To the Graduate Council:

I am submitting herewith a dissertation written by Reinhard W. Menzel entitled "Fundamental Properties of Holographically Reconstructed Images." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Physics.

A. A. Mason, Major Professor

We have read this dissertation
and recommend its acceptance:

R. A. Beth

W. E. Deeds

J B. Dent

Accepted for the Council:

Hilton A. Smith
Vice Chancellor
Graduate Studies and Research

THESIS
75D
M359
Cop. 2

FUNDAMENTAL PROPERTIES OF HOLOGRAPHICALLY RECONSTRUCTED IMAGES

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee

Reinhard W. Menzel

December 1975

1269155

ACKNOWLEDGMENTS

The author is indebted to many people for their assistance in the course of this work. Preeminently, the author wishes to express his gratitude to his wife for sharing the vicissitudes of a graduate career and for aiding greatly in the preparation of the manuscript. Discussions with members of Advanced Concepts of ARO, Inc., helped to clarify many points pertaining to various aspects of the work; in particular, those with Dr. R. A. Belz and H. T. Bentley were most helpful. The author would like to express appreciation to A. E. Lennert, Branch Manager of Advanced Concepts, for his administrative support. Also, the dedicated and meticulous work of F. K. Hornsby, who helped with many details of the experimentation, is thankfully acknowledged.

The author is particularly indebted to his advisor, Dr. A. A. Mason, for his dedicated teaching, his discussions and suggestions during the course of this work, and for his help in bringing the work to a successful conclusion. Special appreciation is offered to Judy Wright for her careful typing of the manuscript.

The work reported herein was sponsored by Headquarters, Arnold Engineering Development Center, Arnold Air Force Station, Tennessee, within the terms of Contract

F40600-75-C-0004 with ARO, Inc., under Project B32S-04A.
Further reproduction is authorized to satisfy the needs of the U.S. Government.

ABSTRACT

The fundamental characteristics of images obtained by holographic reconstruction are presented. These characteristics are discussed in terms of a holographic reconstruction model based on scalar diffraction theory. The model is used to derive the point and line spread functions and the images of a few experimentally important objects which have analytical solutions. The properties which characterize the details of a reconstructed image are shown to be multiparametric functions. The parameters which most significantly influence the resolution and the intensity within the image are shown to be the object-to-reference beam intensity ratio and the hologram size. The experimental variables which influence the hologram size are discussed in detail. Finally, by the linear systems theory, all the holographic image characteristics are implicit in the spread function of the system in question. For that reason, some holographic system spread functions are experimentally obtained, and the results are compared with the theoretical model and with experimentally observed resolution characteristics.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
II.	SCALAR DIFFRACTION THEORY	7
III.	CHARACTERISTICS OF HOLOGRAPHIC SPREAD FUNCTIONS	14
IV.	CHARACTERISTICS OF IN-LINE AND OFF-AXIS RECONSTRUCTIONS	29
V.	RECONSTRUCTION DETAIL AND HOLOGRAM SIZE	54
VI.	IMAGE CLARITY AND RESOLUTION	85
VII.	EXPERIMENTAL DETERMINATION OF HOLOGRAPHIC LINE SPREAD FUNCTIONS	111
VIII.	CONCLUSIONS AND DISCUSSIONS	139
BIBLIOGRAPHY		143
VITA		147

LIST OF FIGURES

FIGURE		PAGE
1.	Definition of Parameters	17
2.	The Reconstructed Images of a Microscope Reticle Which Show the Reversed Field Intensity of In-Line and Off-Axis Holography	38
3.	The Off-Axis Reconstruction of a Resolving Power Test Target Which Shows the Intensity Detail Within the Image of Large Objects	39
4.	The In-Focus, Off-Axis Reconstructions of a Circular Aperture from Holograms of Different Sizes	42
5.	The Out-of-Focus, Off-Axis Reconstructions of a Circular Aperture	46
6.	A Comparison of the In-Line and Off-Axis Reconstructions of the Resolving Power Test Target Which Illustrates the Detrimental Effects of the Spatially Unseparated Images of the In-Line Reconstruction	50
7.	The In-Focus, Off-Axis Reconstructions of a Slit Aperture from Holograms of Different Sizes . . .	52
8.	Theoretical Normalized Intensity of the In-Focus Reconstructed Image of a Particle (25)	55

FIGURE PAGE

9. The Intensity Near the Edge of the Reconstructed Image of a Particle 58

10. The Slope of the Intensity Near the Edge of a Reconstructed Image of a Particle 59

11. The Normalized Intensity at the Center of the Reconstructed Image of a Particle 61

12. Diffraction Efficiency as a Function of Exposed Film Density and the Object-to-Reference Beam Intensity Ratio β Obtained from Holograms for Which $\psi = 3°$ 67

13. The Amplitude Transmission of the Holograms as a Function of the Exposing Energy and the Object-to-Reference Beam Intensity Ratio β . . . 68

14. The Diffraction Efficiency Versus β Curves Obtained from the $\psi = 3°$ Holograms, Which Show the Relative Independence of the Diffraction Efficiency from the Exposed Film Density 70

15. Reconstruction from a Thick Hologram 72

16. Diffraction Efficiency as a Function of Exposed Film Density and the Object-to-Reference Beam Intensity Ratio β Obtained from Holograms for Which $\psi = 90°$ 75

FIGURE PAGE

17. The Diffraction Efficiency Versus β Curves Obtained from the ψ = 90° Holograms, Which Show the Relative Independence of the Diffraction Efficiency from the Exposed Film Density 76

18. The Normalized Energy Within the Entire Image (w/a = 0) and Near the Edge of the Image (w/a = .9) of a Generalized Particle Reconstruction 79

19. The Normalized Energy Within the Entire Image (w/a = 0) and Near the Edge of the Image (w/a = .9) of a Generalized Wire Reconstruction 82

20. The In-Focus, Off-Axis Reconstructions of a Circular Aperture for Different Reference Beam Intensities and β 88

21. Comparison of the Photographic and Off-Axis Reconstructed Image of the Resolving Power Test Target 91

22. The Off-Axis Reconstruction of the Resolving Power Test Target for Different β and Normal Exposure of the Object and Reference Beams . . . 93

23. The Off-Axis Reconstructions of the Resolving Power Test Target for Different β, Under-Exposed Reference Beams, and Normally Exposed Object Beams 94

FIGURE PAGE

24. The Off-Axis Reconstructions of the Resolving Power Test Target for Different β, Normally Exposed Reference Beams, and Over-Exposed Object Beams 95

25. The In-Line Reconstruction of the Resolving Power Test Target for Different Exposed Film Densities 99

26. The Image Intensity of a Bar Chart and Its dc Filtered Image for (a) an Idealized Case With Infinite Resolution and (b) a Realistic Case With Finite Resolution 101

27. Experimental Determination of the Resolution Defining Parameter η 104

28. The Intensity Distribution Resulting from Two Line Spread Functions Separated by η = 2.78 . . . 105

29. Microdensitometer Traces of (a) a Knife Edge and (b) and (c) the Reconstructed Image of a Knife Edge from the A and B Holograms Respectively . . 124

30. The Modulation and Phase Transfer Functions of the Microdensitometer 125

31. The Normalized Line Spread Function of the Microdensitometer; the Solid Line Is the Inverse Fourier Transform of the System's OTF, and the Circles Are a Direct Measurement of the Slope of the Knife Edge Trace 128

FIGURE PAGE

32. The Modulation and Phase Transfer Functions of the A Hologram Minus the Microdensitometer Effects 130

33. The Modulation and Phase Transfer Functions of the B Hologram Minus the Microdensitometer Effects 131

34. The Normalized Line Spread Function of the A Hologram Minus the Microdensitometer Effects 132

35. The Normalized Line Spread Function of the B Hologram Minus the Microdensitometer Effects 133

36. The Image of a 30μ Wide Bar Obtained from the Convolution Integral Using the A Hologram Line Spread Function 137

NOMENCLATURE

A	Object beam amplitude
a	Particle radius
B_1	Reference beam amplitude
B_2	Reconstruction beam amplitude
C	General constant
d	Fringe spacing
E	Energy
$\overline{E}$	Electric field vector
E_N	Normalized energy
F_1, F_2	Reconstruction integrals
$\mathfrak{F}$	Fourier transform
f_s	Spatial frequency
G	Green's function
H	One-half the length of the recording medium
I	Intensity
I_H	Intensity in the hologram plane
I_1	Intensity distribution in the hologram plane from an off-axis particle
I_2	Intensity distribution in the hologram plane from an in-line particle
I_3	Intensity distribution in the hologram plane from an off-axis aperture
i	$\sqrt{-1}$
$i(x_2, y_2)$	Image function

J_n	nth order Bessel function
K	General constant
k	Propagation vector $2\pi/\lambda$
L	One-half the hologram width in the x_2 direction
LSF	Line spread function
M	Transfer function
MTF	Modulation transfer function
m	Integer
N	Integer
n	Integer
$\hat{n}$	Unit vector
$o(x_1, y_1)$	Object function
P	Point of observation
PTF	Phase transfer function
p	A fraction
$\bar{p}$	Dipole moment
R	Hologram radius
r_1	Radial distance in the (x_1, y_1) plane
r_2	Radial distance in the (x_2, y_2) plane
r_3	Radial distance in the (x_3, y_3) plane
S	Spread function
s	Distance from a radiating point source
s_1	Distance between points in the (x_1, y_1) and (x_2, y_2) planes
s_2	Distance between points in the (x_2, y_2) and (x_3, y_3) planes
TT	Air Force Resolving Power Test Target 1951

t	Film thickness
U	Radiation field amplitude
U_H	Holographic amplitude transmission field
U_R	Reconstruction amplitude field
U_1	Reconstruction amplitude field from the hologram of I_1
U_2	Reconstruction amplitude field from the hologram of I_2
U_3	Reconstruction amplitude field from the hologram of I_3
u	Generalized diffraction parameter
v	Generalized diffraction parameter
w	Point on the r_3 radius
x	Generalized coordinate
(x_0, y_0)	Light source plane
(x_1, y_1)	Aperture or diffracting plane
(x_2, y_2)	Film or hologram plane
(x_3, y_3)	Reconstructed image plane
z	Generalized distance
z_1	Distance between the (x_1, y_1) and (x_2, y_2) planes
z_2	Distance between the (x_2, y_2) and (x_3, y_3) planes
α	kaR/z
β	Object-to-reference beam intensity ratio
γ	Slope of film sensitivity curve
Δ	An increment, use with another parameter
ε	Nondimensional variable of integration
ζ	A fixed z distance

η	kx_3L/z_2
θ_i	Incident wave angular direction
θ_r	Reflected wave angular direction
λ	Wavelength
λ_1	Wavelength during holographic recording
λ_2	Wavelength during holographic reconstruction
μ	Micron = 10^{-4} cm
ρ	Nondimensional variable of integration
σ	Impulse response
Φ	Phase
ϕ_1	Angular direction in the (x_1, y_1) plane
ϕ_2	Angular direction in the (x_2, y_2) plane
ϕ_3	Angular direction in the (x_3, y_3) plane
$\phi(x_1)$	Generalized phase function
χ	Difference coordinate, $x_2 - x_1$
ψ	Angle between the object and reference beams
ψ_1	Angle of the reference beam
ψ_2	Angle of the reconstruction beam
Ω	Amplitude of the modulation

CHAPTER I

INTRODUCTION

The work in this investigation is a result of part of the research and application of holographic technology at Arnold Engineering Development Center (AEDC), Arnold Air Force Station, Tennessee. As such, the theoretical discussions and experimental results presented within these pages are derived from specific needs for detailed information that became apparent during the development of various holographic visualization and data gathering techniques.

One technique that has been used extensively at AEDC since the inception of the holographic program in 1967 is the holographic recording of small particle fields. The facility with which holography allows the reduction of size, position, distribution, density, and other information without physically interfering with the conditions which established the particle fields has made this technique a keenly appreciated tool for analyzing various flow diagnostic problems. Some of the successful applications of particle field holography have been the evaluation of dust erosion facilities (1),[1] the simulation of icing conditions on an

---

[1]Numbers in parentheses refer to similarly numbered references in the Bibliography.

air breathing engine in a high water droplet flow rate (2), the interaction of reacting liquid sprays (3), and the visualization of combustion phenomena (4).

