns $\pm$.453 microns (the 12.5 micron aperture was not measured). The error bars were obtained from a 2% uncertainty in the magnification. Indeed the apertures were smaller than requested.

Figures 17 through 24 are field comparisons for the 250 micron aperture, even Fresnel patterns 6-20. Overall, these fits are extremely good. There seems to be a few variations in terms of the intensity of each pattern, however, this deviation can be attributed to the fashion in which the patterns were normalized. Aside from the intensity, the patterns compare very well. The overall shape is correct, and even the smaller details of the curves match quite well. The initial goal of the experiment was to document which obliquity factor from Chapter 2 was indeed correct. However, for the 250 micron aperture, all three factors produce the identical theoretical result.

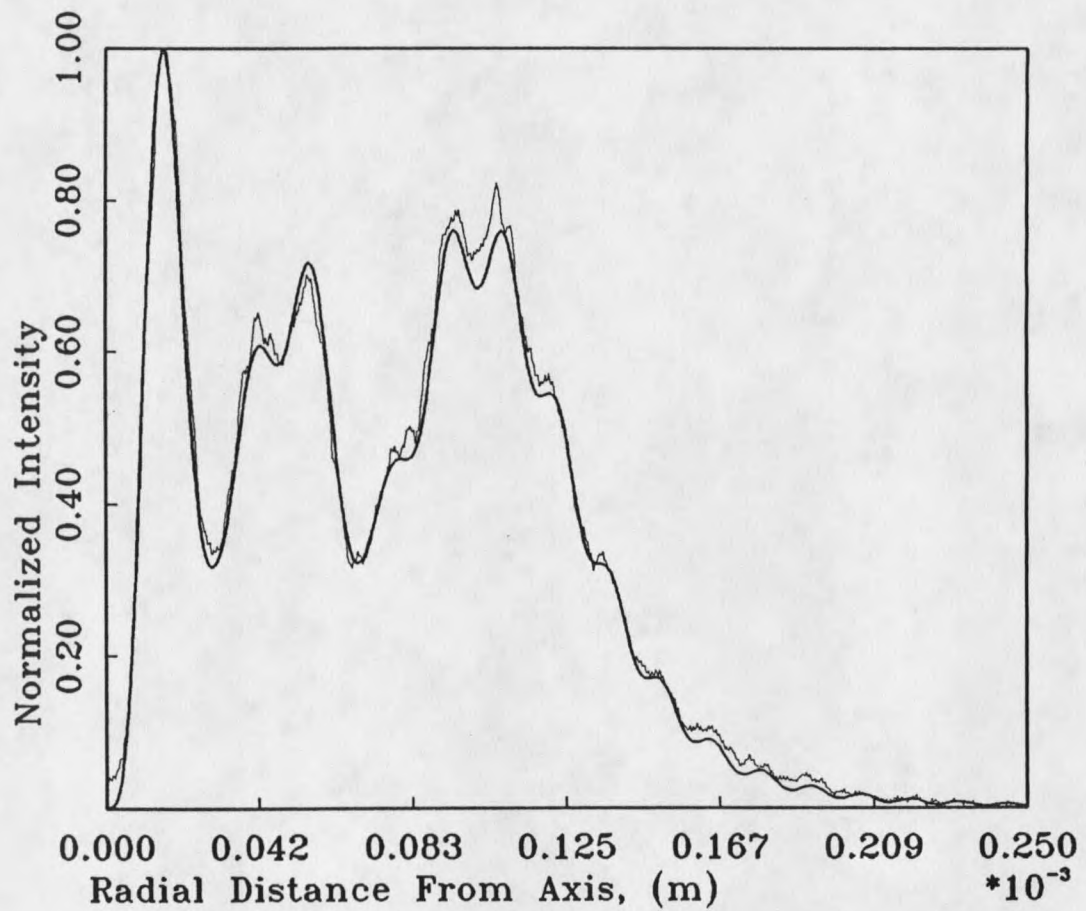


Figure 17 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 6 pattern. The pattern was observed at a distance of $5.128 \text{ mm} \pm .040 \text{ mm}$ away from the 250 micron aperture.

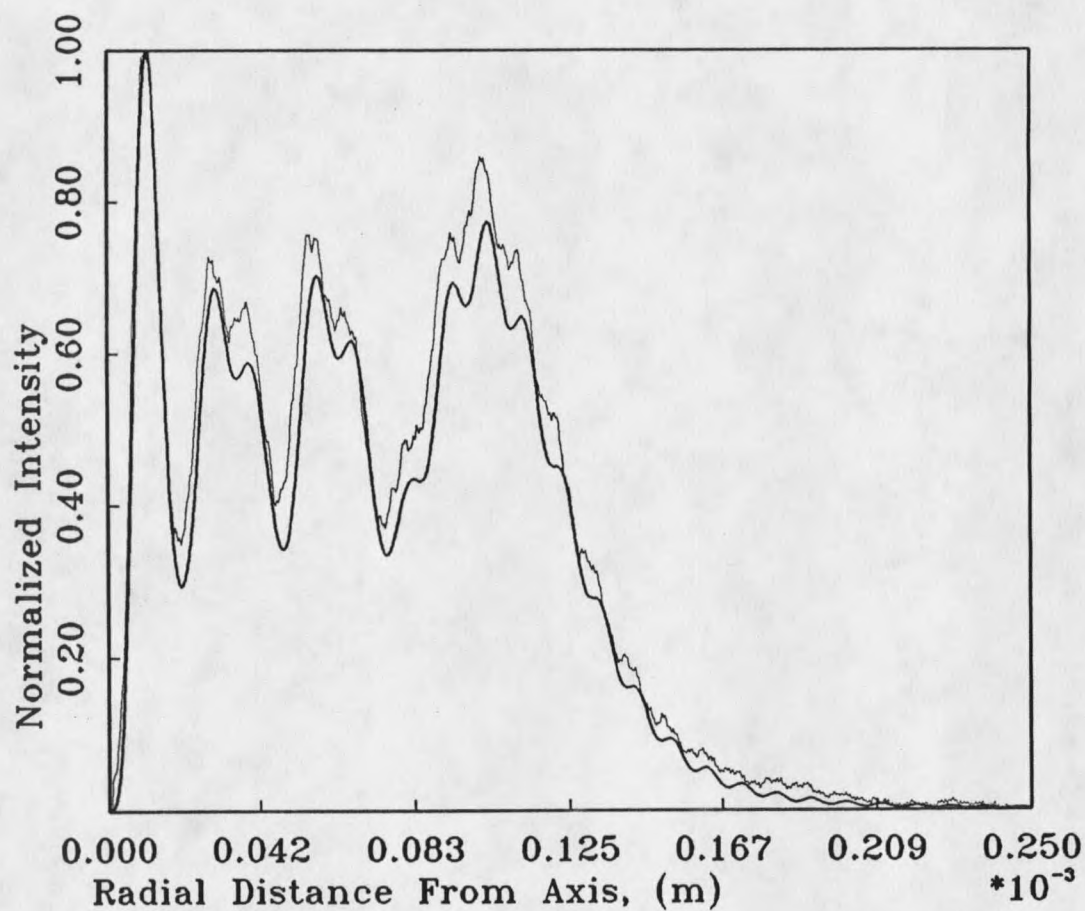


Figure 18 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 8 pattern. The pattern was observed at a distance of $3.600 \text{ mm} \pm .015 \text{ mm}$ away from the 250 micron aperture.

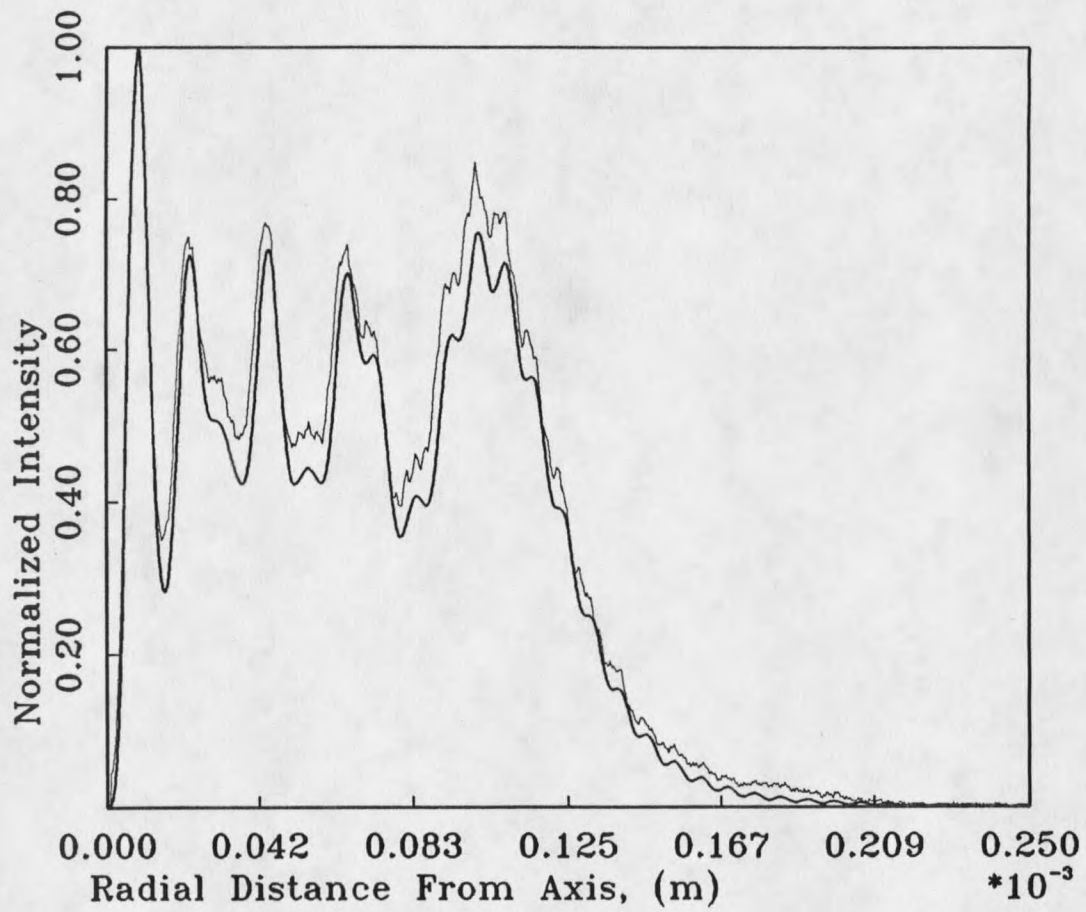


Figure 19 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 10 pattern. The pattern was observed at a distance of $2.788 \text{ mm} \pm .008 \text{ mm}$ away from the 250 micron aperture.

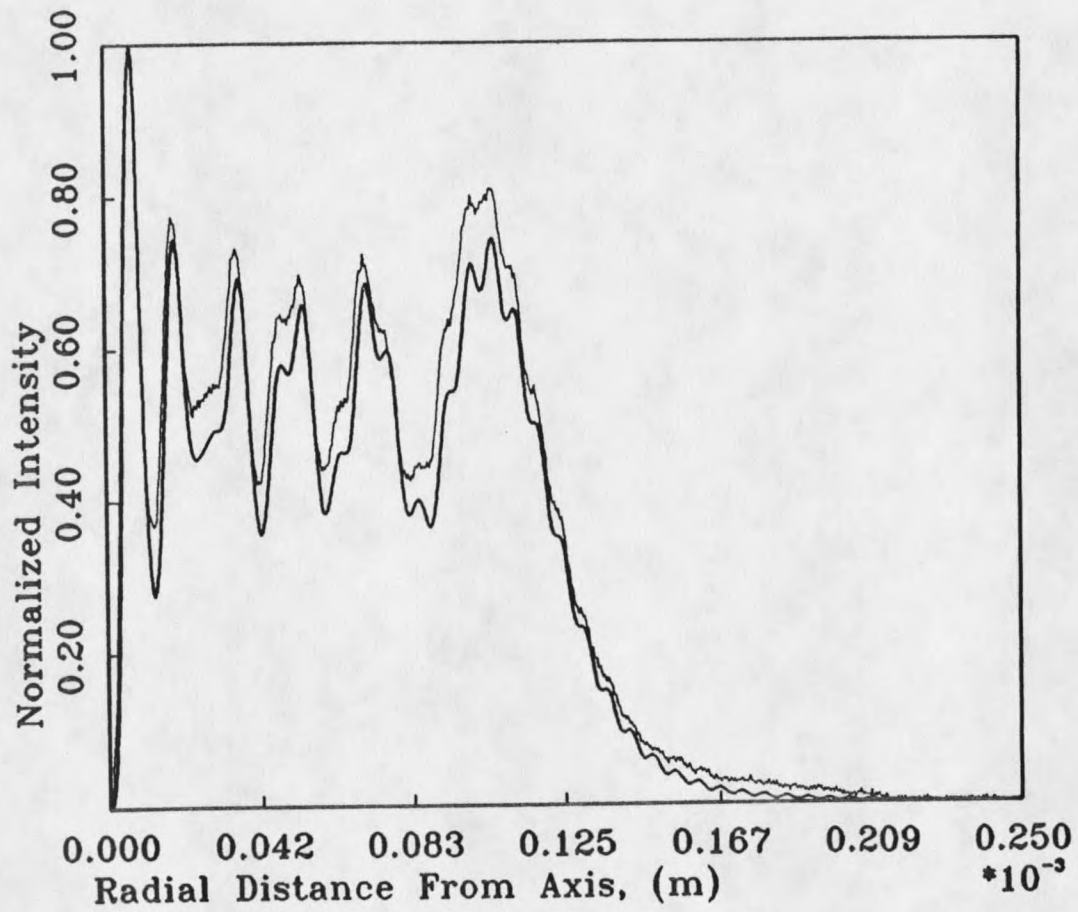


Figure 20 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 12 pattern. The pattern was observed at a distance of $2.274 \text{ mm} \pm .006 \text{ mm}$ away from the 250 micron aperture.

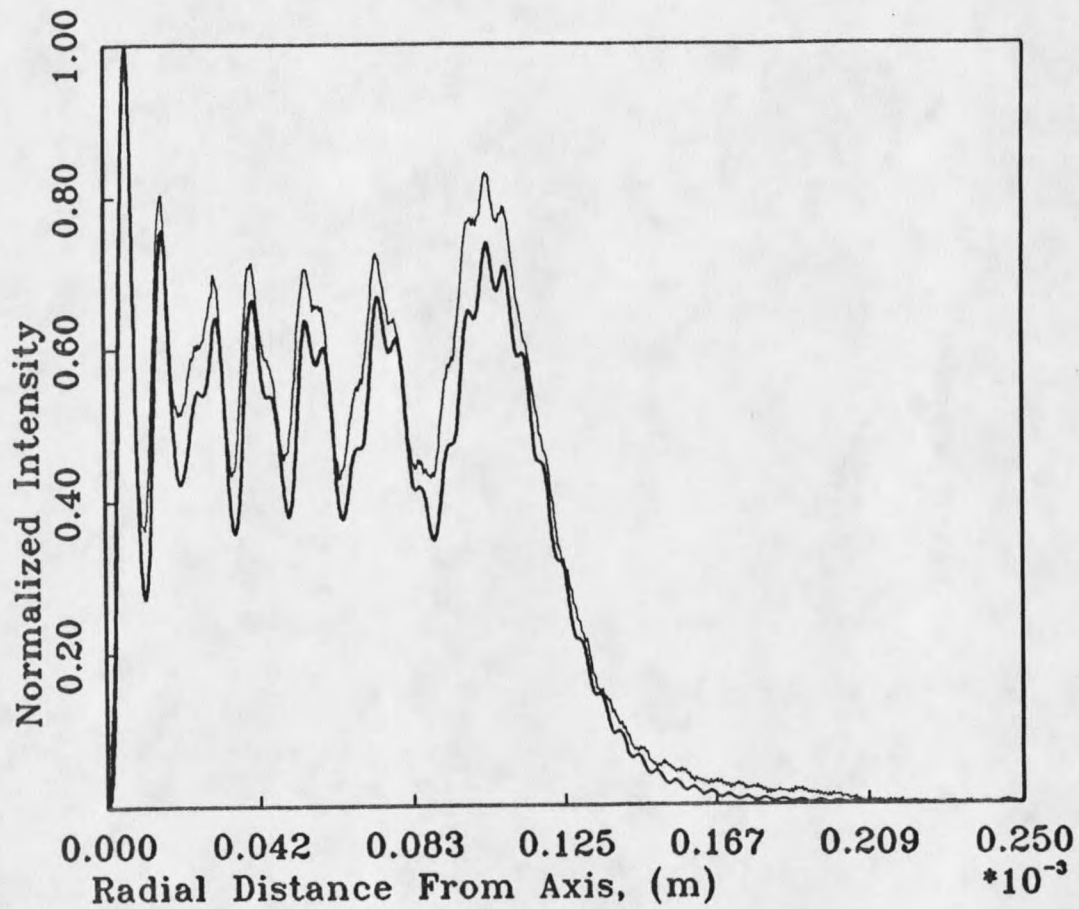


Figure 21 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 14 pattern. The pattern was observed at a distance of $1.920 \text{ mm} \pm .006 \text{ mm}$ away from the 250 micron aperture.