Three essential questions which resulted from the data reduction of each of these experiments and which effectively set boundaries to the application of the holographic format were:

1. What particle field densities may be allowed in a holographic recording so that unambiguous data may be obtained?
2. What are the optical characteristics of the reconstructed images as they apply to data recognition?
3. How are these properties to be applied in determining the resolution available from the holographic imaging systems?

After some preliminary analysis on the first question, it was apparent that the details relating to questions two and three were necessarily prerequisite. Indeed, the problem proved to be more intransigent than anticipated. A mathematical model based on recording a minimum of diffracted information from each diffracting center could either be solved only for nonrealistic assumptions about the uniformity and homogeneity of the size and spacing of the particles, or else it required more variables than could be solved with the known functional relationships. A literature survey did not reveal any similar or more tractable

models and thus simply served to reinforce the impressions about the problem's difficulty. The approach method of others either assumed conditions suitable for a statistical analysis (5) or simply favored the empirical solution obtained from observational deductions (6).

From this perspective, the problems of the second and third questions were addressed with the sense of re-examining the properties of holographic imaging. Specifically, the work within these pages incorporates a theoretical and experimental study of the fundamental characteristics of a holographically reconstructed image, the comparison of these characteristics with other optical imaging systems, and the isolated effects of various significant parameters. No further discussion is presented relative to particle field densities since that problem was judged to be beyond the scope of this work. Nevertheless, the impetus of this work stems from that problem, and the emphasis in the choice of the direction of the research comes from the anticipation of its eventual application. In this manner, the parameters chosen for study proved not to have been discussed sufficiently in the open literature. A general list of references on holographic phenomena is not presented at this point--as is otherwise customary--because such a list is easily found in almost any book on holography. The papers relating directly to the present work are referenced throughout the body of this text.

Since the theoretical model of holography is based on diffraction phenomena, Chapter II discusses the general validity of the fundamental diffraction theory. The emphasis of Chapter II is that classical scalar diffraction theory has a more solid theoretical foundation than is generally stressed in textbooks on physical optics. A conceptual derivation of the basic equations, which shows the physical reasons for the mathematical form which diffraction theory takes, is also discussed.

A mathematical description of the holographic recording and reconstruction in terms of scalar diffraction theory is presented in Chapter III. The emphasis is on deriving the point and line spread functions of the reconstruction mode as they can be used by a linear systems approach to holographic analysis. The similarity of the derived spread functions with those that are obtained from imaging properties of lenses is stressed. This result is quite important since it places the holographic imaging phenomenon as an "equal" member in the larger class of optical imaging systems. The mathematics of this chapter introduces the nomenclature and the geometry for the subsequent experiments. For this reason, it is presented rather early in the text even though the derivational results are not juxtapositioned with experimental results until Chapter VII.

The mathematical model of holography is applied in Chapter IV to deriving the characteristics of the

reconstructed images of small particles for both the in-line and off-axis configuration. Besides representing an object field of actual experimental importance, the small particle has a sufficiently simple geometry so that a closed form solution of the mathematical model may be obtained. Experimental results are shown which verify the analytical predictions for in-focus and out-of-focus properties of the reconstructed particle images. The similarity, stressed in Chapter III, of the image properties obtained from the in-line and off-axis holographic configuration is presented in detail.

The most important boundary condition with respect to diffraction phenomena is the aperture which is considered to cause the diffraction. During the reconstruction mode, this aperture is the size of the area on the hologram which gives an effective contribution to the reconstructed image. In Chapter V, some of the parameters influencing the hologram size are discussed. These include the intensity and spatial resolution of the film used and the diffraction efficiency which may be expected in the reconstruction mode as a function of the film exposure level and object-to-reference beam intensity ratio. The effect of these parameters on the clarity of detail in the reconstructed image is also discussed.

Chapter VI extends the results of Chapter V to include the specific effects observed in the image clarity when the recording parameter was such that linearity in the

recording medium was forfeited. The results show that clear images with higher reconstruction intensities can be obtained in the nonlinear recording configuration than can be obtained by the more conventional techniques where the linearity constraint is imposed. The second part of Chapter VI discusses the experimentally observed resolution in the holographic reconstruction. Comparing the experimental results with the theoretical model shows that the Sparrow resolution criterion accurately describes the resolutions observed. The experimental results of the focal tolerances in the reconstruction are also presented.

In Chapter VII, the application of the spread function by using the linear systems approach is discussed. The experimentally obtained line spread function of the real image reconstruction of a holographic system is presented. The resolutions implied by the experimental results and by the theoretical model are found to be similar. The significance of these results for considering the holographic imaging technique as a member of a larger class of imaging techniques rather than as unique onto itself is discussed.

Finally, in Chapter VIII, a synopsis of the results of this work is presented together with suggestions for possible future work.

CHAPTER II

SCALAR DIFFRACTION THEORY

The propagation of light diffracted by an object is governed by the Maxwell equations. These are coupled vector equations describing the electric and magnetic fields, $\overline{E}$ and $\overline{H}$, subject to the boundary conditions imposed by the diffracting object. Their rigorous solution is mathematically complicated and in only a very few cases can complete solutions be realized. In contrast, scalar diffraction theory, which neglects the vector properties of diffracted light, allows for relatively simple solutions to a broader range of diffraction problems. As the scalar theory was historically developed, the justification of its applications was generally neglected except for the heuristic argument that it led to remarkably accurate results when applied within certain approximation limits.

The scalar representation of the electromagnetic fields of light has been extensively and thoroughly investigated (7, 8, 9, 10, 11). Theimar, et al. (7) justified the calculation of the intensity in optical images from the square of a single scalar wave function subject to the approximation of the far-field dipole radiation. Green and Wolf (8) showed that, in a region free of currents and charges, the electromagnetic field may be derived rigorously from a single complex scalar wave function. Wolf generalized

this derivation (9) by deriving a new (1959) vector representation of an unpolarized quasi-monochromatic electromagnetic field. He thereby showed that the complex potential of this representation is identifiable with the complex disturbance of classical scalar diffraction theory subject to the "usual conditions." Wolf states (9):

> This analysis places the optical scalar theory on a rigorous foundation and the theory is found to have a much wider range of validity than is usually believed to be the case.

Wolf (10) further derives a description of the image space of an optical system in terms of an integral representation that has a physical interpretation in terms of a modified Huygens-Fresnel principle. Richards and Wolf (11) apply these results of imaging optics to lens systems. For the limiting case of narrow angular spreads about the light propagation direction, they show that the integral representation is then completely characterized by a complex scalar function identical with scalar theory. The significance of this result is that they have arrived at the scalar theory from an investigation of the structure of the electromagnetic field near the focus of a point source imaging system. They also determine that the light source polarization is preserved during diffraction when conditions satisfy the scalar theory approximations. This result is most applicable when considering diffraction of highly polarized laser light.

Clearly, the foundations of scalar diffraction theory are satisfactorily traced to first principles. Nevertheless, a conceptual derivation of the theory is useful, particularly for keeping light source functions categorically separated from functions generated in the process of mathematical solution (i.e., Green's functions). Several contributions are particularly important. Weiner's experiments demonstrated that the photochemical action of film is an $\overline{E}$ field energy detector, thereby suggesting that the $\overline{H}$ field can be neglected. Richards and Wolf (11) show analytically that for an optical imaging system--not just restricted to small apertures--the time-averaged magnetic energy density is the same as the time-averaged electrical energy density, but the distributions are orthogonally independent. Maxwell's equations require that the $\overline{E}$ field satisfies the wave equation

$$\nabla \cdot \nabla\overline{E} = \varepsilon_o \mu_o \frac{\partial^2 \overline{E}}{\partial t^2} \qquad (2\text{-}1)$$

which separates into three scalar equations involving a scalar Laplacian in Cartesian coordinates. Scalar quantities now serve as a description for electric energy detector systems.

Historically, Huygens' principle served as the basis of scalar diffraction theory which was mathematically formulated by Kirchhoff and Helmholtz using Green's theorem.

Green's theorem can be derived from Gauss' theorem.

$$\int_S \bar{A} \cdot d\bar{a} = \int_V \nabla \cdot \bar{A}\, dv \qquad (2\text{-}2)$$

which gives the total rate of flow of a vector "fluid" out of a volume of integration. By analogy, Green's theorem,

$$\int_V (U\nabla \cdot \nabla G - G\nabla \cdot \nabla U)dv = \int_S (U\nabla G - G\nabla U)da \qquad (2\text{-}3)$$

is interpreted to give the rate of radiation across the surface of a volume of integration due to scalar sources of functions U and G. Using a Green's function of the form,

$$G = \frac{e^{iks}}{s} \qquad (2\text{-}4)$$

where $k = \frac{2\pi}{\lambda}$, the Helmholtz-Kirchhoff theory gives the light amplitude $U(P_2)$ at an observation point P_2 due to radiation from a source function $U(P_1)$ at point P_1 as,

$$U(P_2) = 1/(4\ \pi) \int_{area} \left[\frac{\partial u(P_1)}{\partial n} G - U(P_1)\ \frac{\partial G}{\partial n}\right] da \qquad (2\text{-}5)$$

where $\hat{n}$ is normal to the volume surface. The basic assumption that a source in a finite region should produce only outward traveling waves imposes the Sommerfeld radiation condition:

$$\lim_{s\to\infty} \frac{1}{s}\left(\frac{\partial u(s)}{\partial n} - ikU(s)\right) = 0 \qquad (2\text{-}6)$$

This states that there are no radiation sources at infinity. Consequently, the above integration area can be reduced to an aperture area.

In their excellent introductory discussion of Green's functions, Morse and Feshbach (12) point out that Green's functions can be used to solve field problems where either the source or the boundary conditions are given. This gives the Green's function the dualism of depicting source distributions for certain boundary conditions or vice versa. In the diffraction case, the boundary conditions of the diffracting aperture dictate corresponding "source" distributions. This is analogous to obtaining image charges when describing the field of a charge near a conducting surface held at a fixed potential. That the Green's function "image sources" has the mathematical form given above is conceptually understood by comparison with the radiation in the far-field from an electric dipole,

$$\bar{E} = k^2 \frac{\exp(iks)}{s} (\hat{n}x\bar{p})x\hat{n} \qquad (2\text{-}7)$$

where $\bar{p}$ is the dipole moment. This analogy further suggests that scalar diffraction theory should be restricted to the radiation zone restrictions. Thus, the radiation wavelength should be small compared with the diffracting aperture which, in turn, should be small compared to the observation distance.

The Fresnel-Kirchhoff diffraction theory gives very satisfactory results. However, Kirchhoff's derivation is not entirely satisfactory since it requires both $U(P_1)$ and $\partial U(P_1)/\partial n$ to be specified in the diffracting aperture. From potential theory, it is known that a potential is zero over an entire plane if, on any segment of the plane, both the Dirichlet and Neumann boundary conditions are satisfied. This is, of course, contrary to the physical conditions of the diffraction phenomenon. Sommerfeld was able to remove this inconsistency from Kirchhoff's theory by showing that a Green's function of the form,

$$G = \frac{\exp(iks)}{s} \pm \frac{\exp(iks')}{s'} \tag{2-8}$$

satisfies the requirements of the problem when using either of the two boundary conditions. The two sources of this Green's function are mirror images of each other with respect to the aperture plane, s and s' being distances of equal magnitude oppositely directed from the aperture plane. With the additional paraxial requirement of considering only small solid angles centered on the direction of the source propagation through the aperture, the diffracted fields predicted by each of the theories are given by

$$\text{Kirchhoff:}\quad U(P) = -\frac{iA}{2\lambda} \int_{\text{aperture}} \frac{e^{ik(s_0+s_1)}}{s_0 s_1}[\cos(\hat{n} \cdot \bar{s}_1) - \cos(\hat{n} \cdot \bar{s}_0)]da \tag{2-9}$$

and

$$\text{Sommerfeld:}\quad U(P) = -\frac{iA}{\lambda}\int_{\text{aperture}} \frac{e^{ik(s_0+s_1)}}{s_0 s_1}\cos(\hat{n},\bar{s}_1)\,da \qquad (2\text{-}10)$$

where s_0 (s_1) correspond to the illuminating source (Green's) function radiation distance. The difference between these two fields is still a source of active investigation (13). However, within the paraxial approximation, the cosine terms are replaced by unity (of appropriate sign); and the two theories become the same.