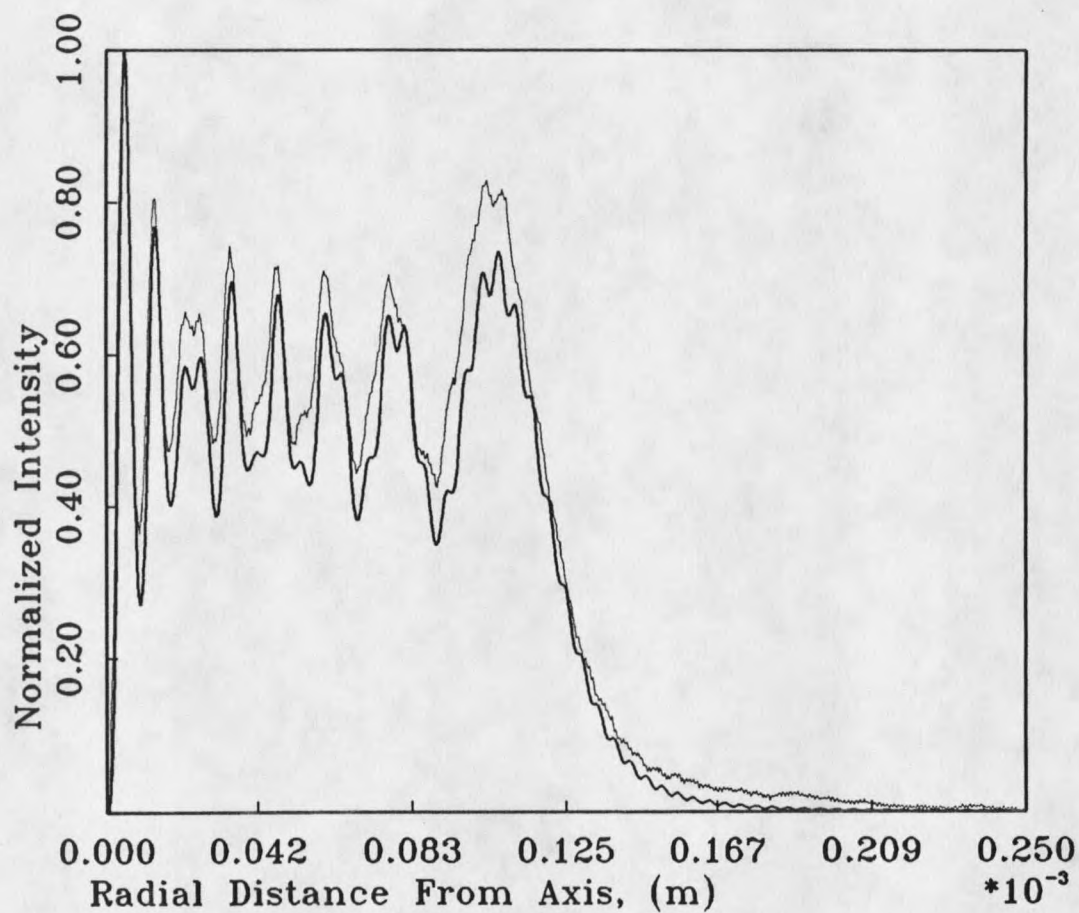


Figure 22 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 16 pattern. The pattern was observed at a distance of $1.664 \text{ mm} \pm .006 \text{ mm}$ away from the 250 micron aperture.

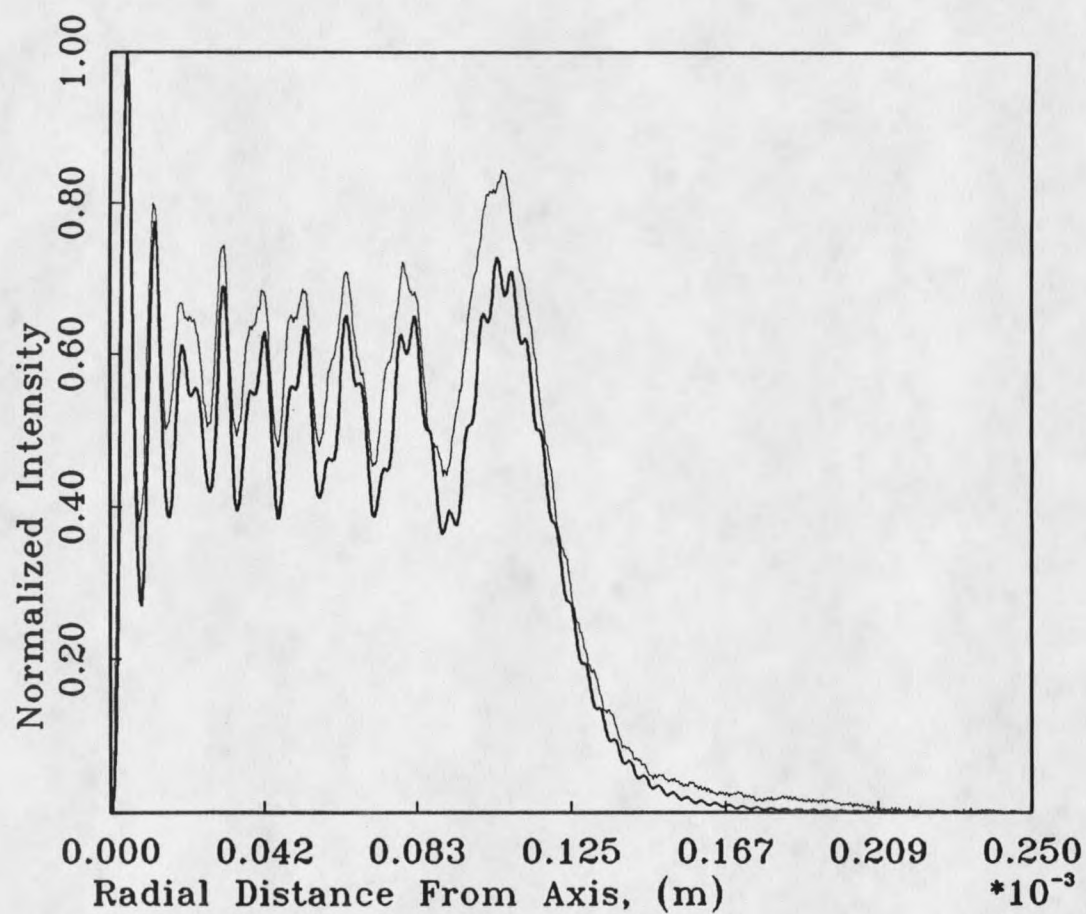


Figure 23 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 18 pattern. The pattern was observed at a distance of $1.462 \text{ mm} \pm .004 \text{ mm}$ away from the 250 micron aperture.

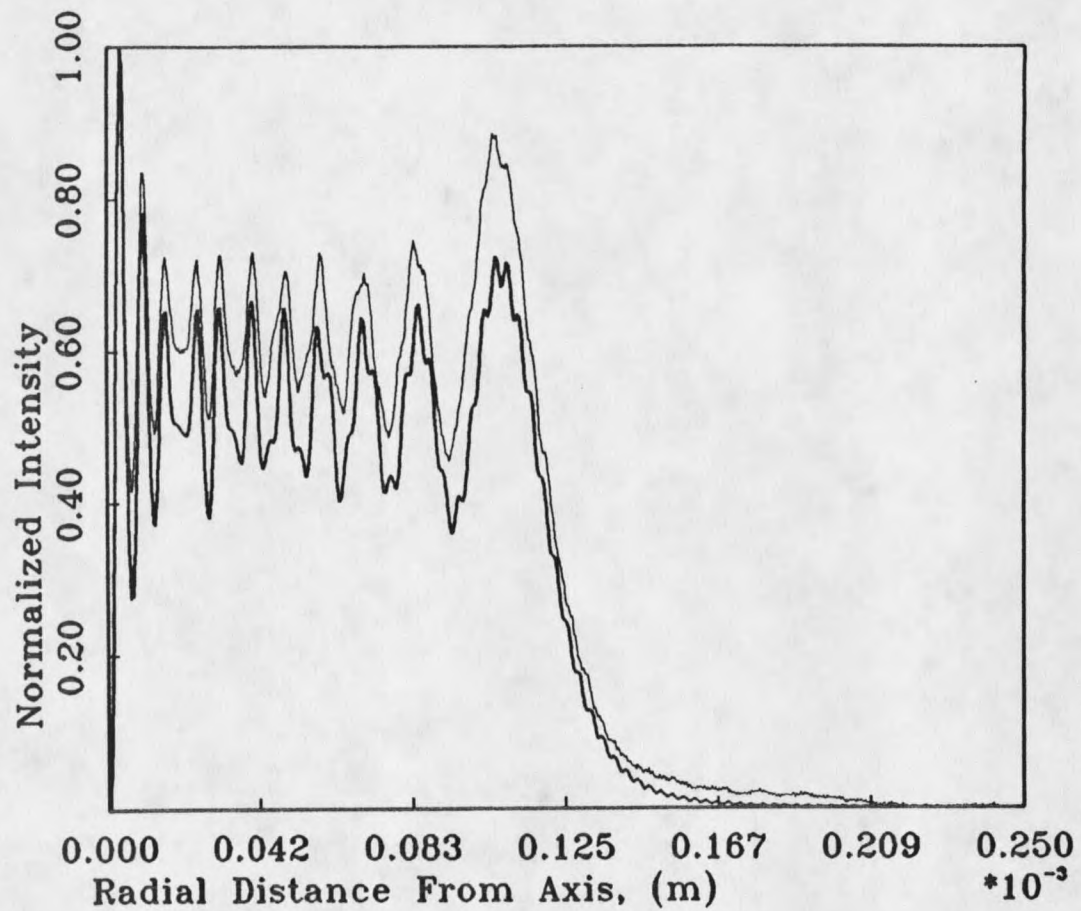


Figure 24 Theory vs Experiment plot for equation 2:1. This pattern is commonly referred to as a Fresnel number 20 pattern. The pattern was observed at a distance of $1.304 \text{ mm} \pm .004 \text{ mm}$ away from the 250 micron aperture.

The next comparison to be made was that of the locations of the Fresnel patterns as predicted by the Fresnel equation. The four larger apertures were utilized for this experiment. Figures 25 through 29 show how well the Fresnel equation predicts the pattern locations. In these figures, a slow breakdown appears to be occurring. In terms of notation, the solid theoretical line is the theory from the measured aperture size. The dotted lines arise from the error bar on the aperture size. The data points themselves also have error bars, but if no error bar appears, then the error bar is smaller than the physical size of the data point.

The results for the 250 micron aperture are extremely good. Even in the region very close to the aperture, Figure 26, the theory and experiment agree to within the error bar. In the plots concerning the other three apertures, a deviation is evident. This deviation becomes even more pronounced for the smaller apertures. Finally, Figure 30 enables the respective apertures to all appear on the same graph. Except for at the far left of the graph, the theory for all four apertures lies along the same line. The relative deviation between the four apertures is clear in this graph. While the 250 micron aperture shows no significant deviation from the theory, the other three apertures show a distinct difference between theory and experiment.

A good question at this point would be "Why does the theory break down?". A few references suggest the area in which the theory will begin to fail.^{2,6} In chapter 5, some of the basic ideas of the theory are presented and some questions are posed. Hopefully, these results will prompt additional work to be done on the problem of diffraction with small apertures.

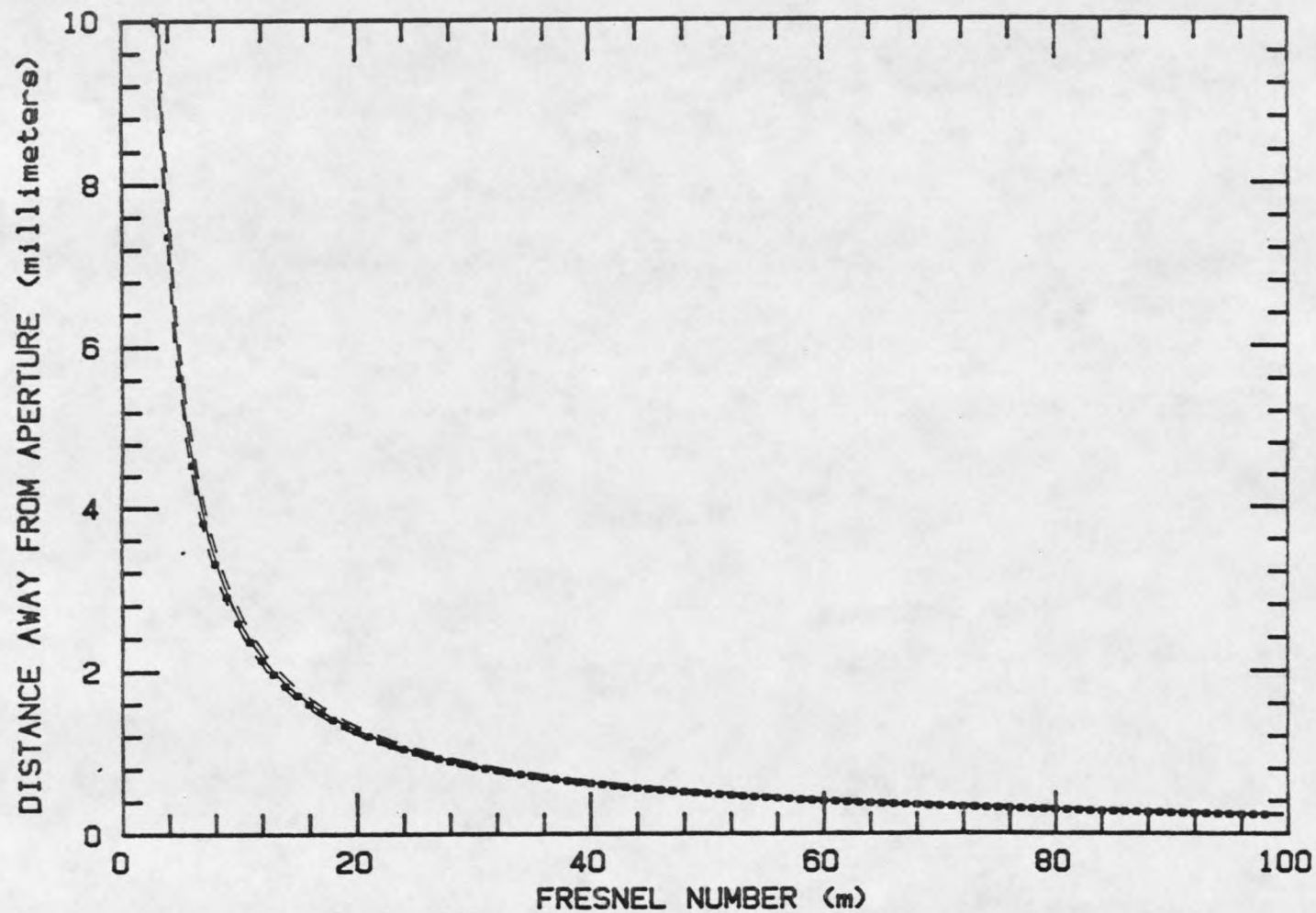


Figure 25 Theory vs Experiment plot for equation 2:4. Note that the theory and experiment agree as far as can be observed. This data is for the 250 micron diameter aperture.

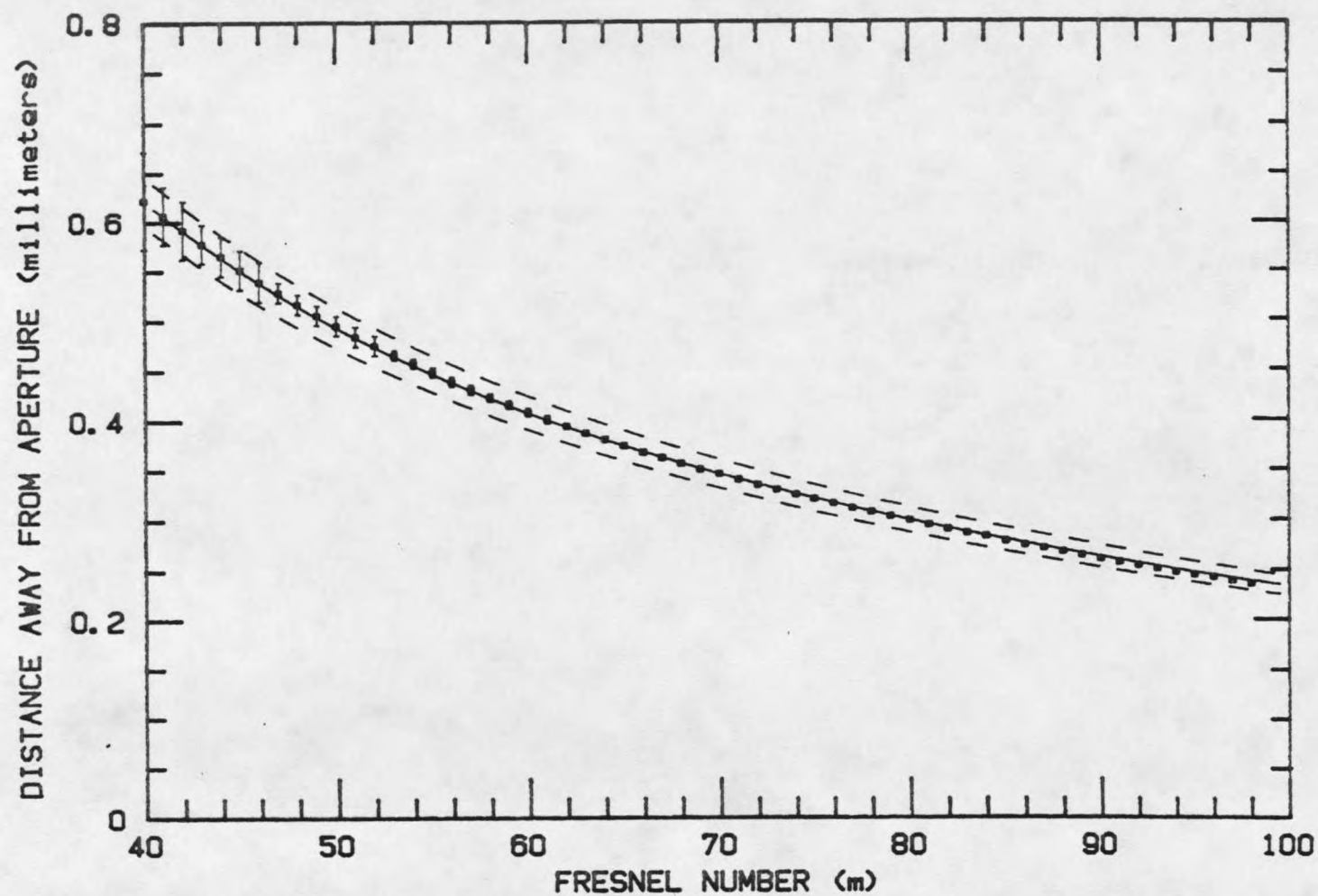


Figure 26 Theory vs Experiment plot for equation 2:4. This is an expanded view of the 250 micron aperture data. The data agrees with the theory within the error bar.