The Sommerfeld diffraction integral provides the basis for the holographic derivations within these pages.

CHAPTER III

CHARACTERISTICS OF HOLOGRAPHIC SPREAD FUNCTIONS

By the principle of superposition for linear systems (14, 15), the presence of any one excitation is not affected by the response due to other excitations. A complicated input function can be considered to be composed of multiple simpler functions such as the components of a Fourier spectrum, and the response of the system can be studied for each component. One such approach is to study the system's transform characteristics to an impulse function input. The output due to a complicated input is then the superpositioning integral (mathematically, the convolution) of the impulse transform function over the input function domain. This is discussed at greater length in Chapter VII.

The optical analogue of an impulse function is a point source of radiation. Invoking Huygen's principle, a radiating surface illuminance is the integration of a point source over the radiating area. By considering an optical system's transfer response to a point radiating source, the essential features of the system's resolution characteristics can be found.

To describe the point spread function (or line spread function) of a holographic system, it is necessary to formulate the holographic recording of a point source of

radiation (or a line source of radiation) and the subsequent reconstruction of the image of the point source (or line source) from the hologram. Using scalar diffraction theory, the point source is described by Green's function,

$$G = A \frac{e^{iks}}{s} \tag{3-1}$$

where A is the amplitude and s is the distance between the point source and the point of observation (i.e., detection).

Since the derivations of this chapter use a plane reference beam, it is necessary to point out that Green's function is not adequate to describe plane wave radiation. Plane waves would, by Green's function, need to originate from a point infinitely distant from the observation plane; the amplitude would be reduced to zero. The essential amplitude and phase information of plane waves is adequately described by

$$G_p = B \cdot e^{i\Phi} \tag{3-2}$$

where Φ is defined geometrically relative to an arbitrary plane along the propagation direction. By the Poynting theorem, observed intensity is obtained from the product of the radiating field times its complex conjugate. The units of A and B are determined by requiring that the intensities resulting from plane waves and point source radiation be the same.

A point source of radiation can be approximated experimentally by the light emanating from an illuminated pinhole. The parameters for recording and reconstructing the hologram of such a point source are defined in Figure 1. The (x_i, y_i) planes with $i = 0, 1, 2, 3$ are respectively the light source plane, the aperture plane (containing the pinhole), the film plane in which the hologram is recorded, and the reconstructed image plane. The optical axes z of all the coordinate systems are collinear to satisfy the far-field radiation paraxial requirement. The wavelengths of light in the recording and reconstruction modes are λ_1 and λ_2. The distance between arbitrary points in the (x_1, y_1) and (x_2, y_2) planes is s_1; similarly, between the (x_2, y_2) and (x_3, y_3) planes it is s_2.

The holographic recording is made by combining in plane (x_2, y_2) the light from the point source located at $(x_1 = 0, y_1 = 0)$ and the plane wave propagating in the sense of the positive z direction at an angle ψ_1 in the (x, z) plane. Choosing the (x_2, y_2) plane for zero reference phase, the field in the hologram plane is

$$U = B_1 \exp(ik_1 x_2 \sin \psi_1) + (A/s_1) \exp(ik_1 s_1) \qquad (3\text{-}3)$$

By the far-field radiation condition, the radiation is restricted to the paraxial region. This, in turn, allows expanding s_1 by the binomial expansion

$$s_1 = (z_1^2 + r_1^2)^{1/2} = z_1 + 1/2(r_1^2/z_1) \qquad (3\text{-}4)$$

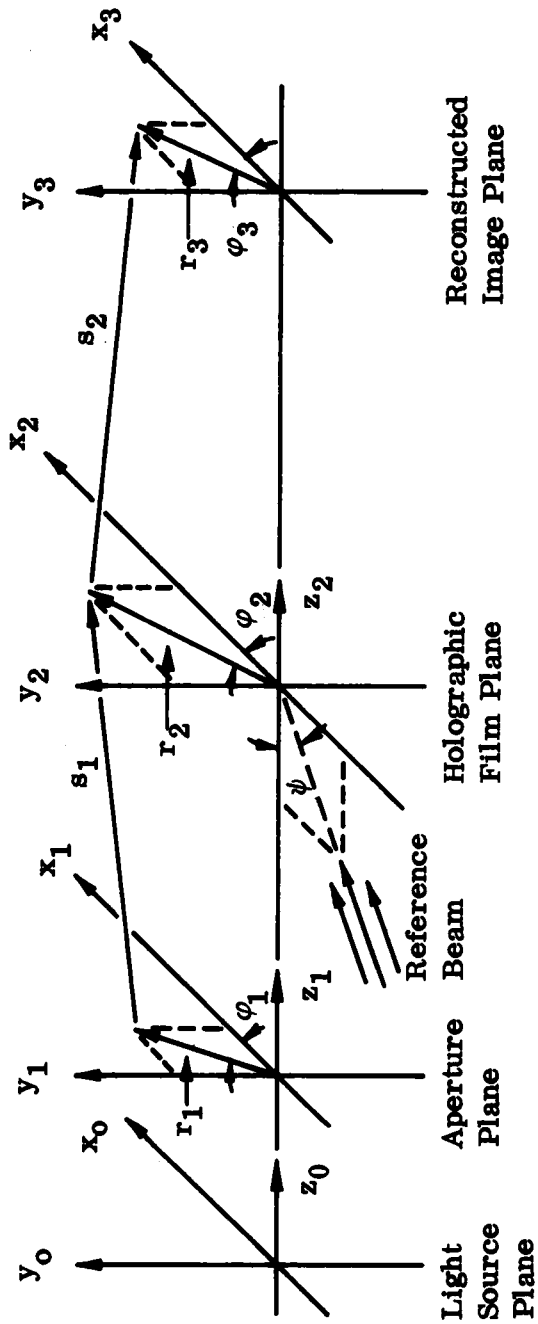


Figure 1. Definition of parameters.

where $r_2^2 = x_2^2 + y_2^2$. Higher order terms of the expansion are assumed to be small compared to λ_1 so that they contribute negligibly to the phase $k_1 s_1$. The intensity in the hologram plane is

$$I_H = B_1^2[1 + (A/z_1B_1)^2 + 2(A/z_1B_1)\cos(k_1z_1 + k_1r_2^2/2z_1 - k_1x_2 \sin\psi_1)] \quad (3\text{-}5)$$

When the plane reference beam is normally incident ($\psi_1 = 0$), the intensity distribution is similar to a Fresnel zone plate (16).

The reconstructed amplitude field $U_R(x_3, y_3)$ resulting from an illuminated hologram plane $U_H(x_2, y_2)$ is modeled on the Sommerfeld diffraction equation,

$$U_R = -\frac{i}{\lambda_2} \iint_{\text{aperture}} U_H \frac{e^{ik_2s_2}}{s_2} \cos(\hat{n}, \bar{s}_2)\, dx_2dy_2 \quad (3\text{-}6)$$

The hologram is centered on the origin of the (x_2, y_2) plane. U_H is the amplitude field transmitted by the illuminated hologram. It is given by the product of the incident field times the complex amplitude transmittance of the hologram. When random phase shifts due to random film emulsion thicknesses can be neglected, the hologram's complex amplitude transmittance is a linear mapping of the complex amplitude during exposure, provided the film exposure level is chosen properly (15). Then, for plane waves incident on the hologram at angle ψ_2,

$$U_H = [B_2 \exp(ik_2 x_2 \sin \psi_2)] \, K[I_H]^{-\gamma/2} \qquad (3\text{-}7)$$

where K is a proportionality constant and γ is the slope of the linear region of the film's density versus log exposure curves, the Hurter-Driffield curves. Assuming that the intensity can be written in the form $I = (1 + \delta)$, then for δ less than one, a binomial expansion of the intensity can be used. Of course, no approximation to a linear expansion is necessary if $\gamma = 2$, a condition which can be experimentally achieved. Techniques for linear optimization are described by Yu (17) based on maximizing the reconstruction in the first-order diffracted image. Generally, phase information is more important than amplitude information. The effects of nonlinear intensity recording causes several higher diffraction orders; however, the phase of the first-order diffraction is not affected (18). Amplitude distortion has little apparent effect visually. Therefore, when the first order diffraction can be spatially separated from the higher order diffraction modes, the nonlinearities of film have no serious deleterious effects on information recovery from the first-order image (17). This applies for off-axis holography; for in-line, the first-order and higher order images are superimposed, and the effects of the higher order images can be thought of as providing intermodulating distortion to the desired first-order image.

Using the exponential form of the trigonometric function in Equation 3-5, then

$$U_H = KB_2B_1^{-\gamma}\exp(ik_2x_2 \sin\psi_2)[1 - (\gamma/2)(A/z_1B_1)^2]$$
$$+ KB_2B_1^{-\gamma}\gamma(A/z_1B_1)\exp(ik_1z_1)\exp(ik_1r_2^2/2z_1)$$
$$\cdot \exp(ik_2x_2 \sin\psi_2 - ik_1x_2 \sin\psi_1) + KB_2B_1^{-\gamma}\gamma(A/z_1B_1)$$
$$\cdot \exp(-ik_1z_1)\exp(-ik_1r_2^2/2z_1)\exp(ik_2x_2 \sin\psi_2$$
$$+ ik_1x_2 \sin\psi_1) \qquad (3\text{-}8)$$

= (effect of dc bias on the film)
+ (the virtual image wave front contribution)
+ (the real image wave front contribution)

From the binomial expansion,

$$s_2 = z_2 + 1/(2\ z_2)[r_2^2 + r_3^2 - 2r_2r_3\cos(\phi_3 - \phi_2)] \qquad (3\text{-}9)$$

Also, the application of the paraxial requirement allows approximately

$$\cos(\hat{n}, \bar{s}_2) \simeq 1 \qquad (3\text{-}10)$$

Using the integral definition of the Bessel function, there results

$$U_R(r_3) = -i(k_2/z_2)\exp(ik_2z_2)\exp(ik_2r_3^2/2z_2)\int_0^R U_H J_0(k_2r_3r_2/z_2)$$
$$\cdot \exp(ik_2r_2^2/2z_2)r_2dr_2 \qquad (3\text{-}11)$$

where R is the maximum dimension of the hologram, effectively the aperture of the hologram. It is determined by a combination of the physical size of the film plane, the

film emulsion's amplitude sensitivity, its spatial frequency response, and the reference (and/or the reconstruction) beam cross-sectional area.

Using the change of variable, $r_2 = R\rho$, the reconstructed amplitude field can be written as

$$U_R(r_3) = \sum_{j=1}^{3} K_j \int_0^1 J_0(v\rho)\, \exp(iu_j\rho^2)\rho d\rho \qquad (3\text{-}12)$$

where

$$v = k_2 r_3 R/z_2$$

$$u_1 = x_2 k_2 \sin\psi_2$$

$$u_2 = (R^2/2)(k_2/z_2 + k_1/z_1) + x_2(k_2 \sin\psi_2 - k_1 \sin\psi_1)$$

$$u_3 = (R^2/2)(k_2/z_2 - k_1/z_1) + x_2(k_2 \sin\psi_2 + k_1 \sin\psi_1)$$

$$K_1 = C[1 - (\gamma/2)(A/z_1B_1)^2]$$

$$K_2 = C\gamma(A/z_1B_1)\exp(ik_1z_1)$$

$$K_3 = C\gamma(A/z_1B_1)\exp(-ik_1z_1)$$

$$C = -iKB_2B_1^{-\gamma}R^2(k_2/z_2)\exp(ik_2z_2)\exp(ik_2r_3^2/2z_2) \qquad (3\text{-}13)$$

The three integrals represent successively the reconstruction of the dc bias on the film, the virtual image, and the real image. Because of the angles ψ_1 and ψ_2, the three images are separated in space. The K_j contains information pertinent to the intensities of the images in the reconstruction. The ratio of the object-to-reference beam intensities is $(A/B_1)^2$. A proportional relationship of the reconstructed image intensities is

$$I_R \propto \gamma^2 (\pi R^2)^2 (A/B_1)^2 \qquad (3\text{-}14)$$

Thus, the reconstructed image intensities depend on the film contrast, the hologram area, and the ratio of the object-to-reference beam intensity ratio while making the hologram.

Considering only the real image reconstruction integral, considerable mathematical simplification results when

$$k_2 \sin \psi_2 + k_1 \sin \psi_1 = 0 \qquad (3\text{-}15)$$

This condition is somewhat loosely regarded as requiring that the reconstruction beam be the same as the reference beam. When the condition is not satisfied, aberrations result in the reconstruction (19).

The integral,

$$U = \int_0^1 J_0(v\rho) \exp(iu\rho^2) \rho d\rho \qquad (3\text{-}16)$$

can be solved using Lommel functions (20). The intensity obtained from UU* can be graphically represented by a set of isophote diagrams (20, 21) in which contour lines of constant intensity near the focal volume of the optical system are plotted on a (u, v) coordinate system. The importance of these isophote diagrams is that the Equation 3-16 solution they represent was obtained by analyzing the three-dimensional light distribution near the focus of a

diffraction-limited lens. In other words, the intensity field in the focused region of the reconstruction of the hologram of a point source is identical in form to the intensity field near the focal volume of a diffraction-limited lens. Therefore, the len's focusing characteristics apply equally well to the holographic reconstruction. In particular,

1. The fraction of the total energy which falls within a given area about the optical axis for any constant u is the same for point images obtained by lens projection and for holographic reconstruction.
2. The phase distribution near focus is the same in both cases, i.e., in the focal waist the phase surfaces propagate like plane waves.

The similarity of lens and holographic imaging shows that the phenomena of refractive and diffractive wave-shaping do not give intrinsic contributions to the image characteristics. Indeed, the derivation of the diffraction-limited lens system begins by assuming that the refractive phenomena has produced perfectly converging coherent spherical waves which are limited in extent only by an aperture. The subsequent results are all due to diffraction by the aperture. Experimental verification (22) has led to universal acceptance. When either the lens or the holographic system yields results that are less than potentially realizable, it is due respectively to being unable to

produce spherically converging wave fronts of sufficient accuracy and being unable to duplicate the reconstruction beam wave fronts from the reference beam wave fronts.

Two special solutions to Equation 3-16 are available. When the focusing condition, u = 0, is satisfied, the normalized intensity becomes

$$I_{u=0} = [J_1(v)/v]^2 \tag{3-17}$$

where $v = k_2 r_3 R/z_2$. This intensity distribution, which is the point spread function, is the familiar Airy pattern associated with diffraction from an aperture of radius R.

The second solution is for points along the optical axis, v = 0. Then,

$$I_{r_3=0} = \text{sinc}^2[(R^2/4)(k_2/z_2 - k_1/z_1)] \tag{3-18}$$

where sinc (x) = sin (x)/x. When

$$\lambda_1 z_1 = \lambda_2 z_2 + \lambda_2(\Delta z/2) \tag{3-19}$$

the focal depth can be considered to be the distance Δz from out-of-focus position $(-\Delta z/2)$ to $(+\Delta z/2)$ such that the intensity of these two positions does not drop below a given part p of the maximum at $\Delta z = 0$. By expanding the sine function and using the first two terms, it follows that

$$\Delta z = \lambda_2 (z_2/R)^2 (4/\pi)[6(1 - (1 - p)^{1/2})]^{1/2} \tag{3-20}$$

For the conventional case that the intensity is not allowed to drop lower than 80 percent of the maximum (20), then $p = .2$, and

$$\Delta z \simeq \lambda_2 (z_2/R)^2 \tag{3-21}$$

which is the familiar result that the focal depth increases with increasing f-number.

The derivation of the line spread function follows the same form as the point spread function derivation. Consider a line source of radiation along the y_1 axis. Invoking cylindrical symmetry, the paraxial expansions of s_1 and s_2 are

$$s_1 = z_1 + 1/2(x_1^2/z_1) \tag{3-22}$$

and

$$s_2 = z_2 + 1/2(x_2 - x_1)^2/z_2 \tag{3-23}$$

When the resulting hologram is exposed symmetrically about the z axis, the reconstruction integral limits are from (-L) to (L). There results

$$U_R = \sum_{j=1}^{3} K_j \int_{-1}^{+1} \exp(iu_j x_2^2) \exp(iv_j x_2)\, dx_2 \tag{3-24}$$

where

$$u_1 = -k_2 x_3 L/z_2 + (k_2 \sin \psi_2)\, L$$

$$u_2 = -k_2 x_3 L/z_2 + (k_2 \sin \psi_2 - k_1 \sin \psi_1)L$$

$$u_3 = -k_2 x_3 L/z_2 + (k_2 \sin \psi_2 + k_1 \sin \psi_1)L$$

$$v_1 = k_2 L^2/2z_2$$

$$v_2 = 1/2(k_2/z_2 + k_1/z_1)L^2$$

$$v_3 = 1/2(k_2/z_2 - k_1/z_1)L^2$$

$$K_1 = C[1 - (\gamma/2)(A/z_1 B_1)^2]$$

$$K_2 = -C(\gamma/2)(A/z_1 B_1) \exp(ik_1 z_1)$$

$$K_3 = -C(\gamma/2)(A/z_1 B_1) \exp(-ik_1 z_1)$$

$$C = -iKB_2 B_1^{-\gamma} L(z_2 \lambda_2)^{1/2} \exp(i\pi/4) \exp(ik_2 z_2) \exp(ik_2 x_3^2/2z_2) \quad (3\text{-}25)$$

L is one-half the width of the hologram in the x_2 direction.

Consider again only the real image. The integral can be solved only for $\psi_1 = -\psi_2$ and only when the focusing condition is satisfied. The resulting normalized intensity, which is the line spread function, is

$$I_{u=0} = \text{sinc}^2[k_2 x_3 L/z_2] \quad (3\text{-}26)$$

Comparing the widths of the line spread function to the point spread function, for equal arguments, at their respective first zero value shows

$$x_2 \cong 0.82\ r_2 \quad (3\text{-}27)$$

For holograms of the same size, the spread function for objects with circular symmetry is about 18 percent "broader" than for objects with linear symmetry. In other words, for equal parameters, linear objects will have a better

resolution than circular objects. This result is developed further in Chapter V.

The point and line spread function derivations can be used to characterize the distinctions between in-line and off-axis holography. These differ only in the geometry of combining the object beam with the reference beam. When $\psi_1 = 0$, the in-line geometry results. As can be seen from the v_j parameters (for both point and line derivations), the three images are superimposed. The image of interest (whether real or virtual) then rides on a field of "noise" produced by the other images. When $\psi_1 \neq 0$, the three images are not overlapping. Then the "noise" of the in-line images does not occur, and small contrasts in the image will be more readily detectable. It is interesting to note that when the reference and reconstruction beams satisfy the condition,

$$k_2 \sin \psi_2 + k_1 \sin \psi_1 = 0 \qquad (3\text{-}28)$$

the in-line and off-axis reconstruction integrals are the same. Thus, no inherent difference is expected in the resolution of the images between the two holograms. Should apparent resolution differences be observed, they will be due to character recognition problems caused by other superimposed images in the in-line reconstruction. The problem of character recognition as related to resolution criterion has multifarious difficulties for which no general solution

has yet been found (23). The experimental aspects of this problem are discussed in Chapter VI.

CHAPTER IV

CHARACTERISTICS OF IN-LINE AND OFF-AXIS RECONSTRUCTIONS

Since the availability of the laser, holography has become a major tool for optical measurement and testing. In-line and off-axis holography, which differ primarily in the geometry of the object and reference wave front combinations, each seems to have specific advantages in various applications. For example, the experimental simplicity of in-line holography seems well suited to recording object fields of small particles. This is particularly so because the in-line configuration minimizes beam coherence problems and the rigid stability requirements of the object-reference beam geometry. On the other hand, as the particle density becomes large, the ability of the in-line hologram to render readable reconstructed images decreases because a larger portion of the reference beam is blocked. Then off-axis holography, with its undisturbed reference beam, gives recognizably superior image reconstruction.

Clearly, a detailed comparison of in-line and off-axis reconstructed images is necessary. While the derivations of the point and line spread functions in Chapter III are of fundamental importance, they do not relate directly to any specific image. For these reasons, the in-line and off-axis reconstruction of a particle will be

derived. Also, because of the similarity between the diffraction due to a particle and a circular aperture of like cross-section (from Babinet's principle), the reconstruction of the aperture will also be derived.

The geometry of the holographic process is the same as in Figure 1, page 17. The particle of radius a (or the equivalent aperture) is centered on the origin in the (x_1, y_1) plane; it is illuminated by plane waves. Again the plane reference wave front is incident on the film plane (x_2, y_2) at an angle ψ_1, and the resulting hologram in the (x_2, y_2) plane is illuminated by a reconstruction plane wave front from an angle ψ_2.

The diffraction integral, Equation 2-10 with plane incident waves, becomes

$$U(x_2,y_2) = -\frac{iA}{\lambda_1} \iint_{\text{aperture}} \frac{e^{ik_1 s_1}}{s_1} dx_1 dy_1 \qquad (4\text{-}1)$$

To solve this diffraction integral requires that attention be given to diffracted light only in the paraxial region about the incident illumination's propagation direction. The diffraction from an aperture of radius a, using

$$s_1 = z_1 + 1/(2\ z_1)[r_1^2 + r_2^2 - 2r_1 r_2 \cos(\phi_2 - \phi_1)] \qquad (4\text{-}2)$$

becomes

$$U_{\text{aperture}}(r_2) = -iA(k_1/z_1)\exp(ik_1 z_1)\exp(ik_1 r_2^2/2z_1) \cdot \int_0^a J_0(k_1 r_2 r_1/z_1)\exp(ik_1 r_1^2/2z_1) r_1 dr_1 \qquad (4\text{-}3)$$

Two observations about Equation 4-3 are important. First, when a is sufficiently small that the argument of the exponential term is small compared with 2π, then Fraunhofer diffraction is obtained. This condition results when

$$4a^2/\lambda_1 << 8z_1 \tag{4-4}$$

The quantity $4a^2/\lambda_1$ is defined as the far-field. Experimentally, Equation 4-4 is satisfied when z_1 equals approximately one far-field. When Equation 4-4 is not satisfied, Equation 4-3 gives the Fresnel diffraction field of the aperture. Secondly, Equation 4-3 has an exact solution as a becomes infinite,

$$U(r_2)_{a\to\infty} = A \exp(ik_1 z_1) \tag{4-5}$$

This simply represents the incident plane wave front. This is consistent with expectation since $a \to \infty$ represents the condition that no diffracting aperture is in the (x_1, y_1) plane. However, in obtaining this result, the paraxial condition was strongly violated. It is nevertheless satisfying to observe the internal consistency in the mathematics of the Fresnel diffraction integral of Equation 4-3. Without it the form of Equation 4-3 would not be correct. An alternative form including higher order terms in the exponential argument would render the mathematics insoluble except in numerical form.

When the condition of Equation 4-4 applies, the exponential in Equation 4-3 is replaced by unity. The

diffraction integral reduces to the mathematically simpler Fraunhofer form of which the exact solution becomes

$$U_{aperture}(r_2) = -iA \exp(ik_1 z_1) \exp(ik_1 r_2^2/2z_1)(a/r_2) \cdot J_1(k_1 a r_2/z_1) \quad (4\text{-}6)$$

The shape of the resulting intensity curve is generally called the Airy pattern. When the condition of Equation 4-4 does not apply, Equation 4-3 has the form of Equation 3-16. Although the Fresnel paraxial conditions can be relaxed for the solution of Equation 4-3 as evidenced by Equation 4-5, the conditions are still part of Equation 4-6 in the form of the r_2^2 exponential term. This term becomes important in the holographic reconstruction equations.