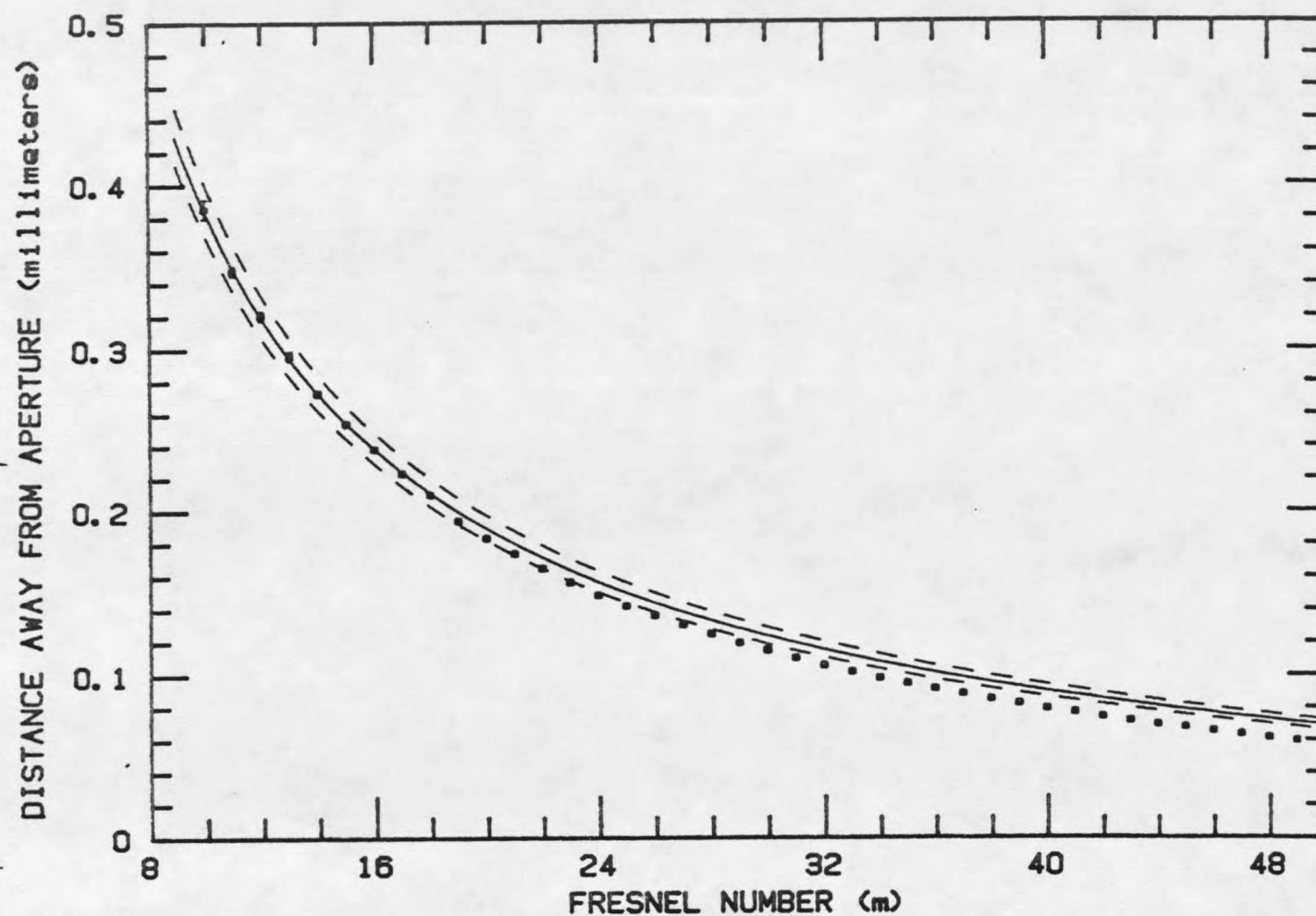


Figure 27 Theory vs Experiment plot for equation 2:4. This plot shows the data for the 100 micron aperture. Note that even with the error bars, the data deviates from the theory as the observation point approaches the aperture.

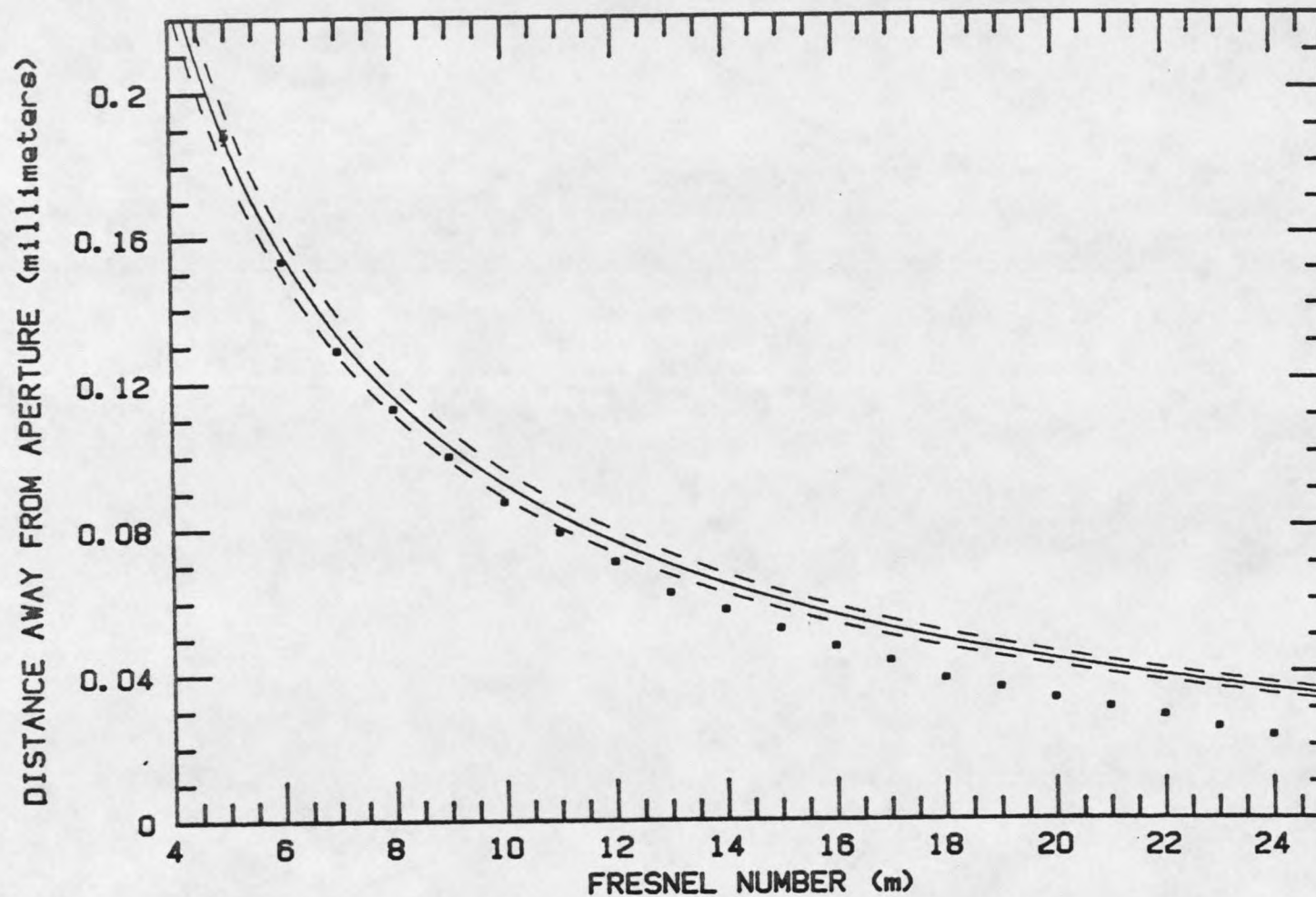


Figure 28 Theory vs Experiment plot for equation 2:4. This plot shows the data for the 50 micron aperture. Note that again, with the error bars, there is a distinct deviation from theory as the observation point approaches the aperture.

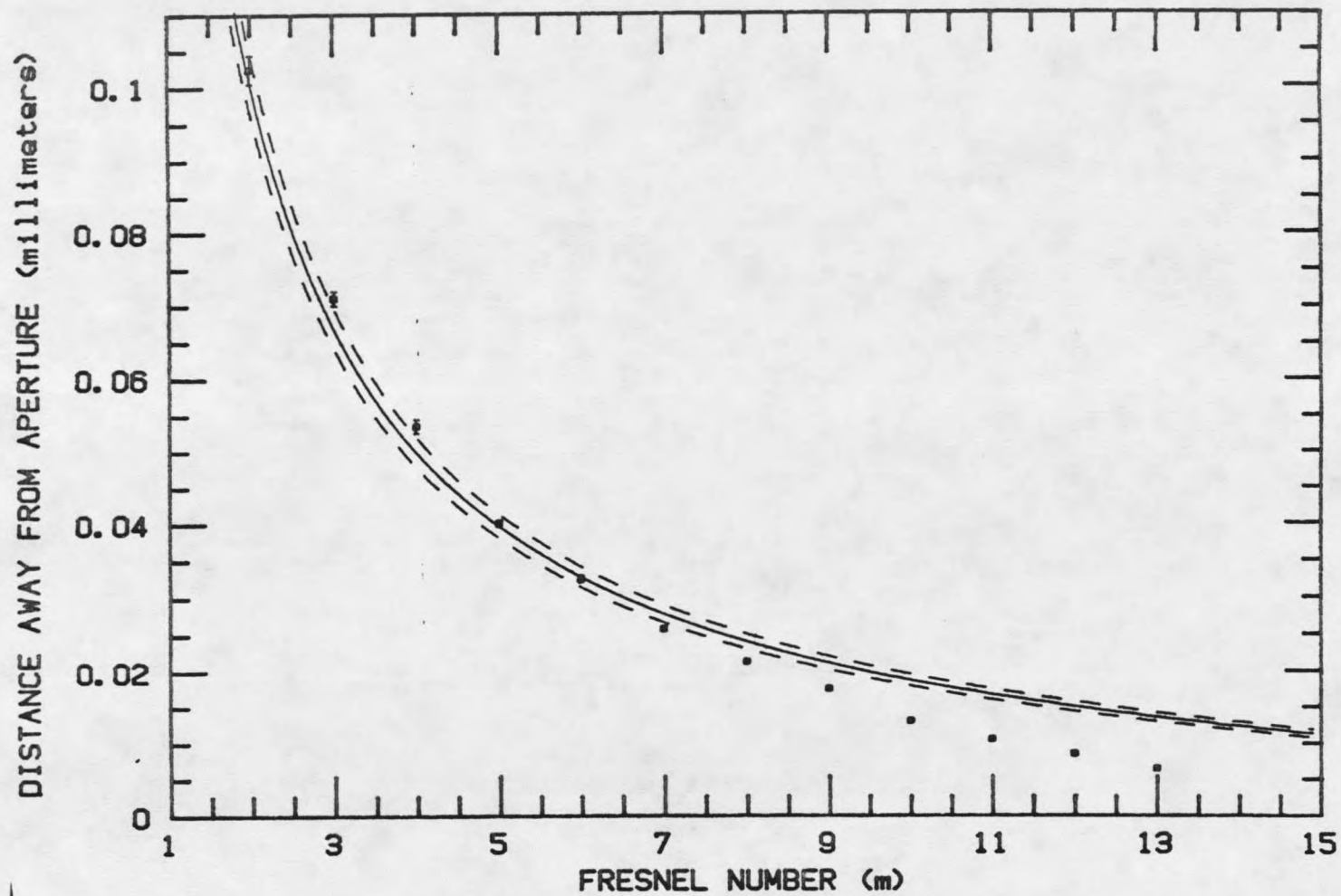


Figure 29 Theory vs Experiment plot for equation 2.4. This plot shows the data for the 25 micron aperture. Again, the deviation is present as the observation point approaches the aperture.

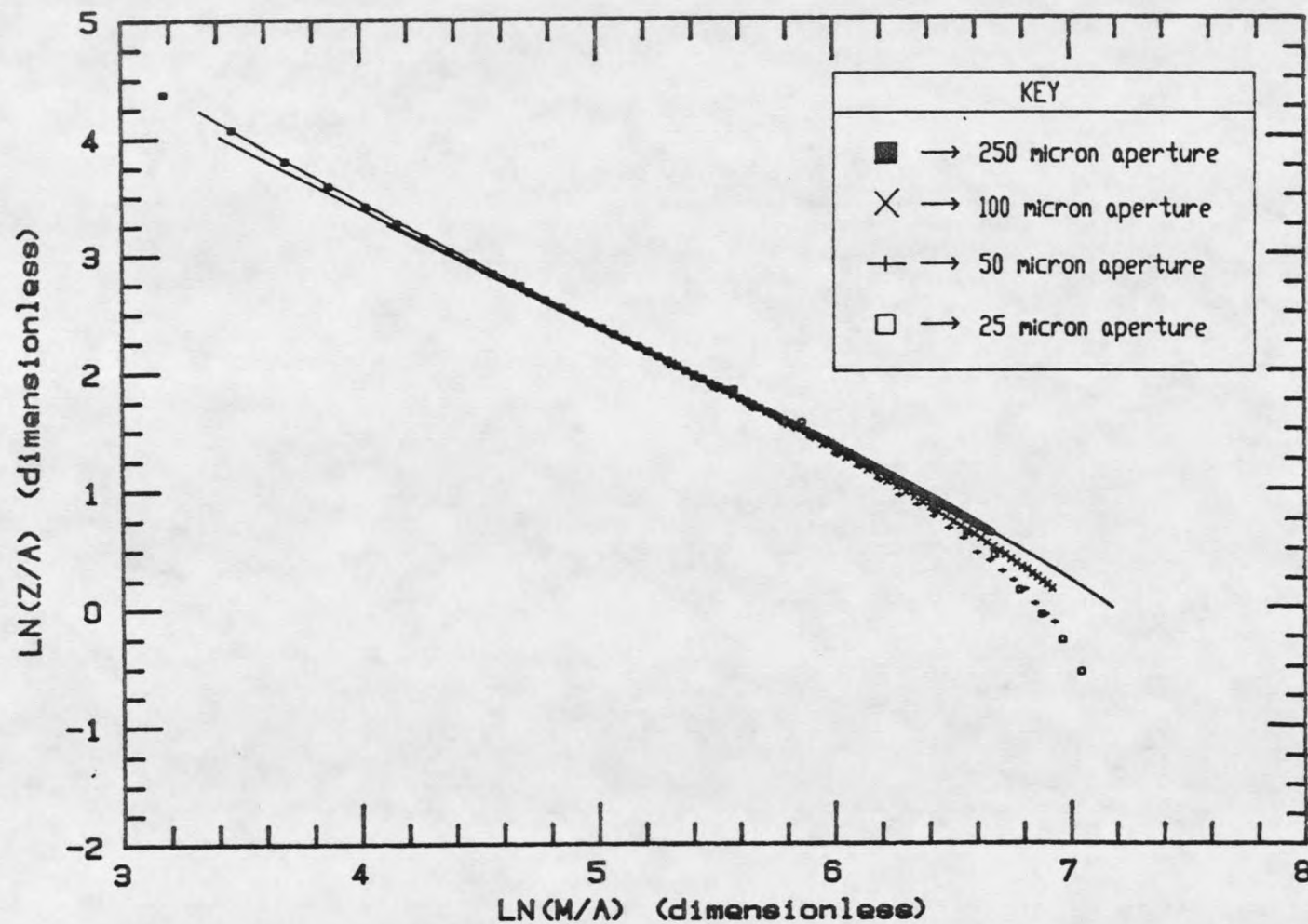


Figure 30 Theory vs Experiment plot for equation 2:4. By converting the axis to dimensionless parameters, all four apertures can appear on the same plot. At the right side of the plot, all four theoretical curves converge to the same line. The relative deviations from the theory can be observed here.

CHAPTER 5

DISCUSSION

In chapter 4 it was shown that for the 250 micron aperture, scalar diffraction theory predicts accurate results. However, as the aperture size decreases, the theory begins to fail as the observation point approaches the aperture. In an attempt to answer the question as to why the theory may fail, the basis of the scalar theory must be discussed. By understanding the assumptions which lead to the decoupling of the electric and magnetic fields, it will become evident that this experiment was probing an area not intended for scalar theory.

In order to be consistent with Maxwell's equations, the intensity at a point in the diffraction field is related to the electric and magnetic fields by

$$I = \frac{c}{4\pi} |\langle \vec{E} \times \vec{H} \rangle| \quad (5:1)$$

where E is the electric field, H is the magnetic field and c is the speed of light. It can be shown² that with the appropriate assumptions, this intensity can be expressed as the square of the modulus of a scalar field.

$$I = C|U|^2 \quad (5:2)$$

where C is a constant and U is a scalar wave function.

Note that the vectorial intensity is simply the Poynting vector. And, the poynting vector in our problem is propagating away from and perpendicular to the plane of the aperture. By equating the intensity with the Poynting vector, the theory acknowledges that all angles relative to the direction of propagation are to be small. One of the appropriate assumptions which accompanies the change from a vector theory to a scalar theory is indeed that all angles within the system are to be small.² The Born and Wolf optics text states that the angles involved are to be no more than 10 degrees in order for this transformation to be valid.