The diffraction from an aperture is used to obtain the mathematical form of a hologram of a particle. Incident on a film plane in the (x_2, y_2) plane is the diffracted radiation from the particle and the reference beam incident from angle ψ_1. The diffraction from a particle has the form of Equation 4-3 but with different limits,

$$U_{particle}(r_2) = \int_a^\infty \cdots dr_1 = \int_0^\infty \cdots dr_1 - \int_0^a \cdots dr_1$$

$$= A \exp(ik_1 z_1)[1 + i \exp(ik_1 r_2^2/2z_1)(a/r_2) \cdot J_1(k_1 a r_2/z_1)] \quad (4\text{-}7)$$

The Fraunhofer diffraction expressed in Equation 4-6 has been assumed in obtaining this result. The intensity

recorded by the hologram is

$$I_1(r_2) = (U_{reference} + U_{particle})(U_{reference} + U_{particle})^* \tag{4-8}$$

Thus,

$$\begin{aligned} I_1(r_2) = B_1^2\{1 &+ (A/B_1)^2[1 + (a/r_2)^2 J_1(k_1 a r_2/z_1) - 2(a/r_2) \\ &\cdot J_1(k_1 a r_2/z_1) \sin(k_1 r_2^2/2z_1) \\ &+ (A/B_1)\cos(k_1 z_1 - k_1 x_2 \sin\psi_1) \\ &- 2(A/B_1)(a/r_2)J_1(k_1 a r_2/z_1)\sin(k_1 z_1 + k_1 r_2^2/2z_1 \\ &- k_1 x_2 \sin\psi_1)\} \end{aligned} \tag{4-9}$$

In its component parts Equation 4-9 represents

$$\begin{aligned} I_1(r_2) = \{&(\text{bias intensity}) \\ &+ (\text{the in-line hologram of the particle}) \\ &+ (\text{intensity modulation due to the reference beam}) \\ &- (\text{intensity of the interference between the} \\ &\quad \text{reference beam and the diffraction from an} \\ &\quad \text{aperture of radius a})\} \end{aligned} \tag{4-10}$$

From Equation 4-9, the in-line hologram of a particle is obtained for $B_1 = 0$,

$$\begin{aligned} I_2(r_2) = A^2[1 &+ (A/B_1)^2(a/r_2)\, J_1^2(k_1 a r_2/z_1) \\ &- 2(A/B_1)(a/r_2)\, J_1(k_1 a r_2/z_1)\sin(k_1 r_2^2/2z_1)] \end{aligned} \tag{4-11}$$

Similarly, by eliminating the proper modulating terms, the

hologram of an aperture is obtained from Equation 4-9,

$$I_3(r_2) = B_1^2[1 + (A/B_1)^2(a/r_2)^2 J_1^2(k_1 a r_2/z_1) - 2(A/B_1)(a/r_2) \cdot J_1(kar_2/z_1) \sin(k_1 z_1 + k_1 r_2^2/2z_1 - k_1 x_2 \sin \psi_1)] \quad (4\text{-}12)$$

In Equations 4-9, 4-11, and 4-12, the J_1^2 term is negligible everywhere except near the origin.

The holographic reconstruction phenomenon is modeled analytically by using Fresnel diffraction. The necessary assumptions for the validity of this approach are threefold: (1) that the geometrical conditions for Fresnel diffraction apply; (2) that the film emulsion thickness be sufficiently thin that Bragg reflection is not important; and (3) that the hologram be exposed in the linear region of the film's amplitude transmission. These assumptions are necessary for a viable mathematical description of the holographic reconstruction. Experimentally, the conditions can be considerably relaxed.

The reconstructed field is given by Equation 3-11. Using the same assumptions about film characteristics as in Chapter III and requiring that the off-axis condition, Equation 3-15, be satisfied, the real image reconstructed fields become

$$U_1(r_3) = i(\gamma/4)(A/B_1)B_2 B_1^{-\gamma} K(k_2/z_2)\exp[i(k_2 z_2 - k_1 z_1)] \cdot \exp(ik_2 r_3^2/2z_2)[F_1 - iF_2] \quad (4\text{-}13)$$

$$U_2(r_3) = -iB_2KA^{-\gamma}(k_2/z_2)\ \exp(ik_2z_2)\ \exp(ik_2r_3^2/2z_2) \cdot [F_1 + i(\gamma/4)F_2] \tag{4-14}$$

$$U_3(r_3) = i(\gamma/4)(A/B_1)B_2B_1^{-\gamma}K(k_2/z_2)\ \exp[i(k_2z_2 - k_1z_1)] \cdot \exp(ik_2r_3^2/2z_2)[F_2] \tag{4-15}$$

where

$$F_1 = \int_0^R \exp(ik_2r_2^2/2z_2)\ J_o(k_2r_3r_2/z_2)r_2dr_2 \tag{4-16}$$

$$F_2 = 2a \int_0^R \exp[(i/2)(k_2/z_2 - k_1/z_1)r_2^2]J_o(k_2r_3r_2/z_2) \cdot J_1(k_1ar_2/z_1)dr_2 \tag{4-17}$$

The differences in the reconstructions of an off-axis particle (subscript 1), an in-line particle (subscript 2), and an off-axis aperture (subscript 3) are due primarily to the combination of the integrals of Equations 4-16 and 4-17. Through these equations, the importance of the effective hologram size determined by R is also apparent.

Equations 4-16 and 4-17 have no general closed-form solution. Some of the reconstruction characteristics can be obtained from two special case solutions. Assuming the holograms are perfect so that $R \to \infty$, and that the focusing condition,

$$k_2/z_2 - k_1/z_1 = 0 \tag{4-18}$$

is satisfied, then

$$F_1 = i(z_2/k_2)\ \exp(-ik_2 r_3^2/2z_2) \qquad \text{for } r_3 > 0 \qquad (4\text{-}19)$$

and

$$\begin{aligned} F_2 &= 2(z_1/k_1) && \text{for } r_3 < a \\ &= 0 && \text{for } r_3 > a \\ &= z_2 a/k_2 r_3 && \text{for } r_3 = a \end{aligned} \qquad (4\text{-}20)$$

The intensities obtained with these results for $r_3 < a$ are

$$I_1 = C_1[1 + 1/4 - \cos(k_2 r_3^2/2z_2)] \qquad (4\text{-}21)$$

$$I_2 = C_2[1 + (\gamma/2)^2 + \gamma \cos(k_2 r_3^2/2z_2)] \qquad (4\text{-}22)$$

$$I_3 = C_1[1] \qquad (4\text{-}23)$$

where

$$\begin{aligned} C_1 &= [(\gamma/4)(A/B_1)B_2 B_1^{-\gamma} K]^2 \\ C_2 &= [B_2 K A^{-\gamma}]^2 \end{aligned} \qquad (4\text{-}24)$$

Using the definition of the far-field, the argument of the cosine function becomes $(\pi/4N)(r_3/a)^2$ where N is the number of far-fields. Since $(r_3/a) < 1$, the cosine function will oscillate only when $N < 1$, i.e., when the particle did not meet the Fraunhofer diffraction conditions during the recording process. Under the usual assumption that the Fraunhofer conditions are satisfied, N is larger than one. Then the cosine term does not oscillate within the image of

the particle. From the difference in the sign of the oscillatory terms in Equations 4-21 and 4-22, it is apparent that the image of an off-axis particle is darkened in intensity whereas the in-line image is increased in intensity over the background radiation. For the reconstruction of the aperture, the image is constant for $r_3 < a$ and zero for $r_3 > a$.

Figure 2 shows the intensity field reversal for in-line and off-axis reconstruction of opaque circles on a microscope reticle. The oscillation of intensities within larger images is shown in Figure 3. The image is an off-axis reconstruction with a focal distance of 114 mm of the United States Air Force Resolving Power Test Target 1951. The spatial frequencies of the bars are given in Table I; the ratio of width to length of the bars is one to five.

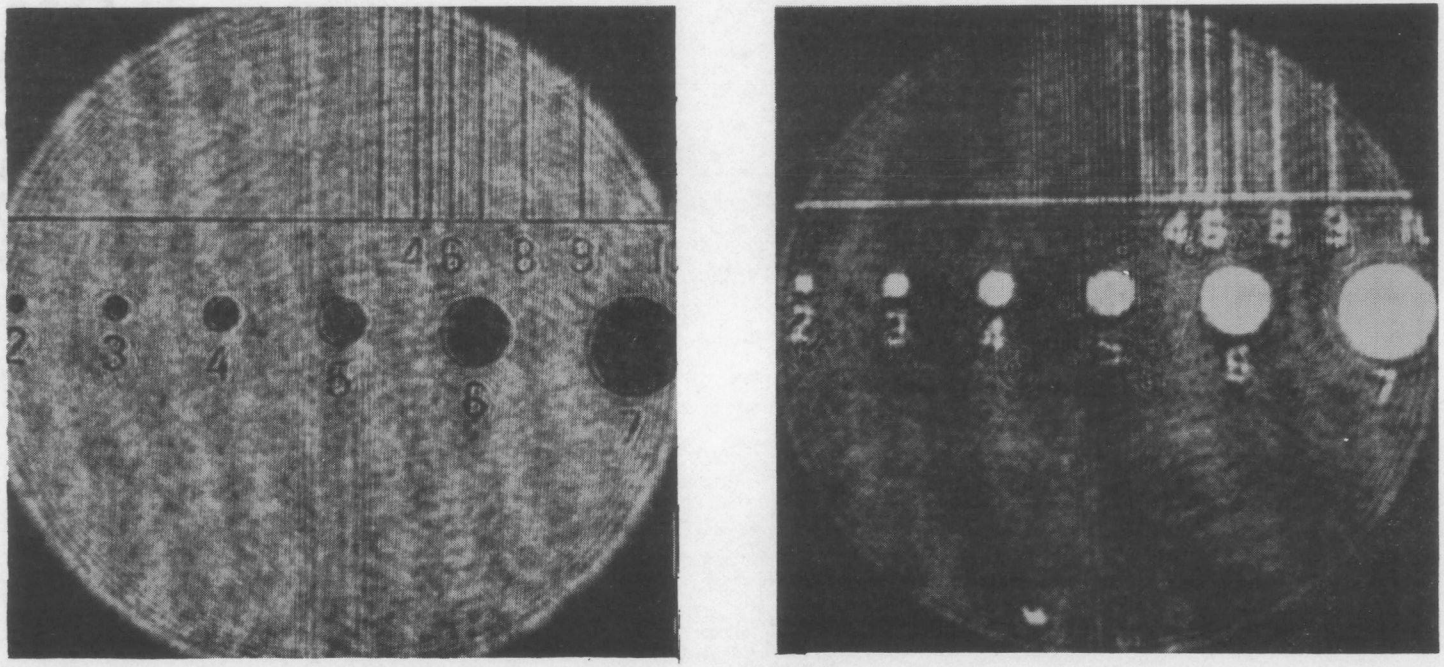
Some observations about the reconstructed intensities can be made when R is finite in Equations 4-16 and 4-17. For finite R, the reconstructed image intensity oscillates along the radius, the number of oscillations being a function of R. This can be derived by noting that each of the above intensities is proportional to F_2^2,

$$I_1 \propto F_1^2 + F_2^2 \tag{4-25}$$

$$I_2 \propto F_1^2 + (\gamma/4)F_2^2 \tag{4-26}$$

$$I_3 \propto F_2^2 \tag{4-27}$$

The maxima and minima (i.e., the extrema) of the intensities is found from



Off-axis Reconstruction

In-line Reconstruction

Figure 2. The reconstructed images of a microscope reticle which show the reversed field intensity of in-line and off-axis holography.

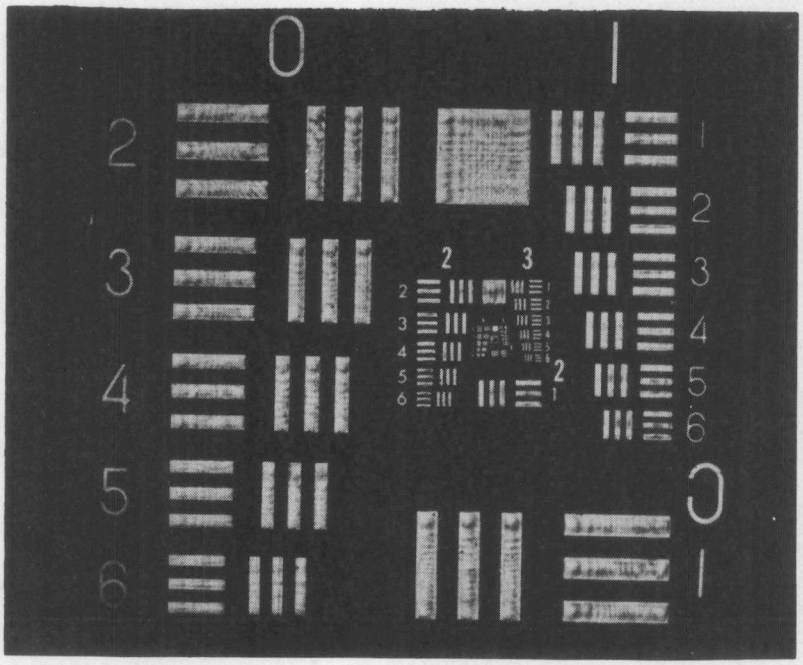


Figure 3. The off-axis reconstruction of a Resolving Power Test Target which shows the intensity detail within the image of large objects.