By observing the data of the experiment, it is clear that the angles involved are well over 10 degrees. For the 250 micron aperture, data was taken up to 1 diameter away from the aperture. This corresponds to angles nearer to 20 degrees. Meanwhile, the other three apertures were probed to within one half of an aperture diameter. The resulting angles there are above 40 degrees. As scalar theory depends explicitly on the criterion that angles are less than 10 degrees, the area of observation is clearly outside of the limits of scalar diffraction theory. This may be an indication that scalar theory should not work.

In conclusion, the experiment was initially set up to find out which obliquity factor from equation 2:2 was the correct factor. However, as the experiment progressed, requirements were set on the aperture size in order to determine the correct obliquity

factor. These requirements led to the investigation of an area where deviations from the Fresnel equation, (2:4), were observed. Although the largest aperture to be investigated in this experiment gave accurate results, the three other apertures gave distinct deviations. The success of the largest aperture helped to reinforce that the experimental process was sound. Investigations of the three smaller apertures led to regions of interest where it seems that scalar diffraction theory is invalid.

APPENDICES

APPENDIX A

DERIVATION OF SCALAR DIFFRACTION THEORY

APPENDIX A

DERIVATION OF SCALAR DIFFRACTION THEORY

In the derivation of scalar diffraction theory^{1-4,6}, both the Kirchhoff and Rayleigh-Sommerfeld formulations make certain bold assumptions about the nature of light. The most important assumption is that the light behaves as a scalar wave. Only the scalar amplitude and phase of one transverse component of the light is treated here. This approach clearly neglects the fact that the fields which compose the light are coupled through Maxwell's equations. It has been shown, however, that this bold assumption produces accurate results when: (1) the diffracting aperture is large compared to the wavelength and (2) the diffracted fields are observed far from the aperture.^{2,6} As was shown in Chapter 3, the experiment which was performed starts to penetrate regions of the diffracted field which do not obey these conditions. This appendix follows Goodman's derivation.⁶

For a starting point of scalar theory, one of the components of the optical wave is required to satisfy the scalar wave equation

$$\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (\text{A:1})$$

By representing the disturbance as a complex function

$$u(p) = U(P) \exp(-i(2\pi \nu t + \phi(P))) \quad (\text{A:2})$$

where $U(P)$ is the position dependent amplitude, and $\phi(P)$ is the position dependent phase, one arrives at the time independent Helmholtz equation

$$(\nabla^2 + k^2) \bar{U} = 0 \quad (\text{A:3})$$

where $k = 2\pi\nu/c = 2\pi/\lambda$ and $\bar{U}$ is the complex function of position

$$\bar{U} = U(P) \exp(-i\phi(P)) \quad (\text{A:4})$$

Calculation of the complex disturbance $\bar{U}(P)$ can be done with the assistance of Green's theorem. (From this point, the bar will be dropped for simplicity.) This theorem states that if $U(P)$ and $G(P)$ are any two complex-valued functions of position and S is a closed surface surrounding a volume V , and if U and G and their first and second partial derivatives are single-valued and continuous on and within S , then

$$\iiint_V (G \nabla^2 U - U \nabla^2 G) dv = \iint_S \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) ds \quad (\text{A:5})$$

The surface of integration is shown in Figure 31. By choosing an appropriate function to represent G , one can arrive at the Kirchhoff and Rayleigh-Sommerfeld formulations of diffraction theory.

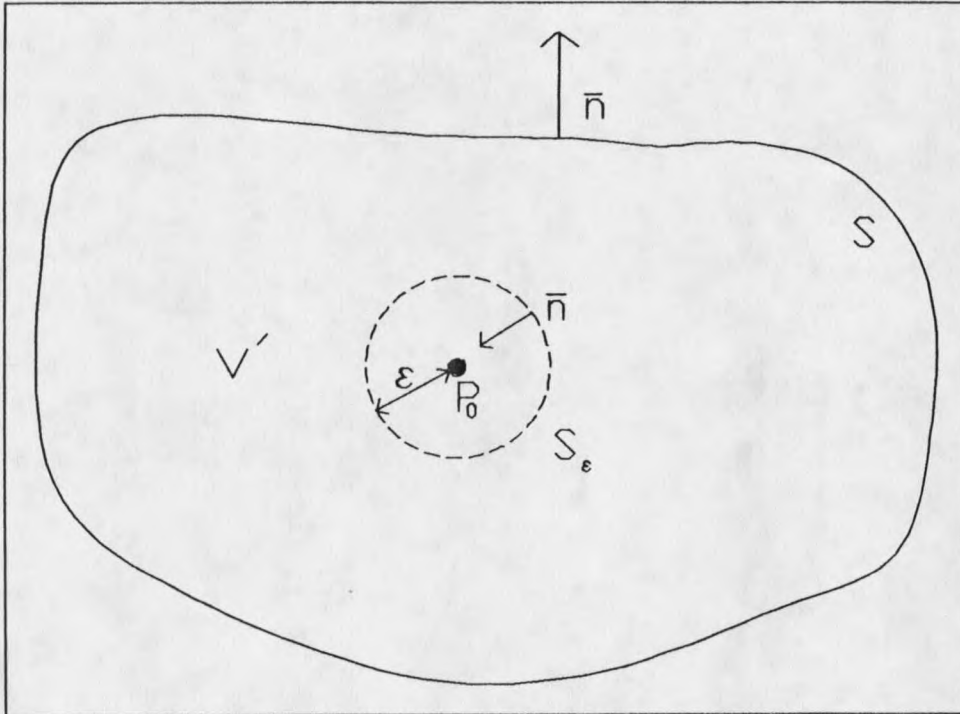


Figure 31 Geometry used in Green's theorem

The point of observation is denoted P_0 and S is an arbitrary closed surface surrounding P_0 . The problem now is to express the disturbance at P_0 in terms of its values on the surface S . Kirchhoff used Green's theorem to solve this problem. The choice of Green's function was that of a spherical wave expanding about P_0 . This is also referred to as the free-space Green's function. The value of G at an arbitrary point P_I will be given by

$$G(P_I) = \frac{\exp(ikr_{0I})}{r_{0I}} \quad (\text{A:6})$$

where the notation is that r_{0I} is the length of the vector pointing from P_0 to P_I .

Green's theorem requires that the Green's function be continuous at all points

within the defined surface. To avoid the discontinuity at P_0 , a small spherical shell, surface S_e , is inserted about the point P_0 . Green's theorem is now applied to the volume V' which lies between S and S_e and the combined surface

$$\dot{S}=S+S_e \quad (\text{A:7})$$

where the normal directions are as indicated in Figure 31.

The free-space Green's function, being an expanding spherical wave, satisfies the Helmholtz equation

$$(\nabla^2+k^2)G=0 \quad (\text{A:8})$$

By substituting the two Helmholtz equations, (A:3) and (A:8), the left hand side of Green's theorem, equation (A:5), is identically zero. The theorem reduces to

$$\int \int_{\dot{S}} (G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n}) ds = 0 \quad (\text{A:9})$$

By taking the integral over each surface separately one finds

$$-\int \int_{S_e} (G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n}) ds = \int \int_S (G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n}) ds \quad (\text{A:10})$$

Referring back to equation (A:6), the differential terms of the Green's function at a point P_1 on S' as

$$\frac{\partial G}{\partial n} = \cos(\vec{n}, \vec{r}_{01}) \left(ik - \frac{1}{r_{01}} \right) \frac{\exp(ikr_{01})}{r_{01}} \quad (\text{A:11})$$

where $\cos(\vec{n}, \vec{r}_{01})$ is the cosine of the angle between the outward normal of the surface and the vector joining P_0 to P_1 . For the case of P_1 on S_ϵ , $\cos(\vec{n}, \vec{r}_{01}) = -1$ and the Green's function and its derivative become

$$G(P_1) = \frac{\exp(ik\epsilon)}{\epsilon} \quad \text{and} \quad \frac{\partial G}{\partial n} = -\frac{\exp(ik\epsilon)}{\epsilon} \left(\frac{1}{\epsilon} - ik \right) \quad (\text{A:12})$$

Letting ϵ become very small, we write

$$\begin{aligned} \iint_{S_\epsilon} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) ds &= 4\pi\epsilon^2 \left[\frac{\partial U(P_0)}{\partial n} \frac{\exp(ik\epsilon)}{\epsilon} - U(P_0) \frac{\exp(ik\epsilon)}{\epsilon} \left(\frac{1}{\epsilon} - ik \right) \right] \\ &\underset{\epsilon \rightarrow 0}{=} -4\pi U(P_0) \end{aligned} \quad (\text{A:13})$$

Substitution of this result into equation A-10 yields

$$U(P_0) = \frac{1}{4\pi} \iint_S \left\{ \frac{\partial U}{\partial n} \left[\frac{\exp(ikr_{01})}{r_{01}} \right] - U \frac{\partial}{\partial n} \left[\frac{\exp(ikr_{01})}{r_{01}} \right] \right\} ds \quad (\text{A:14})$$

This result is formally called the integral theorem of Helmholtz and Kirchhoff.

Namely, it allows one to express the field at any point P_0 in terms of the values of the

field at the closed surface boundary.

By applying the integral theorem, one can address the problem of diffraction by an aperture in an infinite screen. The standard geometry of the problem is illustrated in Figure 32.

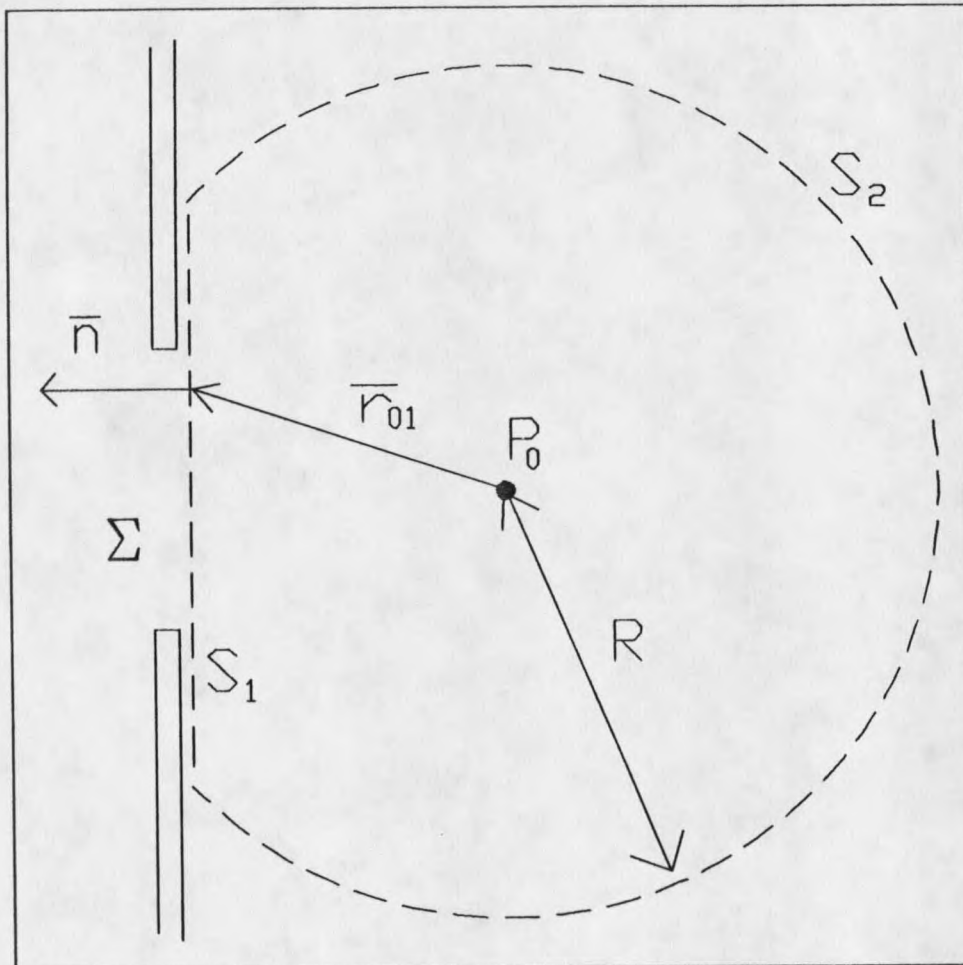


Figure 32 Kirchhoff formulation for diffraction by a plane screen

The wave is assumed to impinge onto the screen from the left and the diffraction field at the point P_0 is to be calculated.

In order to find the field at the point P_0 a convenient surface of integration is

chosen. The surface should reflect the geometry of the problem as much as possible without neglecting the important aspects of the problem. Kirchhoff chose the surface S to consist of 2 parts as shown in Figure 32. The two parts consist of a plane surface, S_1 , which lies directly behind the diffracting screen, and a large spherical surface away from the screen, S_2 . The sphere has a radius R centered at P_0 . Equation (A:14) will appear as

$$U(P_0) = \frac{1}{4\pi} \int \int_{S_1 + S_2} \left\{ \frac{\partial U}{\partial n} \left[\frac{\exp(ikr_{01})}{r_{01}} \right] - U \frac{\partial}{\partial n} \left[\frac{\exp(ikr_{01})}{r_{01}} \right] \right\} ds \quad (\text{A:14b})$$

As with most problems of this geometry, it would be favorable for the contribution from the spherical shell to be exactly zero.

The free space green's function about the point P_0 behaves as $1/R$.

Differentiating the Green's function

$$\frac{\partial G}{\partial n} = \left(ik - \frac{1}{R} \right) \frac{\exp(ikR)}{R} \approx ikG \quad (\text{for large } R) \quad (\text{A:15})$$

and inserting the result into the S_2 part of equation (A:14b), the contribution to the field at P_0 by the spherical shell, as R becomes large, is

$$\int \int_{S_2} \left[G \frac{\partial U}{\partial n} - U(ikG) \right] ds = \int_{\Omega} G \left(\frac{\partial U}{\partial n} - ikU \right) R^2 d\omega \quad (\text{A:16})$$

The quantity GR is finite (and equals $\exp(ikR)$) over the entire boundary. The integral over S_2 will vanish provided that U satisfies

$$\lim_{R \rightarrow \infty} R \left(\frac{\partial U}{\partial n} - ikU \right) = 0 \quad (\text{A:17})$$

This requirement is known as the **Sommerfeld Radiation condition**. Provided that the disturbance U vanishes as fast as a diverging spherical wave, equation (A:17) will hold and the contribution from S_2 will be identically zero.

Now, the integral theorem can be expressed solely in terms of the field in the plane of the aperture.

$$U(P_0) = \frac{1}{4\pi} \iint_{S_1} \left(\frac{\partial U}{\partial n} G - U \frac{\partial G}{\partial n} \right) ds \quad (\text{A:18})$$

Since the screen is opaque, it would seem reasonable that the field at P_0 would consist entirely due to the disturbance which propagates through the aperture. The area of the aperture is denoted as Σ . Kirchhoff adopted the following conditions:

1. Across the surface Σ , the field distribution U and its derivative $\partial U / \partial n$ are the same as they would be in the absence of the screen.
2. Over the portion of S_1 which lies in the geometrical shadow of the screen, the field distribution U and its derivative $\partial U / \partial n$ are identically zero.

These two conditions are known as the **Kirchhoff boundary conditions**. Rather than expressing the integral over the entire aperture plane, the Kirchhoff boundary conditions allow the integral to be expressed as an integral over the aperture itself

$$U(P_0) = \frac{1}{4\pi} \iint_{\Sigma} \left(\frac{\partial U}{\partial n} G - U \frac{\partial G}{\partial n} \right) ds \quad (\text{A:19})$$

The final step in this problem is to express the disturbance inside the aperture in terms of a source. Referring to Figure 33, a typical way to view diffraction is to have the disturbance impinging upon the aperture from the left and observing the diffraction field to the right of the screen.

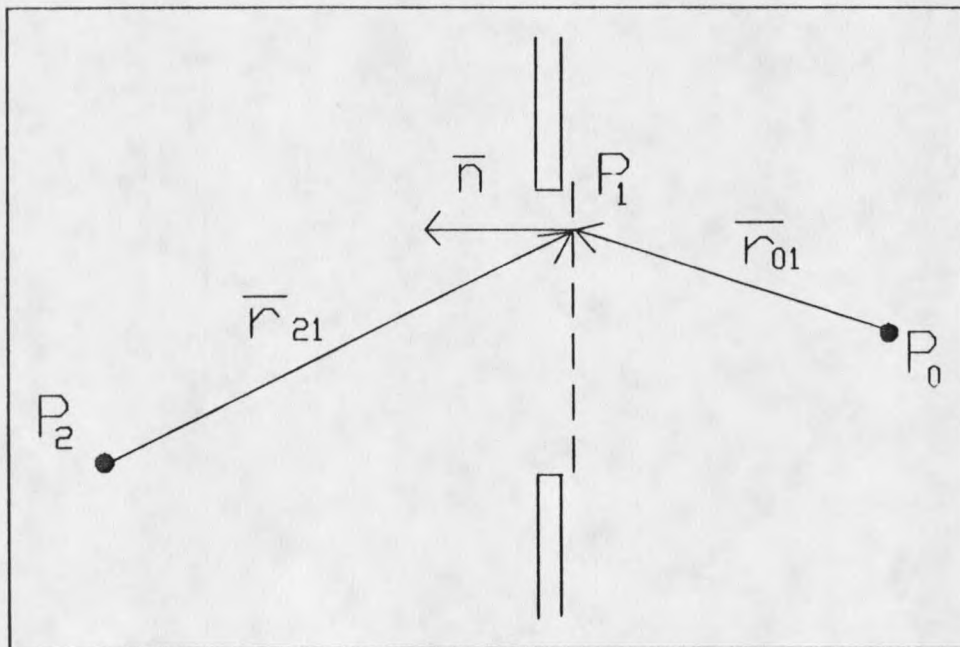


Figure 33 Point source illumination of a plane screen

Note that the normal to the aperture is pointed to the left, since the outward portion of the integration surface points from right to left according to Figure 32. The expression for $U(P_0)$ can be simplified further. With the assumption that the distance r_{01} is many wavelengths away from the aperture the derivative of the free space Green's function in the plane of the aperture becomes

$$\begin{aligned}\frac{\partial G(P_1)}{\partial n} &= \cos(\vec{n}, \vec{r}_{01}) \left(ik - \frac{1}{r_{01}} \right) \frac{\exp(ikr_{01})}{r_{01}} \\ &\approx ik \cos(\vec{n}, \vec{r}_{01}) \frac{\exp(ikr_{01})}{r_{01}}\end{aligned}\quad (\text{A:20})$$

Assuming that the aperture is illuminated by a point source at the point P_2 , such that

$$U(P_1) = \frac{A \exp(ikr_{21})}{r_{21}} \quad (\text{A:21})$$

and assuming that the same type of approximation as made in equation A:20 is valid, the resulting field at the point P_0 becomes

$$U(P_0) = \frac{A}{i\lambda} \iint_{\Sigma} \frac{\exp[ik(r_{21} + r_{01})]}{r_{21}r_{01}} \frac{[\cos(\vec{n}, \vec{r}_{01}) - \cos(\vec{n}, \vec{r}_{21})]}{2} ds \quad (\text{A:22})$$

Equation (A:22) is known as the Kirchhoff diffraction formula. The value of the integral depends explicitly on the phase relationships which exist from the aperture plane to the point of observation. This can also be viewed as a superposition of secondary sourcelets emitting waves from the plane of the aperture. Note that the cosine terms $1/2(\cos(\vec{n}, \vec{r}_{01}) - \cos(\vec{n}, \vec{r}_{21}))$ act to modify the amplitude of these secondary sourcelets. This term is known as the obliquity factor. In essence, this factor allows the wave to

propagate in a forward direction, rather than in all directions, as if the secondary sourcelet acted alone as a point source. Figure 34 shows the effective result of the obliquity factor. In this figure, the rings are intended to represent an amplitude distribution rather than waves. The thicker the ring, the more amplitude the wave possesses in that direction.