TABLE I

THE NUMBER OF LINES PER MILLIMETER IN THE UNITED STATES AIR FORCE RESOLVING POWER TEST TARGET 1951

Element Number	Group Number							
	0	1	2	3	4	5	6	7
1	1.00	2.00	4.00	8.00	16.00	32.0	64.0	128.0
2	1.12	2.24	4.49	8.98	17.95	36.0	71.8	144.0
3	1.26	2.52	5.04	10.10	20.16	40.3	80.6	161.0
4	1.41	2.83	5.66	11.30	22.62	45.3	90.5	181.0
5	1.59	3.17	6.35	12.70	25.39	50.8	102.0	203.0
6	1.78	3.56	7.13	14.30	28.51	57.0	114.0	228.0

$$\frac{dI}{dR} \sim 2F_2 \frac{dF_2}{dR} = 0 \qquad (4\text{-}28)$$

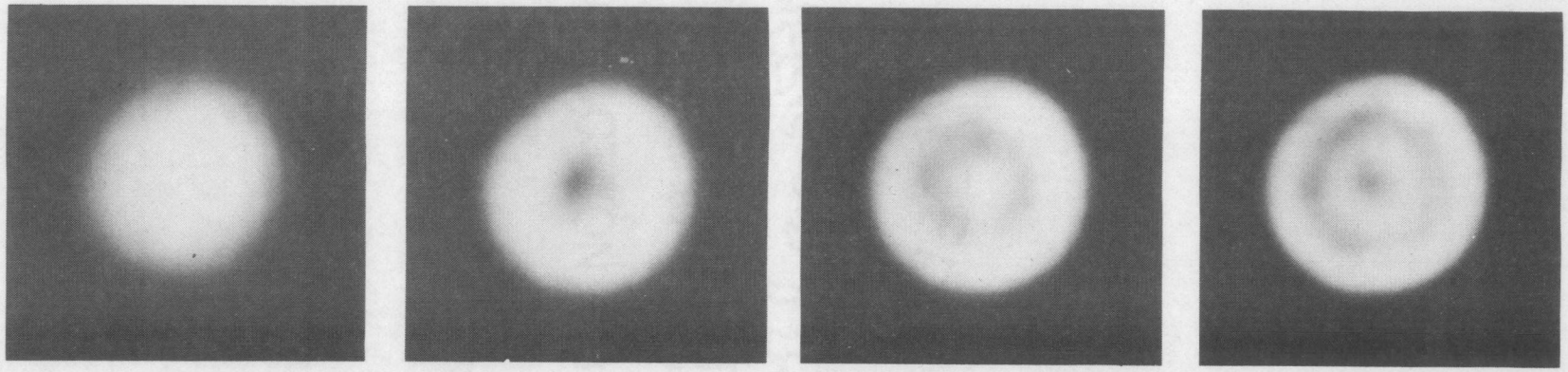
Now, for the in-focus case,

$$\frac{dF_2}{dR} = (2a/R)\ J_o(k_1 r_3 R/z_1)\ J_1(k_1 aR/z_1) \qquad (4\text{-}29)$$

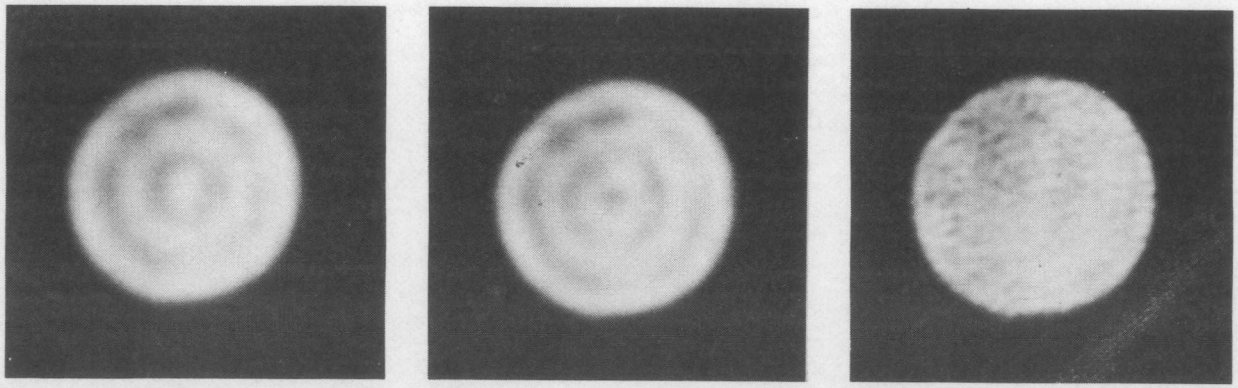
The r_3 locations of the extrema are determined by the J_o zeros. The number of extrema for any given set of k_1, a_1, z_1 and R is given by the number of J_o zeros that are contained within $0 \leq r_3 \leq a$. When the intensity has an extreme value at the center of the image, Equation 4-29 reduces to

$$\left.\frac{dF_2}{dR}\right|_{r_3=0} = (2a/R)\ J_1(k_1 aR/z_2) \qquad (4\text{-}30)$$

Equation 4-30 was verified experimentally. Figure 4 shows the in-focus (z_2 = 35 cm) reconstructions of a 165 μ diameter aperture for various values of R. The reconstruction beam diameter which determined R was chosen such that the image center intensity was an extremum. The number of extrema were found to satisfy the above conditions. The numerical results together with those of other sizes of apertures are given in Table II. It can be observed that the average J_1 arguments obtained correspond very satisfactorily with the J_1 zeros except for the single extremum case (the R = .222 cm reconstruction in Figure 4). The deviation results because of systematic observational errors;



R = 0.222 cm R = 0.311 cm R = 0.425 cm R = 0.565 cm



R = 0.705 cm R = 0.832 cm R = 2.0 cm

Figure 4. The in-focus, off-axis reconstructions of a circular aperture from holograms of different sizes.

TABLE II

EXPERIMENTAL DETERMINATION OF THE $J_1(kaR/z_2)$ ZEROS

2a	Number of Extrema in the Reconstruction					
(μ)	1	2	3	4	5	6
			Radius of the Hologram R (cm)			
44	.813	1.09	1.61			
65	.527	.826	1.14	1.39		
85	.413	.591	.870	1.19	1.45	1.73
98	.349	.508	.667	.978	1.16	1.45
165	.222	.311	.425	.565	.705	.832
204	.171	.235	.343	.470	.572	.660
			kaR/z_2 Where z_2 = 35 cm and λ = .6328 μ			
44	5.07	6.78	10.1			
65	4.86	7.61	10.5	12.8		
85	4.98	7.12	10.5	14.4	17.5	20.8
98	4.85	7.06	9.27	13.6	16.2	20.1
165	5.20	7.28	9.96	13.2	16.5	19.5
204	4.96	6.80	9.92	13.6	16.5	19.1
			Average ± Standard Deviation			
	4.99	7.11	10.04	13.53	16.66	19.88
	.13	.31	.46	.58	.56	.75
			J_1 Zeros			
	3.832	7.016	10.17	13.32	16.47	19.62

the judgemental determination of the single maximum condition cannot be made adequately. This has some bearing on the minimum size of recorded holographic information necessary in order to obtain a meaningfully recognizable reconstructed image. Experimentally, it is observed that the image simply becomes broader and more diffuse when the recording size is smaller than would result in the first zero of J_1. Then, little more information can be deduced than the fact that a particle was present during the recording. The focal plane of such a reconstruction is also difficult to determine. On the other hand, with R large enough that the second J_1 zero is recorded, a fairly accurate estimate of the particle size is available. Then the focal plane is defined by a well localized central intensity minimum.

The second special case solution of Equations 4-16 and 4-17 is found for points along the optical axis. Using $z_2 = z_1 - \Delta z$ and assuming for simplicity that the difference between λ_1 and λ_2 is small, then

$$k_2/z_2 - k_1/z_1 = k_1\Delta z/z_1(z_1 - \Delta z) \qquad (4\text{-}31)$$

For $R \to \infty$ and $r_3 = 0$,

$$F_1(r_3 = 0) = \int_0^\infty \exp(ik_2 r_2^2/2z_1) r_2 dr_2 = i(z_2/k_2) \qquad (4\text{-}32)$$

and

$$F_2(r_3 = 0) = 2a \int_0^\infty \exp(ik_1 r_3^2 \Delta z/2z_1(z_1 - \Delta z))\ J_1(k_1 a r_2/z_1) dr_2$$

$$= 2(z_1/k_1)[1 - \exp((-ik_1 a^2/2z_1)(z_1/\Delta z - 1))] \quad (4\text{-}33)$$

All the resulting intensities oscillate with the same functional dependence,

$$I_1(r_3 = 0) = C_1[1 + 4(1 + 3\Delta z/z_1)\ \sin^2((k_1 a^2/4z_1)(z_1/\Delta z - 1))] \quad (4\text{-}34)$$

$$I_2(r_3 = 0) = C_2[1 + (\gamma/2)(1 + \gamma/2 + (\Delta z/z_1) \cdot (1 + \gamma))\ \sin^2((k_1 a^2/4z_1)(z_1/\Delta z - 1))] \quad (4\text{-}35)$$

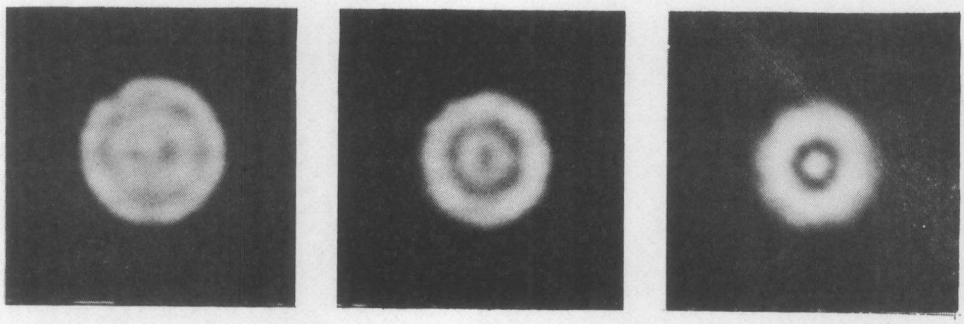
$$I_3(r_3 = 0) = C_1[8(1 + \Delta z/z_1)^2\ \sin^2((k_1 a^2/4z_1)(z_1/\Delta z - 1))] \quad (4\text{-}36)$$

The minima occur for

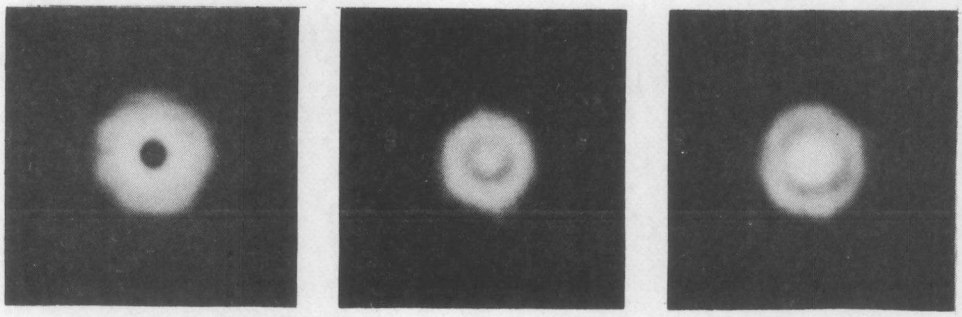
$$\Delta z = z_1 a^2/(2m\lambda z_1 + a^2) \quad (4\text{-}37)$$

where m is an integer. These extrema are easily located experimentally. Therefore, Equation 4-37 provides a method of obtaining particle size information.