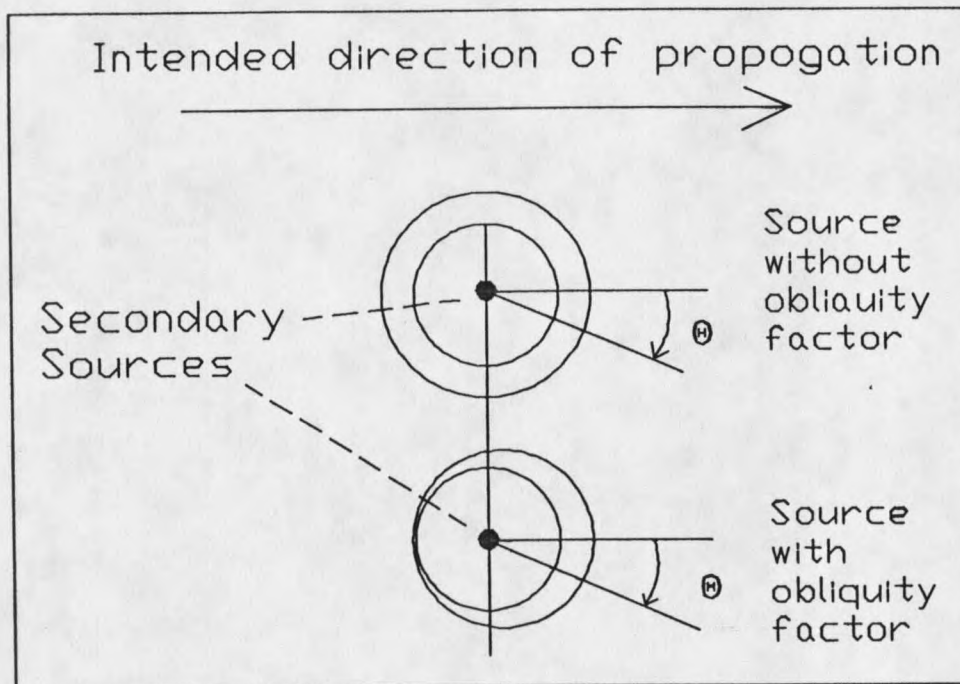


Figure 34 Without the obliquity factor, a secondary sourcelet would not have a preferred propagation direction.

Although several approximations were made to arrive at the Kirchhoff diffraction formula, it has been shown that in certain experimental situations, its ability to predict the diffraction field is quite accurate.^{7,8} Nonetheless, the derivation of equation (A:22) has undergone a certain amount of scrutiny for certain mathematical inconsistencies. Essentially, Kirchhoff over-defined the problem by defining both the value of the field and its derivative behind the aperture, U and $\partial U/\partial n$. In potential theory, there is a

theorem which states that if a two-dimensional potential function and its normal derivative vanish along any curve segment, the potential function must vanish over the entire plane. This theorem theoretically negates the secondary sourcelets as a possible cause for the diffraction field. Another mathematical inconsistency arises when the point of observation approaches the aperture plane. The Kirchhoff formulation fails to recover the assumed boundary conditions.⁶ In light of these inconsistencies, a more mathematically consistent formulation of diffraction theory was sought.

In theory, if Kirchhoff's choice of Green's function could be modified as to lead to equation (A:18) and either G or $\partial G/\partial n$ vanishes over the entire surface S_1 , the necessity of imposing the Kirchhoff boundary conditions on both U and $\partial U/\partial n$ would be removed. This would eliminate the mathematical inconsistencies in the Kirchhoff theory. Sommerfeld pointed out that a Green's function with the required property did exist.⁶ This Green's function is given by

$$G_-(P_1) = \frac{\exp(ikr_{01})}{r_{01}} - \frac{\exp(ik\tilde{r}_{01})}{\tilde{r}_{01}} \quad (\text{A:23})$$

This Green's function can be thought of as not only a point source located at P_0 , but also an identical point source at the position $\tilde{P}_0$. Figure 35 illustrates the two individual sources contributing to the Green's function. The corresponding derivative of this Green's function becomes

$$\frac{\partial G}{\partial n} = \cos(\bar{n}, \bar{r}_{01}) \left(ik - \frac{1}{r_{01}} \right) \frac{\exp(ikr_{01})}{r_{01}} - \cos(\bar{n}, \bar{\tilde{r}}_{01}) \left(ik - \frac{1}{\tilde{r}_{01}} \right) \frac{\exp(ik\tilde{r}_{01})}{\tilde{r}_{01}} \quad (\text{A:24})$$

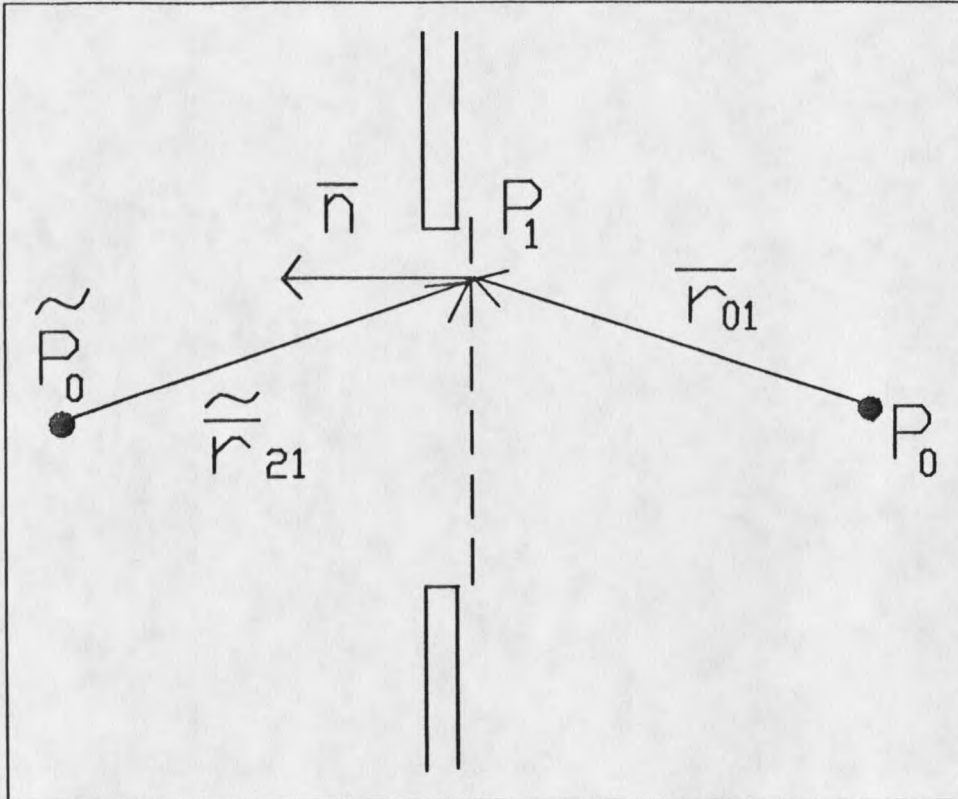


Figure 35 Sommerfeld's dual Green's function

With the points P_0 and $\tilde{P}_0$ being mirror points, the distance from both of these points to a common point in the aperture plane is the same distance. And, noticing that the points occur on opposite sides of the plane makes the two cosine terms in equation (A:24) negatives of each other. Therefore, one obtains for the function and its derivative

$$G_-(P_1) = 0$$

$$\frac{\partial G_-(P_1)}{\partial n} = 2 \cos(\bar{n}, \bar{r}_{01}) \left(ik - \frac{1}{r_{01}} \right) \frac{\exp(ikr_{01})}{r_{01}} \quad (\text{A:25})$$

Upon substitution of these conditions into equation (A:18) one obtains

$$U(P_0) = \frac{1}{i\lambda} \iint_{S_1} U(P_1) \frac{\exp(ikr_{01})}{r_{01}} \cos(\vec{n}, \vec{r}_{01}) ds \quad (\text{A:26})$$

Here, it has been assumed that $r_{01} \gg \lambda$ so the $1/r_{01}^2$ may be dropped. At this point, in order to restrict the integral to the area within the aperture, one must only state that the value of the field, $U(P_1)$ is zero outside of the aperture. At that point, the surface integral will be over the area Σ , rather than the entire aperture plane.

Again, a point source is assumed to diverge about the point P_2 as in Figure 33 so that the integral becomes

$$U(P_0) = \frac{A}{i\lambda} \iint_{\Sigma} \frac{\exp[ik(r_{21} + r_{01})]}{r_{21}r_{01}} \cos(\vec{n}, \vec{r}_{01}) ds \quad (\text{A:27})$$

Equation (2:27) is known as the Rayleigh-Sommerfeld diffraction formula. The integral is exactly what Kirchhoff arrived at with one exception. The obliquity factor is slightly altered, as it depends only on the angle at which the wavelet exits the aperture. The value of the obliquity factor still ranges from 0 to 1, as in the Kirchhoff formulation, but the form difference is still noticeable.

An alternative Green's function which still avoids the mathematical inconsistency is

$$G_+(P_1) = \frac{\exp(ikr_{01})}{r_{01}} + \frac{\exp(ik\tilde{r}_{01})}{\tilde{r}_{01}} \quad (\text{A:28})$$

Its value and derivative in the S_I plane are

$$G_+(P_1) = 2 \frac{\exp(ikr_{01})}{r_{01}} \quad \frac{\partial G_+(P_1)}{\partial n} = 0 \quad (\text{A:29})$$

where, again, the two Green's function sources are mirror images of each other, but oscillate in phase. Equation (A:18) is now written as:

$$U(P_0) = \frac{1}{i\lambda} \iint_{S_1} \frac{\exp(ikr_{01})}{r_{01}} \frac{\partial U(P_1)}{\partial n} ds \quad (\text{A:30})$$

The mathematical inconsistency found in Kirchhoff's theory is again removed, as one only needs to require that the value of the derivative of the incident disturbance is zero at the aperture. Similar to the last two formulations, a point source is assumed to diverge about a point P_2 as in Figure 33. The resulting integral, with the restriction put on the derivative of this disturbance becomes

$$U(P_0) = \frac{1}{i\lambda} \iint_{\Sigma} \frac{\exp[ik(r_{21} + r_{01})]}{r_{21}r_{01}} \cos(\bar{n}, \bar{r}_{21}) ds \quad (\text{A:31})$$

The assumption that $r_{21} \gg \lambda$ has been made to simplify $\partial U / \partial n$. Similarly, this is also the Rayleigh-Sommerfeld diffraction formula. Equations (A:27) and (A:31) are commonly referred to as the Rayleigh-Sommerfeld formulas I and II.

The derivation of scalar diffraction theory led to the formulation of three separate integrals based on the initial Green's functions assumption. For each case the same phase relationship was attained, but the weighting factors are different equations. Summarizing the three derivations, one obtains

$$U(P_0) = \frac{1}{i\lambda} \iint_{\Sigma} \frac{\exp[ik(r_{21} + r_{01})]}{r_{21}r_{01}} \phi ds \quad (\text{A:32})$$

$$\phi = 1/2[\cos(\bar{n}, \bar{r}_{01}) - \cos(\bar{n}, \bar{r}_{21})] \quad (\text{Kirchhoff})$$

$$\phi = \cos(\bar{n}, \bar{r}_{01}) \quad (\text{Rayleigh-Sommerfeld I})$$

$$\phi = \cos(\bar{n}, \bar{r}_{21}) \quad (\text{Rayleigh-Sommerfeld II})$$

Next, the derivation of the Fresnel equation will be considered. This equation predicts the location of some of the more interesting intensity patterns formed by equation (A:32). Namely, as the point of observation, P_0 , is moved along the optical axis of the diffraction system, the intensity of the diffraction field oscillates between a maximum and minimum value. The Fresnel equation predicts the locations of these maximums and minimums.

The equation is based solely on the relative pathlengths between the center of the

diffracting aperture and the edge of the aperture. Figure 36 shows a diverging spherical wave incident on an aperture of radius a . The wave protrudes a distance h through the aperture and has a radius ρ .

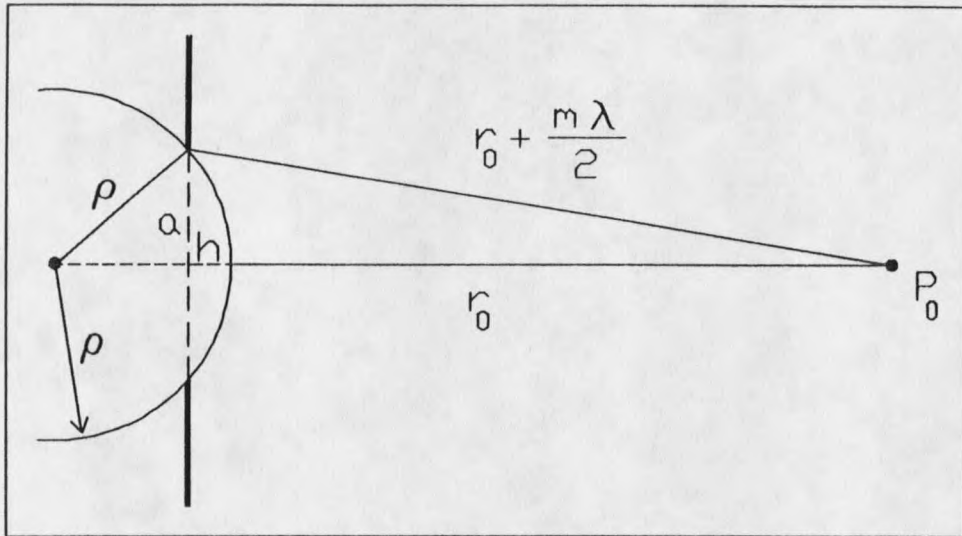


Figure 36 Geometry for derivation of the Fresnel equation

The most interesting effect can be observed when this pathlength difference is an integral number of half-wavelengths. Depending on the phase difference, there will be a maximum or a minimum at the observation point P_0 . A geometrical approach is taken at this point. A simple application of the Pythagorean theorem yields

$$\left(r_0 + \frac{m\lambda}{2}\right)^2 = a^2 + (r_0 + h)^2 \quad (\text{A:33})$$

Using a simple geometrical formula to express h in terms of a and ρ

$$h \approx \frac{a^2}{2\rho} \quad (\text{A:34})$$

and, after a little algebra, the position of the Fresnel pattern m , can be expressed at a distance r_0 away from the front of the wave as

$$r_0(m) = \frac{\lambda^2 m^2 / 4 - a^2}{a^2 / p - m\lambda} \quad (\text{A:35})$$

This result is the Fresnel equation.

APPENDIX B

SPHERICAL ABERRATIONS - COMPUTER SIMULATION STUDY

APPENDIX B

SPHERICAL ABERRATIONS - COMPUTER SIMULATION STUDY

One of the first studies related to the performance of the experiment was that of aberration introduced by spherical lenses. The general theme of the study was to quantify the amount of spherical aberration present in a lens when it is used to focus a gaussian beam. The study hoped to answer the question as to how much of a lens would be termed "diffraction limited" as defined in chapter 3. The study was accomplished with the aid of a computer ray tracing program entitled Beam 3, which is produced by Stellar Software, Berkeley, California. The program allows one to specify lens surfaces, indices, rays and other qualities of virtually any optical system.

Ideally, when concerned with geometrical optics, a perfect lens will focus a parallel beam of light to a geometrical point. However, real lenses do not behave this way. The main reason for the presence of the aberration is simply due to the construction of the lens. The ideal lens has non-spherical surfaces, however, most lenses are ground with spherical surfaces because this process is much easier. The central part of any lens can be close to ideal. However, as the lens increases in size, the spherical surface deviates more and more from the ideal surface. Thus the non-axial rays of light

are focused in a different area than the rays which are extremely close to the axis of the lens. Figure 37 shows a typical image of spherical aberration which was obtained directly from Beam 3. Notice that the rays near the edge of the lens are focused in a region before the paraxial focus.

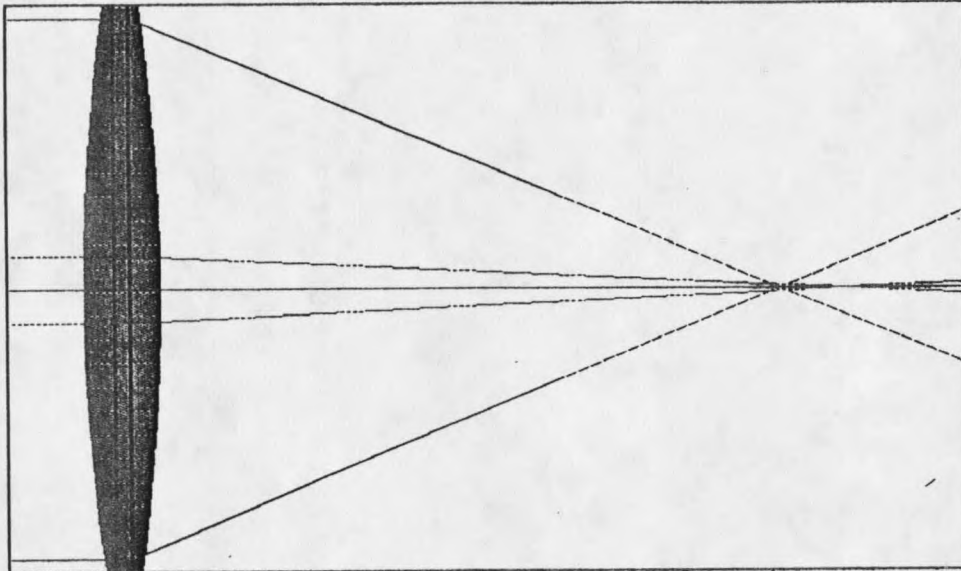


Figure 37 Spherical Aberration. The rays being focused from the edge of the lens do not focus in the same location as the paraxial rays.

To conduct the study, six different, commonly used lenses were programmed into Beam 3. The lenses had focal lengths of 5 cm, 10 cm, 20 cm, 50 cm, 100 cm and 200 cm. This was accomplished by specifying the radius of curvature of each surface and the index of refraction between the surfaces. The index used was 1.52, which corresponds to BK-7 lens glass. Then, parallel beams of increasing diameter were programmed to propagate through the lens. One half the diameter of the beam was considered to be equal to the waist size of the beam. The resulting beam size was measured in the plane of the true focus, as specified by the lens-maker's equation for a

thin lens:

$$\frac{1}{f} = (n_l - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (\text{B:1})$$

where n_l is the lens material's index of refraction and R_1 and R_2 are the radii of curvature of the surfaces of the lens. The size of the beam at "focus", as a function of input waist into the lens, was then compared to the expected waist of a gaussian beam of the same input waist. The gaussian beam waist is given by

$$\frac{\lambda f}{\pi \omega_0} = \omega_f \quad (\text{B:2})$$

Realize that the comparison being made is between a concept which is based in geometrical optics and a concept which is based in gaussian optics. However, when the geometrical spot size deviates from zero, a measure of spherical aberration can be made. Figure 38 shows the comparison graph.

Note that the two curves are in opposite directions, as one is governed exclusively by aberration and the other is governed by gaussian optics. Figure 38 simply illustrates that spherical aberration begins to contribute the spot size of a focused gaussian beam. It was assumed that once the aberration spot size became greater than 10 percent of the predicted gaussian spot size, the focused spot would no longer be diffraction limited. It was then assumed that the actual diameter

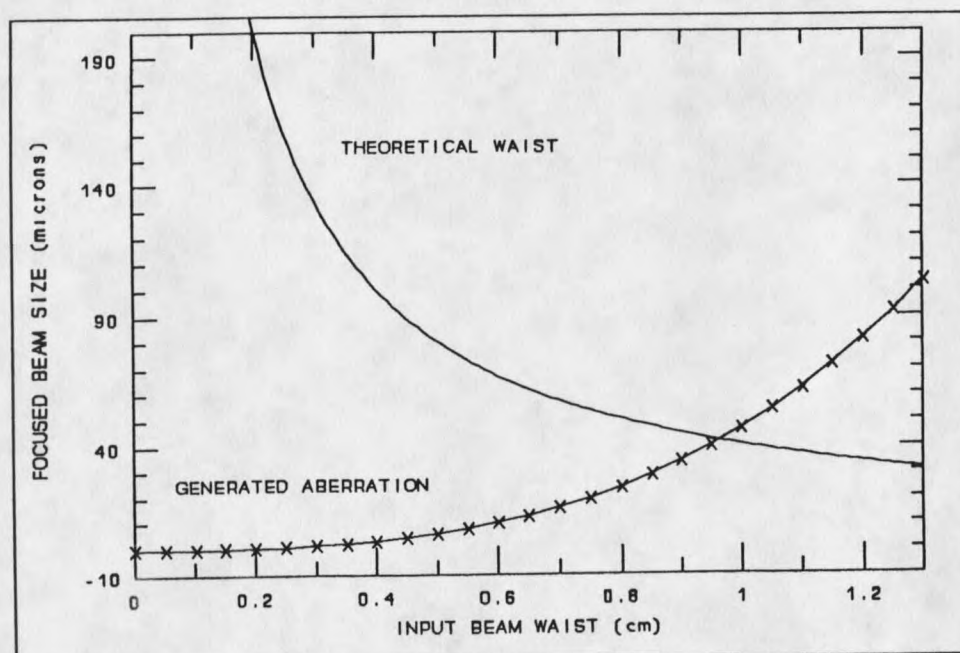


Figure 38 Comparison between geometrical spot size and theoretical gaussian waist; a measure of spherical aberration.

of the lens which could be used would be three times the waist which produced a 10 percent aberration. This assumption is reasonable, as a gaussian beam has a spot diameter about three times its waist. The following data summarizes the findings. These values give the researcher an good idea of how much of a lens to be used can be considered to be diffraction limited.

<i>Lens Focal Length</i>	ω_0	$3\omega_0$
5 cm	0.123 cm	0.369 cm
10 cm	0.2 cm	0.6 cm
20 cm	0.31 cm	0.93 cm
50 cm	0.66 cm	1.98 cm
100 cm	1.18 cm	3.54 cm
200 cm	> 2.2 cm	> 6.6 cm

APPENDIX C

OBSERVING SPHERICAL ABERRATION USING A TWYMAN-GREEN INTERFEROMETER

APPENDIX C

OBSERVING SPHERICAL ABERRATION USING A TWYMAN-GREEN INTERFEROMETER

After doing the computer simulation to quantify spherical aberration, it was hoped that an experiment could be done to see if the calculations were indeed successful at predicting spherical aberration. The usual procedure to test the amount of aberration in a lens is to use a standard wedge collimation tester. By collimating a beam with the lens to be tested, aberrations become apparent when the collimation of the beam is tested, as the anticipated fringes are not straight.¹² However, as the size of the lens becomes smaller and smaller, the collimation tester is not as effective, as the shear introduced by the wedge becomes too large to produce as sufficient interference pattern. At this point, such a small beam can be collimated by sight, simply by observing the beam diameter over a large distance. However, there remains no qualitative information about how much aberration is present on the beam. Another type of interferometer is needed.

The Twyman-Green interferometer can be used to test optical components. It was invented and patented by Twyman and Green in 1916 for the testing of prisms.¹² With a small alteration, it can be used for testing lenses with great ease. Figure 39

shows the appropriate configuration for testing optical components. The interferometer consists of a Michelson interferometer modified by the insertion of the lens to be tested into one of the legs.

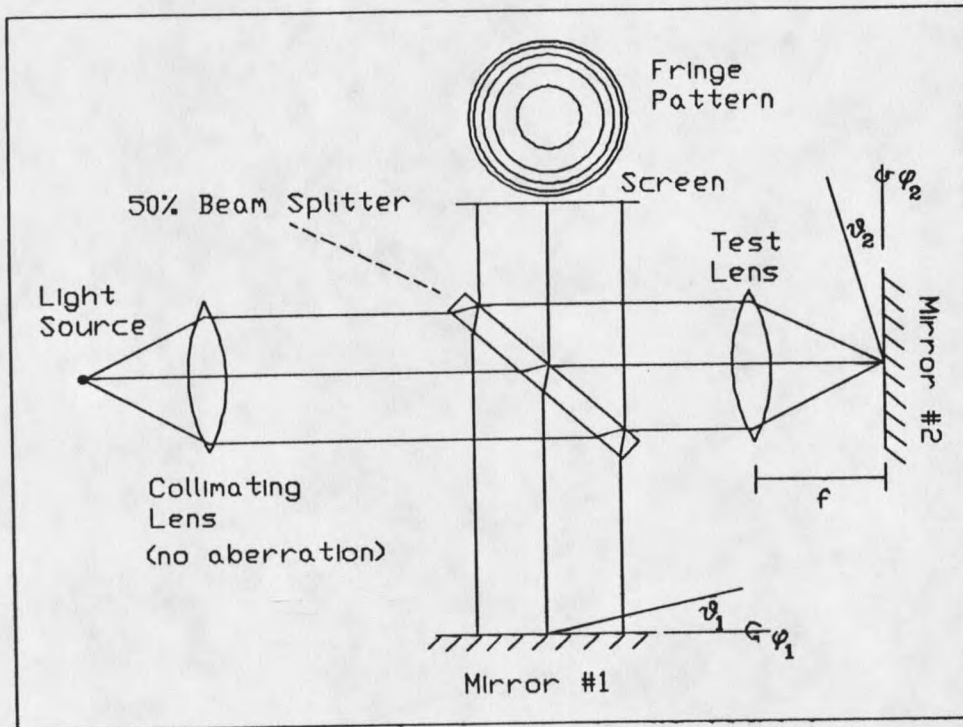


Figure 39 Twyman-Green interferometer (modified Michelson interferometer) configured for testing lenses

Each pass of the beam through the lens introduces an amount of aberration for the spherical lens. The resulting interference pattern consists of concentric rings corresponding to the cylindrical symmetry of the aberration in the lens.

The process of alignment of the interferometer is fairly simple. First, the interferometer is aligned without the lens so that a "null" pattern is present. This null pattern represents the fact that the two beams exiting the interferometer onto the screen are now parallel. The pattern is simply that of one single bright or dark fringe, as the

phase variation across the entire fringe pattern is the same. The alignment consists of varying the angles ϑ and ϕ on both mirrors. One reference point to observe throughout the alignment process is the portion of the beam reflected out of the system back onto the light source. The light from the interferometer will trace its way back out of the entrance leg, and a small spot will appear near the light source. Matching this small spot to the light source will have the beams lined up with such accuracy that only fine tuning will be needed from this point.

Once the null pattern is achieved, the lens to be tested is inserted into one of the legs of the interferometer. Frequently, this process tends to take the system out of alignment. The system is realigned using mirror #1. Since the lens is focusing the beam onto mirror #2, altering mirror #2 would be of little use, as the lens tends to re-collimate the beam regardless of mirror #2's rotation. The null pattern is once again achieved.

However, due to the aberration present in the lens, the null pattern will now consist of the previous bright or dark spot, surrounded by a ring pattern, as illustrated in Figure 39. By measuring the size of the single null fringe in comparison to the size of the lens being used, one can arrive at how much of the lens is "diffraction limited". Using this testing process, it was possible to compare the images from the interferometer with the theoretical data generated in Appendix B. The data shows a remarkable similarity. For example, for a 5 cm focal length lens, the interferometer yields a diffraction limited spot size of .4 cm. Figure 40 shows the interferogram from the 5 cm lens. The clear spot at the center of the lens is the amount of the lens to be considered diffraction limited. Above the image is a millimeter scale. The diameter of the spot is

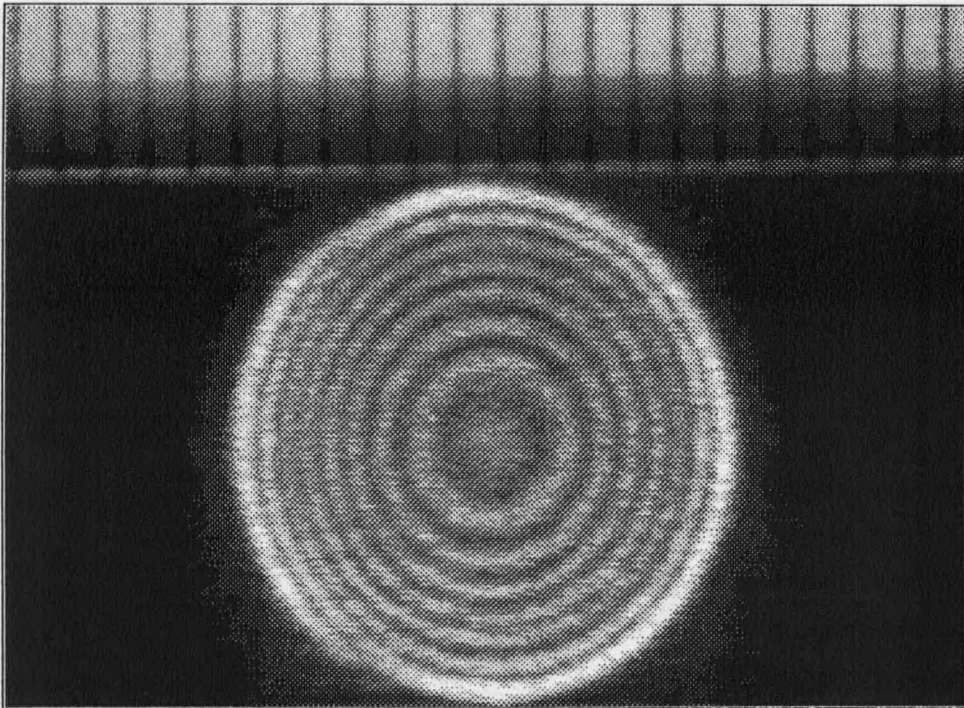


Figure 40 Image of Spherical Aberration as observed by the Twyman-Green interferometer.

about 1 cm, while the diameter of the first circular ring is about .3 - .4 cm. This is very close to the previous computer simulation. As can be seen, the Twyman-Green interferometer is simple, but effective in providing a way to measure the diffraction limit of a lens.

APPENDIX D

THE CONVOLUTION INTEGRAL

APPENDIX D

THE CONVOLUTION INTEGRAL

The formation of an image by an optical system is usually explained by geometrical optics. Through a lens with a given focal length, the image is a point to point mapping of the object, with the proper magnification and position factors taken into account. This process fails to take into account the wave nature of light, as diffraction is left unaccounted for. While geometrical optics sufficiently describes position and magnification, the quality of the formed image is left to be explained by Fourier Optics. In this branch of optics, the mapping which occurs is not the point to point mapping which occurs in its geometrical counterpart. Now, every point in the object plane will be mapped into a corresponding response or "spread" function in the image plane, the sum of the individual spread functions across the image plane forms the image. Several texts include the concepts of image formation and its correlation to Fourier Optics.^{1,2,6}

It is known that the response function takes the form of the Airy pattern.¹ The process which maps each point of the object into the image is now called the **convolution**. Mathematically, the convolution appears as the integral

$$U_i(x_i, y_i) = \int \int_{-\infty}^{\infty} h(x_i, y_i; x_o, y_o) U_o(x_o, y_o) dx_o dy_o \quad (D:1)$$

$U_i(x_i, y_i)$ is the field distribution in the image plane (i), $U_o(x_o, y_o)$ is the field distribution in the object plane (o). And, $h(x_o, y_o; x_i, y_i)$ is the response function, which describes how each point in the object plane is mapped into the image plane. The details of the imaging system are contained in the response function. These details include the magnification and the size of the spread function. A simple one dimensional convolution is demonstrated in Figure 41. In this figure, each point is simply spread out by the response function.

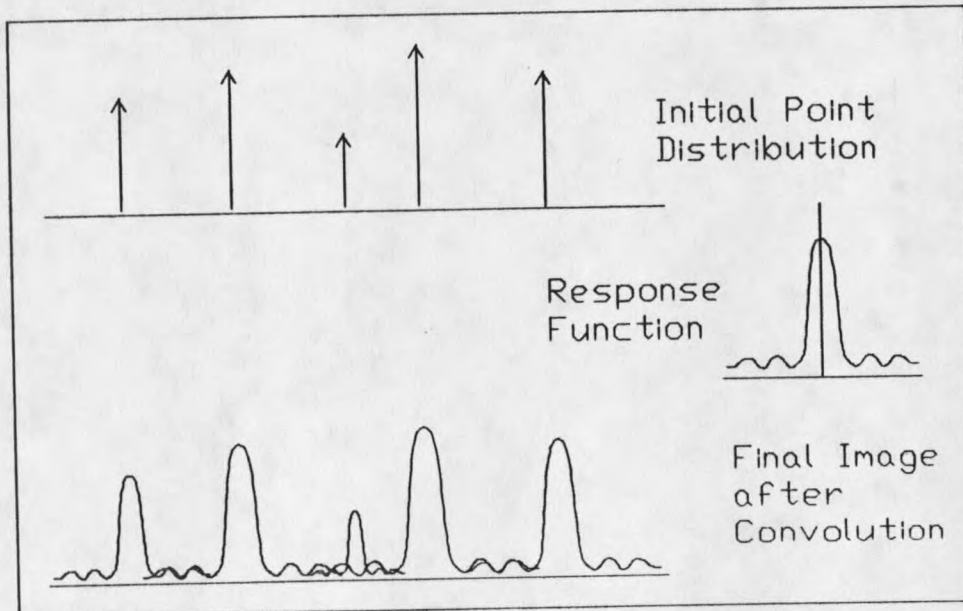


Figure 41 A one-dimensional convolution

In accordance with geometrical optics, the ideal response function appears as

$$h(x_o, y_o; x_i, y_i) = C \delta(x_i + Mx_o, y_i + My_o) \quad (D:2)$$

where C is a complex multiplicative constant. This response function simply maps each object point into an image point with the appropriate magnification factor. On the other hand, Fourier optics predicts the response function to appear as

$$h(x_o, y_o; x_i, y_i) = \frac{1}{\lambda^2 d_o d_i} \int \int_{-\infty}^{\infty} P(x, y) \exp\left\{ \frac{-i2\pi}{\lambda d_i} [(x_i + Mx_o)x + (y_i + My_o)y] \right\} dx dy \quad (D:3)$$

The response function is based on the geometry depicted in Figure 42. Here, the object plane is denoted as x_o, y_o , the image plane is denoted by x_i, y_i and the plane of the lens (the integration plane) is denoted simply as x, y . $P(x, y)$ is called the pupil function. It serves to limit the integral, as its value is 1 or 0 depending if the integration is occurring inside or outside of the boundary of the lens.