Figure 5 shows the reconstruction of a 98 μ diameter aperture from a hologram where R = 2.5 cm and z_2 = 35 cm at various distances Δz away from the focal plane. The Δz = 1.9 mm reconstruction is particularly striking because there is total cancellation of the intensity at the center. Table III summarizes the use of Equation 4-37. The out-of-focus distances Δz of the minima in the reconstruction of



$\Delta z = 0$ $\Delta z = 0.64$ mm $\Delta z = 1.27$ mm



$\Delta z = 1.90$ mm $\Delta z = 2.54$ mm $\Delta z = 3.18$ mm

Figure 5. The out-of-focus, off-axis reconstructions of a circular aperture.

TABLE III

DETERMINATION OF THE APERTURE SIZE FROM THE OUT-OF-FOCUS DISTANCE

2a (μ)	m = 1	m = 2	m = 3	m = 4
	Out-of-Focus Distance Δz (mm)			
44	.394			
65	.800			
85	1.295			
98	1.867	.965		
165	5.398	2.705	1.854	
204	8.446	4.280	2.845	2.172
	Experimentally Determined Diameter (μ)			
44	44.6			
65	63.6			
85	80.8			
98	97.0	98.7		
165	164.0	164.9	167.4	
204	204.3	206.9	207.0	209.1

$\lambda_1 = .6328\ \mu$

$z_1 = 35.0$ cm

variously sized apertures were measured. It is apparent that the correspondingly calculated aperture sizes are quite accurate. The values of Δz were not observably a function of R.

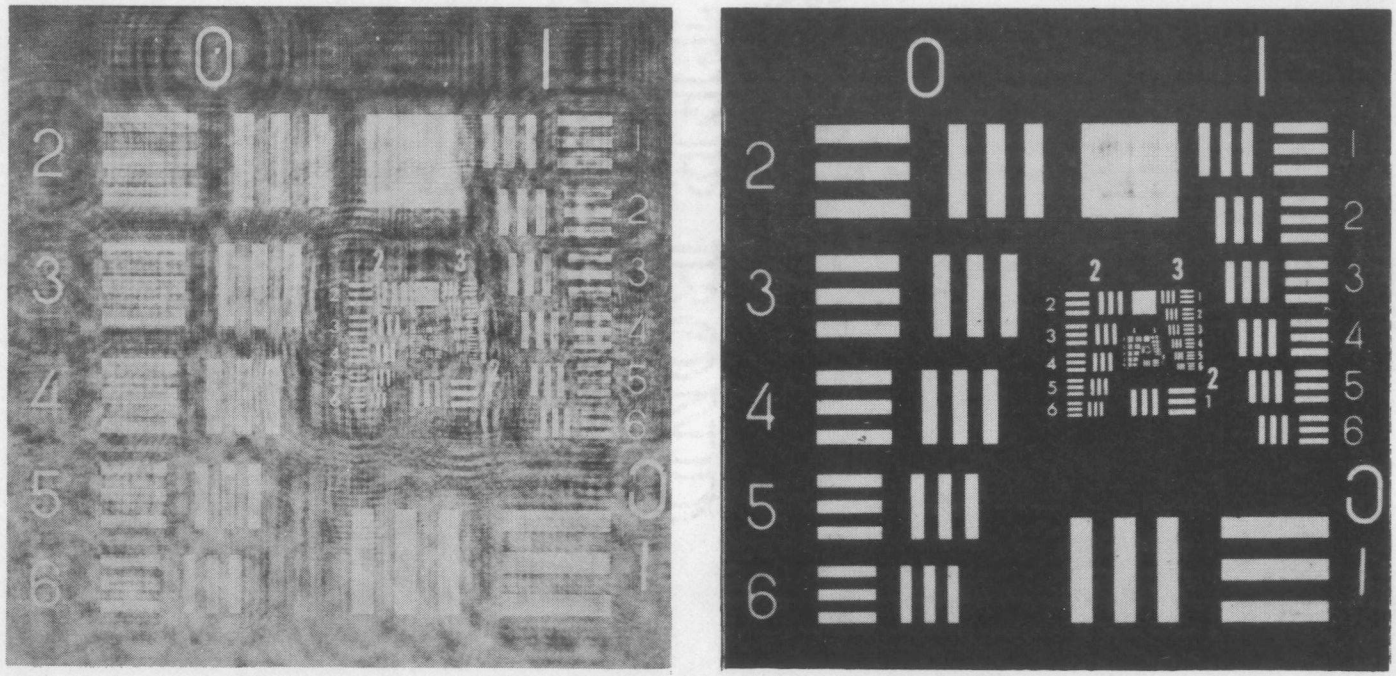
In summary, the derivations of this chapter have shown a basic similarity between the reconstruction of three different types of holograms: (1) the in-line hologram of a particle, (2) the off-axis hologram of a particle, and (3) the off-axis hologram of an aperture of the same cross-sectional area as the particle. With the exception of the intensity levels, the structural characteristics of the images are the same. In-line reconstruction of particle images is discussed in great detail by Belz (24). In particular, Belz presents isophote diagrams of the real image intensities of positions near the focal location (but not including the background radiation). The diagrams show the dependence of the image intensities on the effective size of the hologram R. With the derivations presented here, it becomes apparent that many of the results obtained by Belz can be generalized to off-axis holography.

It is necessary to mention in connection with in-line holography that the intensity derived for the particle reconstruction, Equation 4-22, did not include the out-of-focus virtual image. This image contribution is obtained by including in Equation 4-14 an integral of the form of Equation 4-17 except that the exponential argument becomes $i(r_2^2/2)(k_2/z_2 + k_1/z_1)$. When this is included in the

intensity derivation, the out-of-focus virtual image contributes negligibly to the intensity of the real image within $0 \leq r_3 \leq a$. For $r_3 > a$, the virtual image has the effect of superimposing on the background intensity an intensity pattern that is similar to a large diffraction pattern of the image. For single images this superimposed intensity has no serious detrimental effect on the image quality. However, when the holographic field is composed of many images, the virtual image "noise" has the psychovisual phenomenon of rendering the images less recognizable. The resolution of the images is not seriously affected (see Chapter VI), but the confusion of the virtual image noise makes the observational concentration necessary to ascertain the resolution a good deal more strenuous as compared to viewing the off-axis reconstruction of the same object scene. Figure 6 shows the in-line and off-axis reconstruction with focal distances $z_2 = 35$ cm of the Air Force Resolving Power Test Target 1951 to illustrate this point.

Two additional objects which were examined were the reconstructed images of the complementary aperture pair, a long wire and a long slit. Following the same derivational steps as above, the real reconstructed image can be characterized by,

$$U_{wire} \sim \int_0^1 (1/x)\, \sin[(kaL/z_1)x]\, \exp[-i(kL^2/2) \cdot (1/z_1 - 1/z_2)x^2]\, \exp[-i(kx_3L/z_2)x]dx \qquad (4\text{-}38)$$



In-line Reconstruction

Off-axis Reconstruction

Figure 6. A comparison of the in-line and off-axis reconstructions of the Resolving Power Test Target which illustrates the detrimental effects of the spatially unseparated images of the in-line reconstruction.

where L is one-half the hologram width, a is the wire cross-sectional radius and the other parameters are the same as before. The width of the image is a function of x_3. The length of the image is along the y_3 axis. Because of the (1/x) term in the integrand, the integral is more intractable than the reconstruction integral of the particle image. The only closed form solution available is obtained in the form of an infinite series for the in-focus real image. No special case out-of-focus solutions exist.

Using a differentiation approach similar to Equation 4-28, the maxima in the in-focus real image are found from the transcendental equation,

$$\cot(kaL/z_1)\ \tan(kx_3L/z_1) - x_3/a = 0 \qquad (4\text{-}39)$$

A graphic solution shows the number m of maxima within the image is dependent on the hologram size L by

$$kaL/z_1 = m\pi \qquad (4\text{-}40)$$

In other words, there is one maximum per π distance along kaL/z_1. Figure 7 demonstrates the extrema variation within the reconstructed image of a 236 μ wide slit. The reconstruction distance was 35 cm. The increase in L from one maximum to the next gave an average change in (kaL/z_1) of $3.05 \pm .20$. This value confirms the validity of Equation 4-40. Figure 7 is additionally useful for showing the functional dependence of the clarity of the reconstructed image on the film size. As L increases, the sharpness of

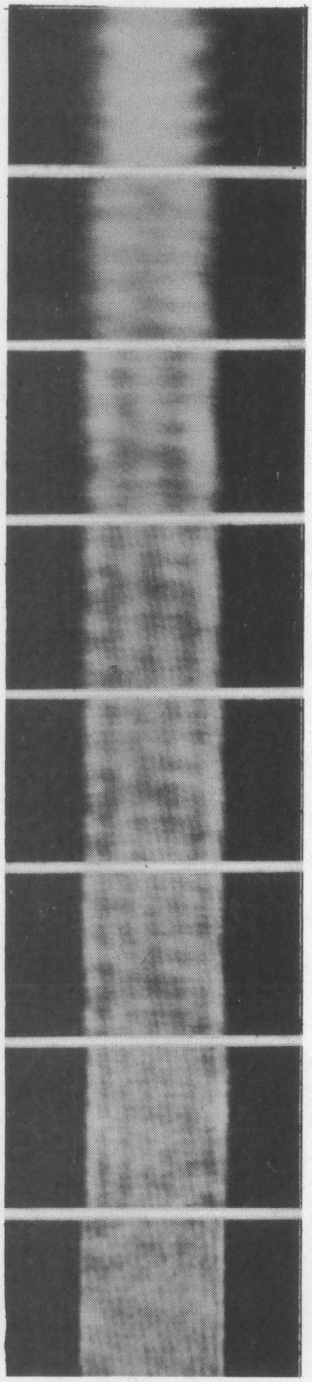


Figure 7. The in-focus, off-axis reconstructions of a slit aperture from holograms of different sizes.

the edge becomes more well defined, the intensity spread within the image becomes more uniform, and the image width increases to indicate accurately the object width. The edge details of the image are discussed at greater length in Chapter V.

CHAPTER V

RECONSTRUCTION DETAIL AND HOLOGRAM SIZE

The effective size of a hologram is that portion of the hologram which gives significant contribution to the intensity and definition of the reconstructed image. The term definition means that aspect of an image that is associated with the clarity by which detail is reconstructed. The necessity of obtaining the effective hologram size for determining reconstruction characteristics has been demonstrated in Chapters III and IV. The influence of hologram size on definition will now be discussed.

Unfortunately, definition has no satisfactory general meaning for describing optical phenomena (23). Rather, definition is usually related to some specific experimental configuration by which images of specific objects are recognized. The foreknowledge of the image that is to be recognized has a great effect on the "intelligibility" of the perceived image. In the case of holographic reconstruction of a particle image, the two most important characteristics that determine "good definition" are found by experience to be the intensity and the sharpness--i.e., the slope near the edge of the image. These quantities can be analytically obtained.