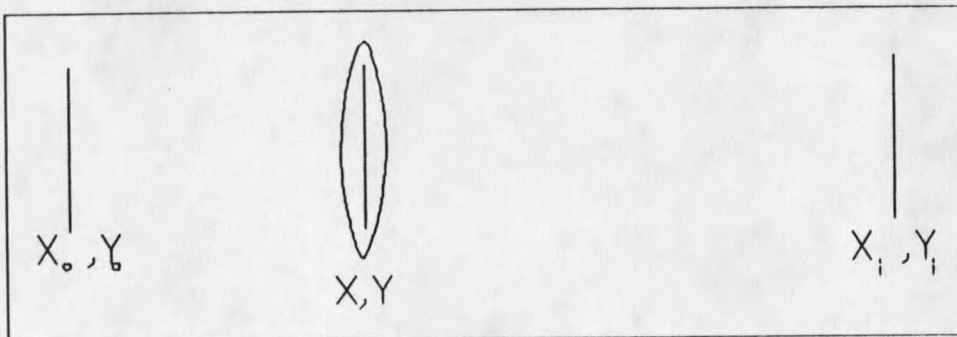


Figure 42 Geometry for calculating the response function

Consider equation D:3. If the pupil has the value of 1 everywhere, which describes an infinite lens, the integral would be identical to equation D-2. Hence, the fact that the

lens is finite leads to the response function's deviation from a delta function. Qualitatively, an expression for the response function is wanted. Also, it would be beneficial to be able to specify the width of the response function as it depends upon a critical parameter of the imaging system. The entrance f-number of the imaging system is the parameter to be used.

Starting with equation D:3, two assumptions will be made. The first assumption is that the response function over the entire aperture will be equivalent to the response function calculated along the axis of the system, hence $x_o=0$, $y_o=0$. Secondly, the magnification will be the ratio of the image and object distances, $M=d_i/d_o$. Now, the response function takes the form

$$h(x_o, y_o; x_i, y_i) = \frac{1}{\lambda^2 d_o^2 M} \int \int_{-\infty}^{\infty} P(x, y) \exp\left\{ \frac{-i2\pi}{\lambda d_o M} [x_i x + y_i y] \right\} dx dy \quad (D:4)$$

To do this integral the substitution

$$x_i = r_i \cos \theta_i \quad x = r \cos \theta \quad y_i = r_i \sin \theta_i \quad y = r \sin \theta \quad dx dy = r dr d\theta \quad (D:5)$$

is used to obtain

$$h(r_o, \theta_o; r_i, \theta_i) = \frac{1}{\lambda^2 d_o^2 M} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=a} \exp\left\{ \frac{-i2\pi r r_i}{\lambda d_o M} [\cos \theta \cos \theta_i + \sin \theta \sin \theta_i] \right\} r dr d\theta \quad (D:6)$$

The pupil function has now restricted the integral to inside the lens aperture. Due to the cylindrical symmetry, the response function will be radial only, hence θ_i is set to 0. The integral over $d\theta$ is separated to be done first

$$h(r_o, \theta_o; r_i, \theta_i) = \frac{1}{\lambda^2 d_o^2 M} \int_{r=0}^a r dr \int_{\theta=0}^{2\pi} \exp\left(\frac{ik}{d_o M} r r_i \cos\theta\right) d\theta \quad (D:7)$$

where the definition of $k=2\pi/\lambda$ is used. The θ integration can be found in standard integral tables. Now the radial integration appears as

$$h(r_o, \theta_o; r_i, \theta_i) = \frac{2\pi}{\lambda^2 d_o^2 M} \int_{r=0}^a r J_0(k'r) dr \quad k' = \frac{k r_i}{d_o M} \quad (D:8)$$

With the appropriate substitutions, one finds that the response function indeed takes on the form of the Airy pattern

$$h(r_o, r_i) = C \frac{J_1\left(\frac{2\pi r_i a}{\lambda d_o M}\right)}{\frac{2\pi r_i a}{\lambda d_o M}} \quad (D:9)$$

where C is a multiplicative constant. Finally, the f-number of the system is defined as

$f/\# = 2a/d_o$. The response function can now be expressed as a function of entrance f-number of the system,

$$h(r_o, r_i) = C' \frac{J_1(q \frac{r_i}{f/\#})}{\frac{qr_i}{f/\#}} \quad (D:10)$$

where C' and q are multiplicative constants.

At this point, the response function is known as a function of the entrance f-number of the optical system. In order to qualitatively inspect equation D-10, the f-number was varied from .3 to 7, and the appropriate airy patterns were generated. The airy patterns produced were compared directly with gaussian curves, since the main airy peak strongly resembles a gaussian. This comparison allowed a gaussian waist to be extracted. Then a plot of response waist vs entrance f-number was generated. This plot appears in Figure 43.

The result is that the response function varies linearly with entrance f-number. The slope of the line was calculated to be about $.6 \lambda$. The y-intercept is zero. Both of these qualities should make sense. If the system's entrance f-number is zero, then the lens is infinite in size, and the response function will be a delta function, or a gaussian with a waist which approaches zero. Also, a smaller wavelength should yield a better resolution, thus a smaller response waist. If two wavelengths are used, the smaller

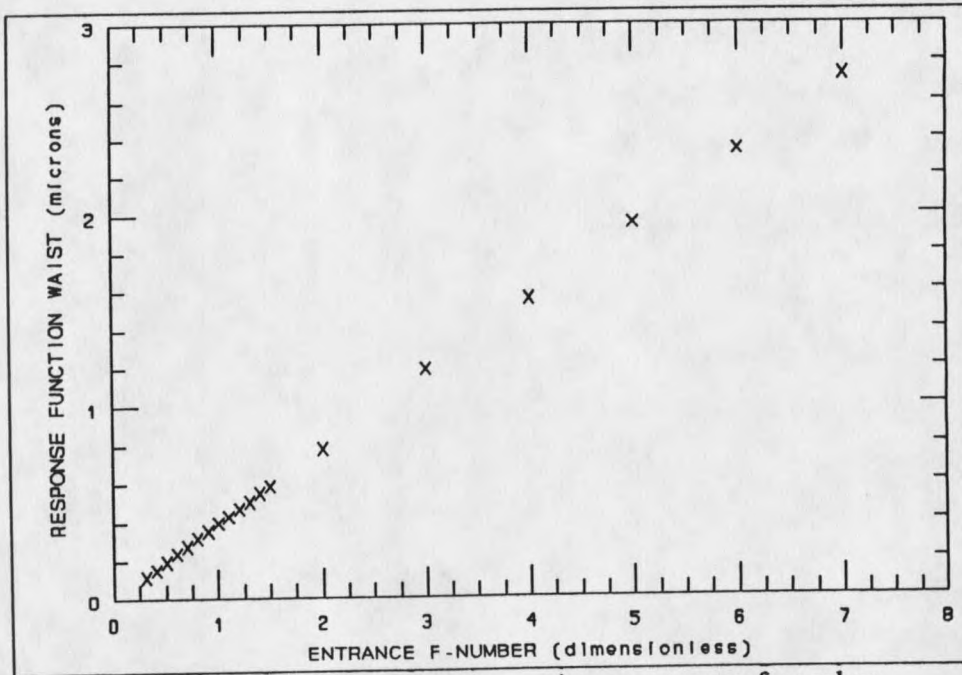


Figure 43 Plot of response function waist vs entrance f-number

wavelength can use less of the lens compared to the larger wavelength in order to obtain the identical resolution.

Finally, there had to be a way to test the resolution of the imaging system. To do this, a United States Air Force test target was obtained. The target consists of a glass plate with several sets of vertical and horizontal lines etched onto a chromium deposit. The target ranges from 1 line pair per millimeter to 228.1 line pairs per millimeter. Figure 44 illustrates the difference between a line set and a single line pair.

In order to test the imaging system, the test target had to be convolved with a response function to obtain an expected image. This process was accomplished through the computer program given in Appendix E. The program does the convolution expressed in equation (D:1). By programming a 1-dimensional slice of the field data

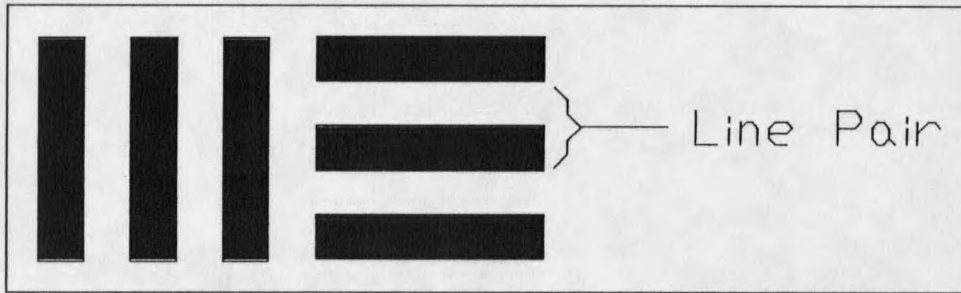


Figure 44 One test target line set with a line pair shown

with several gaussian response functions of different waists, several convolutions were obtained. The computer program produced Figure 45. Notice that the convolution simply acts to smooth the corners of the object. By comparing the relative size of the peaks and troughs of the computer image to a real image, one can obtain a good value for confirming the entrance f-number of the imaging system.

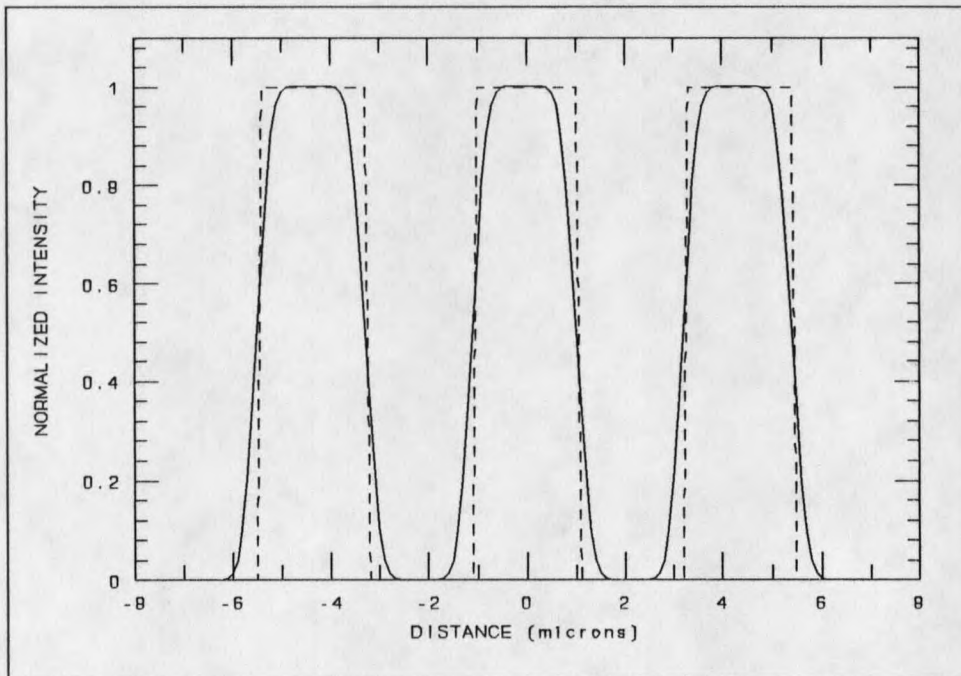


Figure 45 Computer simulated convolution of the 228.1 lp/mm test pattern

In Figure 45 notice that the response function simply provides a smoothing function.

The width between the individual test pattern lines is now less due to the response function. An entrance f-number of .65 was used to determine the gaussian response function. The test target was then imaged onto the diode array in the same fashion as listed in Chapter 3. Figure 46 shows the image obtained from the experimental setup.

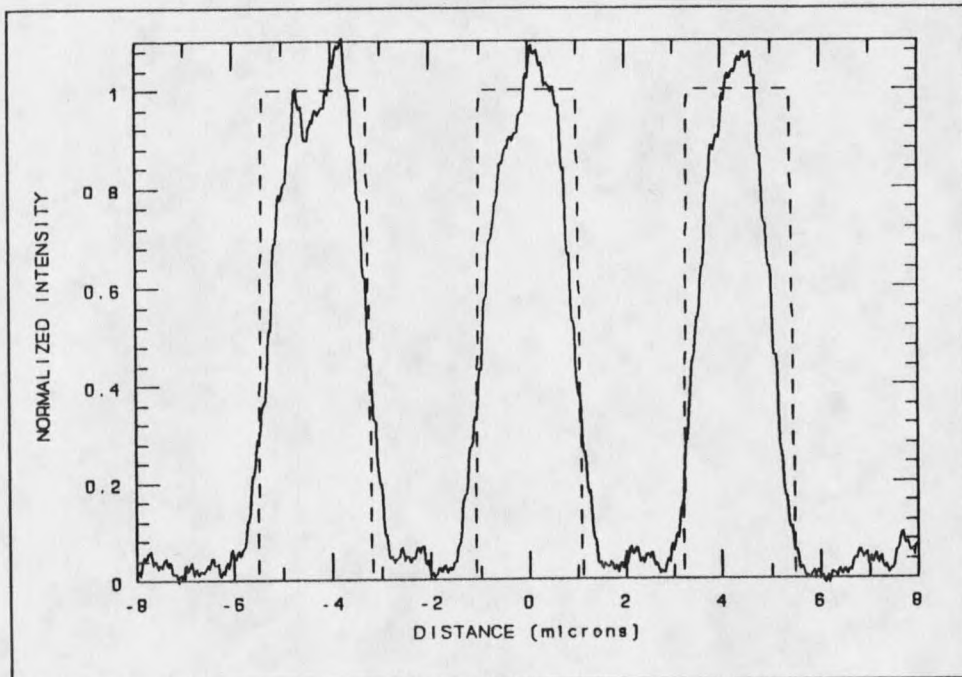


Figure 46 Experimental image of the 228.1 lp/mm test pattern

The experimental image corresponds quite well to the theory depicted in the previous figure. There is some additional "dirt" in this signal due to an interference pattern arising between the two faces of the glass which the target is deposited on. However, it should be noticed that the separation of the test lines is quite similar to that in the theoretical picture.

APPENDIX E**SOURCE CODE FOR COMPUTING THE CONVOLUTION INTEGRAL**

APPENDIX E

Figure 47

SOURCE CODE FOR COMPUTING THE CONVOLUTION INTEGRAL

```

C *****
C
C      PROGRAM:  CONV.FOR
C      AUTHOR:   Andrew Struckhoff
C      REVISED:  September 28,1991
C      DESCRIPTION: This program was created using several subroutines
C                   given in Numerical Recipes. It's purpose is to perform
C                   a convolution between two functions. The convolution
C                   occurs by fourier transforming each function independently,
C                   multiplying the fourier components together, then an inverse
C                   fourier transform completes the process.
C
C      NOTES:    The basic format comes from Numerical Recipes.
C                   However, the authors of the book failed to recognize several
C                   problems with their suggested code. The prescribes
C                   subroutines to be linked together were extremely
C                   incompatible. This program solves the incompatibility
C                   problem. Both the function array and response array
C                   MUST be 1024 characters in length, with the response array
C                   formatted in a wrap-around style, as given in Numerical
C                   Recipes.14
C *****

```

```

dimension data(1024),respns(1024),x1(1024),x2(1024),data5(2048)
character*30 oldfl,impuls,newfl

```

```

parameter(nmax=8192)
complex fft(1024), ans(1024), prod(1024)
n=1024
write(*,20)
20  format(' Enter Data File (MUST BE 1024 DATA POINTS)')
   read(*,'(a30)')oldfl
   write(*,*) ' Accessing file:',oldfl
   open(10,file=oldfl)
   read(10,*,end=30)(x1(i),data(i),i=1,1024)
30  write(*,40)
40  format(' Enter Response Function File')
   read(*,'(a30)')impuls
   write(*,*) ' Accessing Response Function File....',impuls
   open(20,file=impuls)
   read(20,*,end=50)(x2(i),respns(i),i=1,1024)

```

C I USED THE SUBROUTINE CONVLV(DATA,N,RESPNS,M,ISIGN,ANS) AS A
 C TEMPLATE FOR THE MAIN PROGRAM. THIS ROUTINE WILL CONVOLVE
 C THE DATA FILE WITH THE RESPONSE FILE. THE PROGRAM IS TAKEN
 C FROM 'NUMERICAL METHODS'. AT THIS POINT, ISIGN IS SET TO 1 SO
 C THAT THE TRANSFORM OF BOTH DATA SETS IS TAKEN, AND NOT THE
 C INVERSE TRANSFORM

```
50  isign=1
```

```
      call twofit(data,respns,fft,ans,1024)
```

C AT THIS POINT, THERE SHOULD BE TWO COMPLEX ARRAYS (fft) AND
 C (ans) THE NEXT STEP IS TO MULTIPLY THE COMPONENTS TOGETHER.
 C AFTER THIS STEP I WILL TRANSFORM THE COMPLEX ARRAY THAT IS
 C LEFT INTO A SINGLE REAL ARRAY THEN PUT IT THROUGH THE FOURI
 C SUBROUTING AGAIN WITH ISIGN = -1

C HERE IS THE MULTIPLICATION STEP

```

      do 85 i = 1,n
        prod(i)=fft(i)*ans(i)
85  continue

```

C NOW WE TRANSFORM THE COMPLEX ARRAY INTO A REAL ARRAY OF
 C LENGTH 2*N

```

      do 86 i=1,n
        data5(2*i-1)=real(prod(i))

```



```

      data5(2*i)=aimag(prod(i))
86   continue

```

C NOW THE ARRAY IS INVERSE TRANSFORMED TO YIELD THE
C CONVOLUTION

```

      call four1(data5,2*n,-1)

      write(*,90)
90   format(' Enter Name for Convolved Data File:')
      read(*,'(a30)')newfl
c    write(*,*)(data5(i),i=1,2048)
      open(30,file=newfl)
      write(30,100)(i,data5(2*i-1)/1024,i=1,1024)
100  format(i5,f8.4)
      write(*,*) 'Thank You and Enjoy Your Day at Wallyworld!!!'
      end

```

C NEXT I WILL PROGRAM IN THE DOUBLE FOURIER TRANSFORM
C PROGRAM WHICH SHOULD TRANSFORM THE THE TWO FUNCTIONS
C CALLED DATA AND RESPONSE

```

      subroutine twofit(data1,data2,fft1,fft2,n)
      dimension data1(n),data2(n), sntrmd(2048)
      complex fft1(n),fft2(n),h1,h2,c1,c2
      c1=cmplx(0.5,0.0)
      c2=cmplx(0.0,-0.5)

```

C HERE IS WHERE I HAVE DEVIATED FROM THE GIVEN TWOFFT ROUTINE.
C MY OPTION IS TO COMBINE THE TWO FILES INTO A DOUBLE LENGTH
C REAL ARRAY INSTEAD OF A COMPLEX ARRAY SINCE THE FOUR1
C PROGRAM DEALS WITH COMPLEX ARRAYS AS LENGTHENED REAL
C ARRAYS.

```

      do 44 i=1,1024
          sntrmd(2*i-1)=data1(i)
          sntrmd(2*i)=data2(i)
44   continue

      call four1(sntrmd,2048,1)

```

C NOW THE TRANSFORMED ARRAY IS RETURNED, HOWEVER, THE
C TWOFFT ROUTINE NEEDS TO SEPARATE THE ARRAYS. IT DOES THIS BY
C MAKING USE OF SYMMETRIES OF THE FOURIER TRANSFORM OF THE

C COMPLEX ARRAY. THIS NEXT PART SIMPLY INSTALLS THE
 C TRANSFORM IN A COMPLEX ARRAY, SO THAT THE ROUTINE CAN
 C HANDLE THE SEPARATION OF THE DATA TRANSFORM AND RESPONSE
 C TRANSFORM

```

      do 45 j=2,2*n,2
        fft1(j/2)=cmplx(snttrmd(j-1),snttrmd(j))
45    continue

```

C THIS PART OF THE TWOFFT SUBROUTINE IS UNTOUCHED

```

      fft2(1)=cmplx(aimag(fft1(1)),0.0)
      fft1(1)=cmplx(real(fft1(1)),0.0)
      n2=n+2
      do 12 j=2,n/2+1
        h1=c1*(fft1(j)+conjg(fft1(n2-j)))
        h2=c2*(FFT1(j)-conjg(fft1(n2-j)))
        fft1(j)=h1
        fft1(n2-j)=conjg(h1)
        fft2(j)=h2
        fft2(n2-j)=conjg(h2)
12    continue
      return
      end

```

C NEXT I WILL PROGRAM IN THE FOUR1 SUBROUTINE AS GIVEN IN
 C NUMERICAL RECIPES

```

      subroutine four1(data3,nn,isign)
      real*8,wt,wi,wpr,wpi,wtemp,theta
      dimension data3(nn)
      write(*,*) '*****FOUR1 SUBROUTINE ACCESSED*****'
      n=nn
      j=1
      do 11 i=1,n,2
        if (j.gt.i) then
          tempr=data3(j)
          tempi=data3(j+1)
          data3(j)=data3(i)
          data3(j+1)=data3(i+1)
          data3(i)=tempr
          data3(i+1)=tempi
        endif
      do 11 i=1,n,2

```

```

      m=n/2
1      if ((m.ge.2).and.(j.gt.m)) then
          j=j-m
          m=m/2
          go to 1
      endif
      j=j+m
11     continue
      mmax=2
2      if (n.gt.mmax) then
          istep=2*mmax
          theta=6.28318530717959D0/(isign*mmax)
          wpr=-2.d0*dsin(.5d0*theta)**2
          wpi=dsin(theta)
          wr=1.d0
          wi=0.d0
          do 13 m=1,mmax,2
              do 12 i=m,n,istep
                  j=i+mmax
                  tempr=sngl(wr)*data3(j)-sngl(wi)*data3(j+1)
                  tempi=sngl(wr)*data3(j+1)+sngl(wi)*data3(j)
                  data3(j)=data3(i)-tempr
                  data3(j+1)=data3(i+1)-tempi
                  data3(i)=data3(i)+tempr
                  data3(i+1)=data3(i+1)+tempi
12             continue
              wtemp=wr
              wr=wr*wpr-wi*wpi+wr
              wi=wi*wpr+wtemp*wpi+wi
13          continue
          mmax=istep
      go to 2
      endif
      write(*,*)'*****EXITING FOUR1 SUBROUTINE*****'
      return
      end

```

APPENDIX F

**COMPUTING THE CHANGE IN RADIUS OF CURVATURE OF A BEAM
PASSING THROUGH A GLASS PLATE**

APPENDIX F

COMPUTING THE CHANGE IN RADIUS OF CURVATURE OF A BEAM PASSING THROUGH A GLASS PLATE

Propagating a ray of light through an interface which has two different indices of refraction on either side causes the beam to change its direction. Similarly, when two different interfaces are encountered, the beam changes directions twice. If the first interface is from air to glass, and the second is from glass to air, there will be no net change in direction for the beam. However, the beam will not exit the glass at the same relative location as when it entered the glass. This problem was considered to calculate the change in radius of curvature of a gaussian beam which passes through a glass plate. Here, two ways of calculating the real radius of curvature will be presented. These two ways are by simple geometry and by propagation matrices.

The geometrical considerations are depicted in Figure 48. Consider a point source of light located at the point to the left. The rays of this source travel a distance d before encountering the glass mask. The thickness of the mask is denoted as t . The aperture radius is marked a . The propagation of the rays through the substrate make the beam appear to have a radius of curvature much shorter than the separation of the spatial

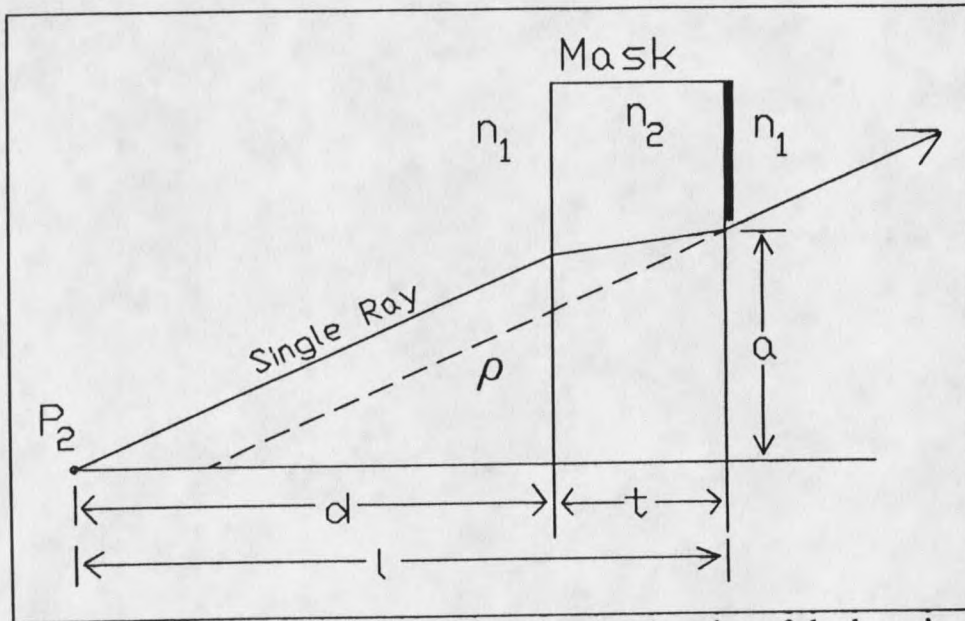


Figure 48 Geometrical considerations for the alteration of the beam's radius of curvature

filter and the plate. Using only the geometry and Snell's Law, the corrected radius of curvature appears in a parametric equation which is given as

$$\frac{d a}{(\rho^2 - a^2)^{1/2}} + \frac{t a n_1}{(n_2^2 \rho^2 - n_1^2 a^2)^{1/2}} = a \quad (\text{F:1})$$

The solution for ρ was obtained numerically.

Another approach to dealing with the glass plate is known as the propagation matrices. The details of propagation matrices can be found in general optics texts.^{1,2,13} This method for letting a beam move throughout space deals with representing a ray of light as a position and a slope. It is also shown that the matrices dealing with a pure geometrical propagation can be applied to the propagation of a gaussian beam.

Consider a ray of light moving in a vertical plane. At any one instance it has a

position, r and a slope, r' . The position and slope are represented by the matrix

$$\begin{vmatrix} r \\ r' \end{vmatrix} \quad (\text{F:2})$$

In order to allow the beam to move, an operation by another matrix is required. The operator matrix is a two-by-two matrix of elements A,B,C and D. For instance, the matrix which represents the propagation over a linear distance d is expressed as

$$\begin{vmatrix} 1 & d \\ 0 & 1 \end{vmatrix} \quad (\text{matrix for propagation over a distance } d) \quad (\text{F:3})$$

The operation is done according to the standard rules of matrix multiplication. The new ray matrix is written as

$$\begin{vmatrix} r_{\text{new}} \\ r'_{\text{new}} \end{vmatrix} = \begin{vmatrix} 1 & d \\ 0 & 1 \end{vmatrix} \begin{vmatrix} r \\ r' \end{vmatrix} = \begin{vmatrix} r+r'd \\ r' \end{vmatrix} \quad (\text{F:4})$$

The beam of light has merely changed its position, according to the slope of the ray, and the slope has remained the same. Figure 49 shows the propagation of the ray as described by the matrix.

In a similar fashion, other matrices exist which describe propagation through other elements in an optical system. This ABCD matrix can also be applied to a gaussian

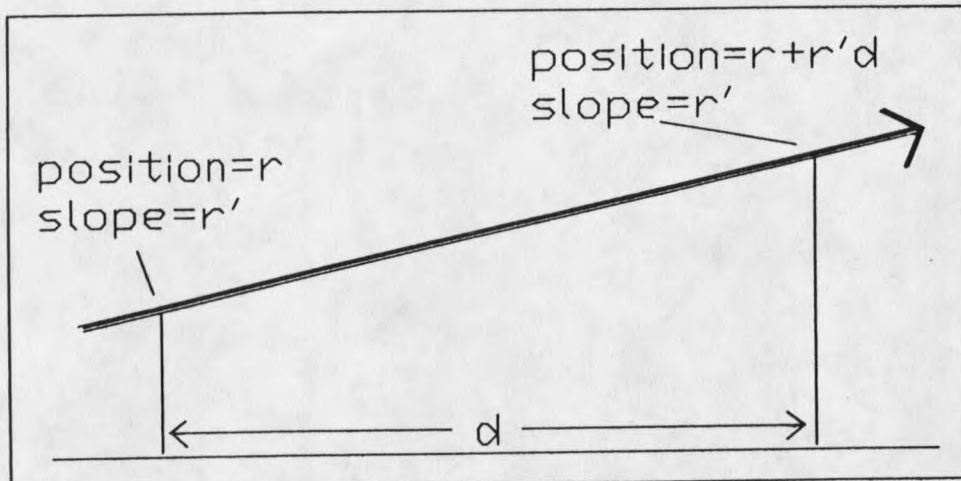


Figure 49 Propagation of a ray by the matrix method

beam. Instead of the slope and position of the ray being used, however, the gaussian waist and radius of curvature are used as the descriptive properties. The application of the matrix also appears in a different manner. To apply the matrices, the *q*-factor is used. This factor is expressed as

$$\frac{1}{q_i} = \frac{1}{R_i} - i \frac{\lambda}{\pi \omega_i^2 n} \quad (\text{F:5})$$

where R is the radius of curvature, ω is the waist of the beam and n is the index of refraction. When a particular ABCD matrix is applied, it is applied as

$$q_{\text{new}} = \frac{Aq_{\text{old}} + B}{Cq_{\text{old}} + D} \quad (\text{F:6})$$

After starting with an initial q , given by the initial radius and waist, the matrix

operations are done according to equation 3:9. A new waist and radius are then extracted through the q-factor. Now, similar to Figure 48, the gaussian beam starts at the spatial filter with a given waist. It propagates a distance d , it encounters the interface at the mask, it then propagates the distance t , then finally re-emerges on the other side of the mask. One nice application of this matrix idea is that the matrices can be multiplied together according to the usual matrix multiplication, before applying them to the beam itself. Thus, only one operation need be done on the beam, with a single matrix representing the entire path of propagation.

Figure 50 shows the propagation and the matrices now used at each part of the path taken by the beam of light. The radius of curvature can be extracted from each of

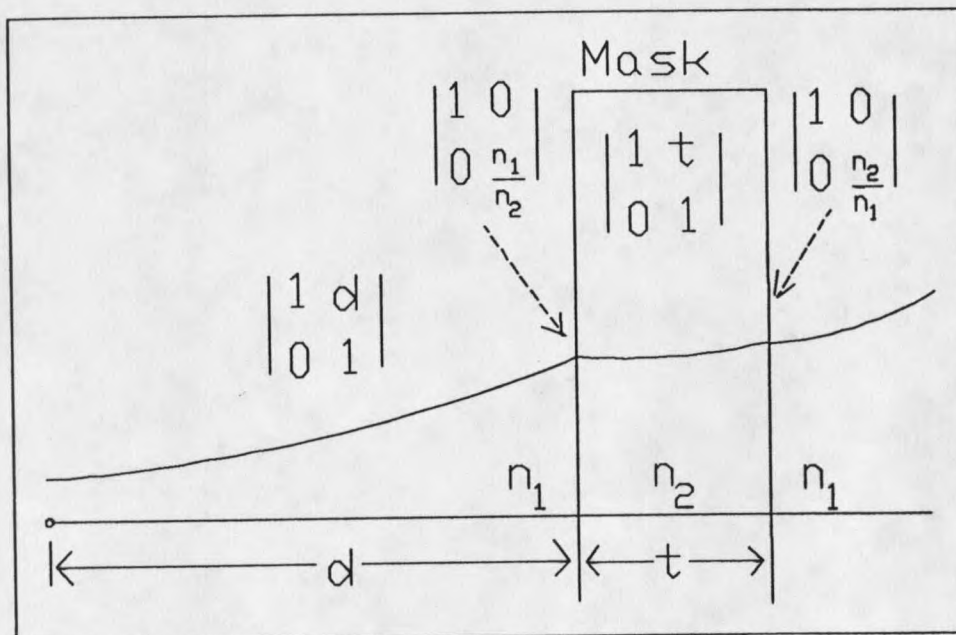


Figure 50 Matrices applied to a propagating gaussian beam

the two methods just presented. Consider the following initial data: $t = 6.32$ mm and $l = 40$ mm. The purely geometrical calculation yields $\rho = 13.2206$ mm and the matrix

calculation yields $\rho = 13.2251$ mm. Interestingly enough, the difference in the two individual methods is only .03 %. So both methods are equally applicable for the geometry being inspected. It will be shown that this amount of error in the radius of curvature produces no noticeable result in the theory.

REFERENCES CITED

REFERENCES CITED

1. E. Hecht, *Optics* (Addison-Wesley, MA, 1978)
2. M. Born, E. Wolf, *Principles of Optics* (Pergamon Press, England, 1965)
3. J.D. Jackson, *Classical Electrodynamics* (Wiley and Sons, NY, 1975)
4. C.F. Meyer, *The Diffraction of Light, X-rays and Material Particles* (University of Chicago Press, IL, 1934)
5. E. Wolf, E.W. Marchland, "Comparison of the Kirchhoff and the Rayleigh-Sommerfeld Theories of Diffraction at an Aperture", *J.Opt.Soc.Am.*, 54, 587 (1964).
6. J.W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, NY, 1968)
7. D.S. Burch, "Fresnel Diffraction by a Circular Aperture", *Am.J.Phys.*, 53, 255 (1985).
8. A.J. Campillo, J.E. Pearson, S.L. Shapiro, N.J. Terrel, "Fresnel Diffraction effects in the Design of High-Power Laser Systems", *Appl.Phys.Lett.*, 23, 85, (1973).
9. H.G. Kraus, "Huygens-Fresnel-Kirchhoff wavefront diffraction formulation - Spherical Waves", *J.Opt.Soc.Am.A*, 6, 1196 (1989).
10. H.G. Kraus, "Huygens-Fresnel-Kirchhoff wavefront diffraction formulation - paraxial and exact Gaussian Laser Beams", *J.Opt.Soc.Am.A*, 7, 47 (1990).
11. R.G. Wenzel, J.M. Telle, J.L. Carlsten, "Fresnel Diffraction in an Optical System Containing Lenses", *J.Opt.Soc.Am.A*, 3, 838 (1986).
12. D. Malacara, *Optical Shop Testing* (Wiley and Sons, NY, 1978).
13. A. Yariv, *Quantum Electronics* (Wiley and Sons, NY, 1989).

14. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1989).

MONTANA STATE UNIVERSITY LIBRARIES



3 1762 10182692 1

HOUCHEN
BINDERY LTD
UTICA/OMAHA
NE.