The intensity distribution of a reconstructed particle image is shown in Figure 8 which has been adapted

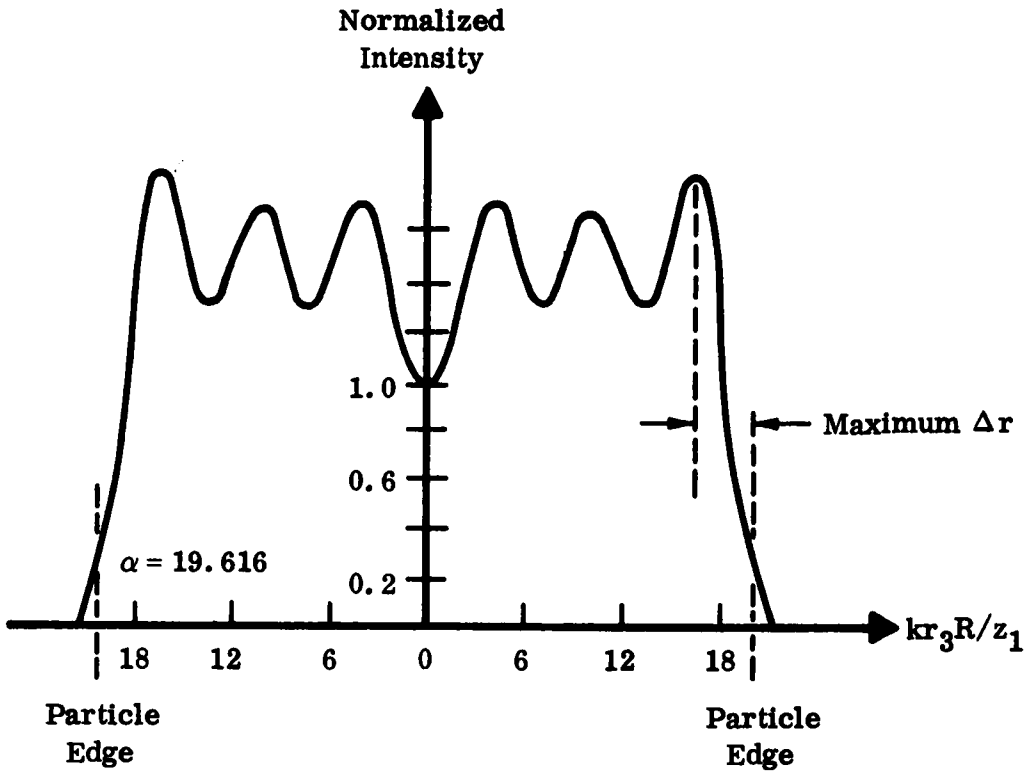


Figure 8. Theoretical normalized intensity of the in-focus reconstructed image of a particle (25).

from Belz (25). The hologram dimension R, particle size a, and recording distance z were such that kaR/z = 19.616 which is the value of the sixth zero of J_1. Consider the intensity in the non-oscillatory region near the edge of the image. Observed visually, this region is perceived as the edge of the image; the actual edge of the object occurs (25) where I = .25. The variation of the intensity at an arbitrary point a small distance Δr inward from the edge and the variation of the slope at that point as a function of R can be obtained from Equation 4-17. For simplicity, the background field F_1 and the out-of-focus contribution to the image are neglected. For the in-focus condition, $k_1/z_1 = k_2/z_2 = k/z$, the amplitude field at $r_3 = a - \Delta r$ is

$$F_2 = 2a \int_0^R J_0(k(a - \Delta r)r_2/z)\, J_1(kar_2/z)\, dr_2 \qquad (5\text{-}1)$$

Using the integral definition of the Bessel function, it may be shown that

$$J_0(x - x') = J_0(x) + x'J_1(x) \qquad (5\text{-}2)$$

when x' is small enough so that the approximations

$$\cos(x' \sin \theta) \simeq 1 \qquad (5\text{-}3)$$

and

$$\sin(x' \sin \theta) \simeq x' \sin \theta \qquad (5\text{-}4)$$

are sufficiently accurate for arbitrary θ. Applying such a restriction to $kR\Delta r/z$, then

$$F_2 = (z/k)[(1 - J_o^2(\alpha)) + \alpha^2(\Delta r/a)(J_o^2(\alpha) + J_1^2(\alpha) - (2/\alpha)\, J_o(\alpha)\, J_1(\alpha))] \qquad (5\text{-}5)$$

where $\alpha = kaR/z$. As before, the intensity is proportional to F_2^2. The point of observation is in the non-oscillatory region of the reconstructed image because of the condition that Δr be small. The slope of the intensity at the point of observation is

$$\left| a \frac{dI(\Delta r)}{dr_3} \right| = (z/k)2\alpha^2 F_2 [J_o^2(\alpha) + J_1^2(\alpha) - (2/\alpha)J_o(\alpha)J_1(\alpha)] \qquad (5\text{-}6)$$

The intensity curves generated by Equations 5-5 and 5-6 are shown in Figures 9 and 10 respectively. The curves have been normalized to exclude the (z/k) term; $(\Delta r/a)$ was chosen arbitrarily to be 0.05. The increase in intensity with increasing α shows that, as the hologram size increases, the signal strength in the region of the image edge increases. Similarly, the increasing slope shows that the edge becomes more well defined, that is to say, more like a step function.

The intensity versus α curve of Figure 9 should not be interpreted to indicate that the general intensity level across the particle image increases with increasing α. Consider the intensity at the center of the image,

$$I_{r_3=0} = (z/k)^2[1 - J_o^2(\alpha)]^2 \qquad (5\text{-}7)$$

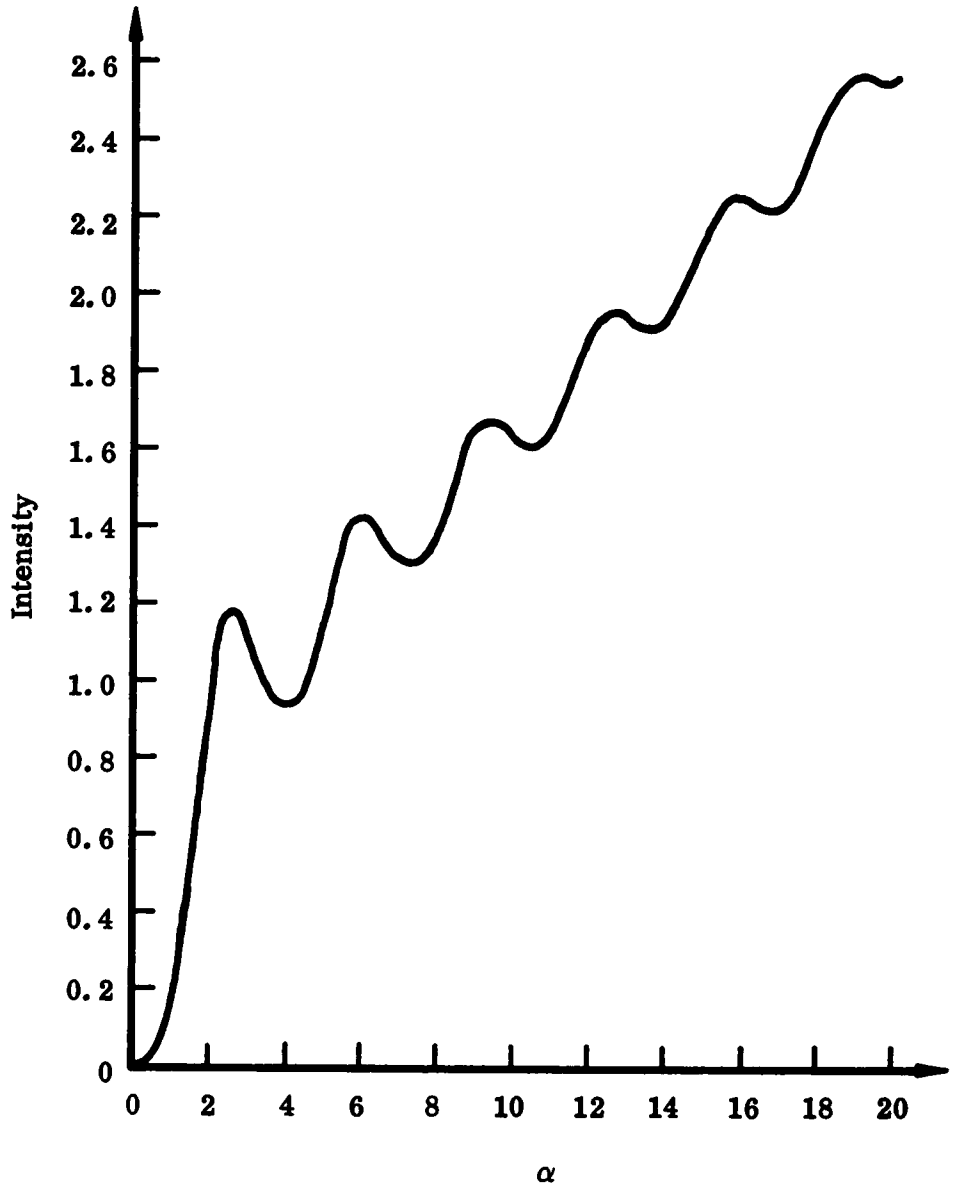


Figure 9. The intensity near the edge of the reconstructed image of a particle.

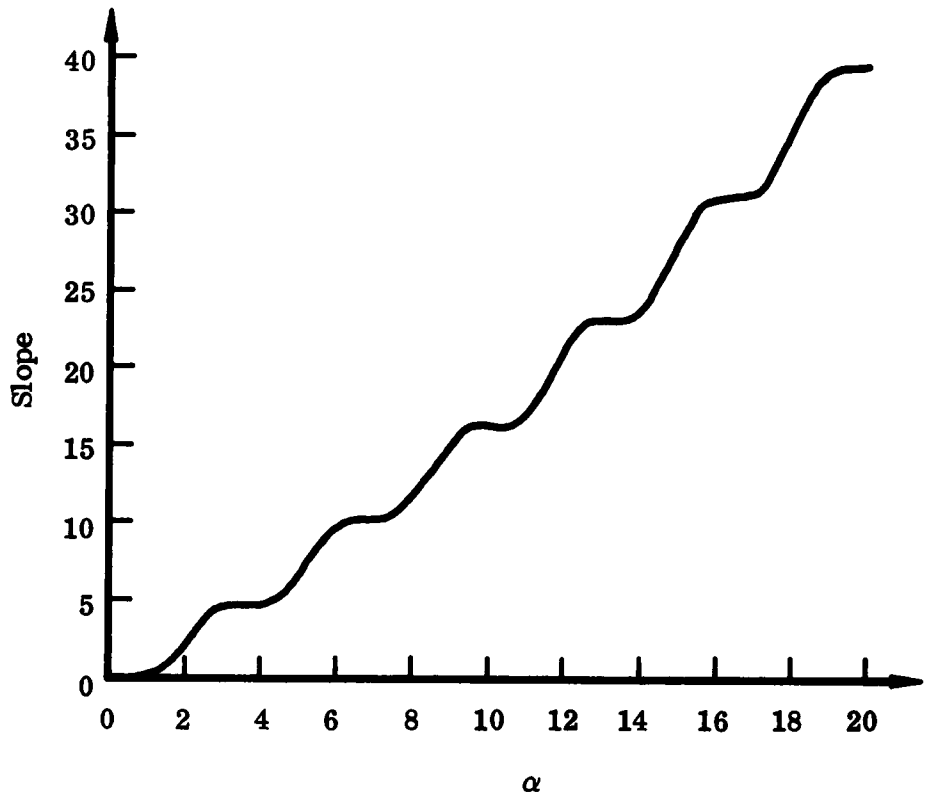


Figure 10. The slope of the intensity near the edge of a reconstructed image of a particle.

which is shown in Figure 11. Again the curve has been normalized to exclude the (z/k) term. The intensity oscillates with increasing α, but the average level of intensity is approximately constant.

Thus, the net effect of increasing hologram size on the reconstructed image of a particle is that an increasing intensity will be found near the image edge. Consequently, the edge becomes more like a step function. These effects can be easily observed experimentally.

The problem of finding the effective hologram size R for evaluating the reconstruction integrals, Equations 4-16 and 4-17, can be illustrated by considering the intensity function which the holographic film must record. Consider superimposing on a film plane two plane waves of intensities A^2 and B^2 such that the angle between them is ψ. The film plane is perpendicular to the angle bisector. Arbitrarily, A^2 is the intensity of the reference wave, and B^2 is the intensity of the signal (object) wave. The intensity in the film plane is

$$I = A^2 + B^2 + 2AB \cos(kx \sin(\psi/2))$$
$$= A^2[1 + \beta + 2\sqrt{\beta} \cos(kx \sin(\psi/2))] \qquad (5\text{-}8)$$

where $\beta = (B/A)^2$ is the object-to-reference beam intensity ratio, and x is the direction in the film plane which is along the plane containing beams A and B.

The normalized bias term $(1 + \beta)$ is modulated by the "information" term $2\sqrt{\beta} \cos(kx \sin(\psi/2))$. The information

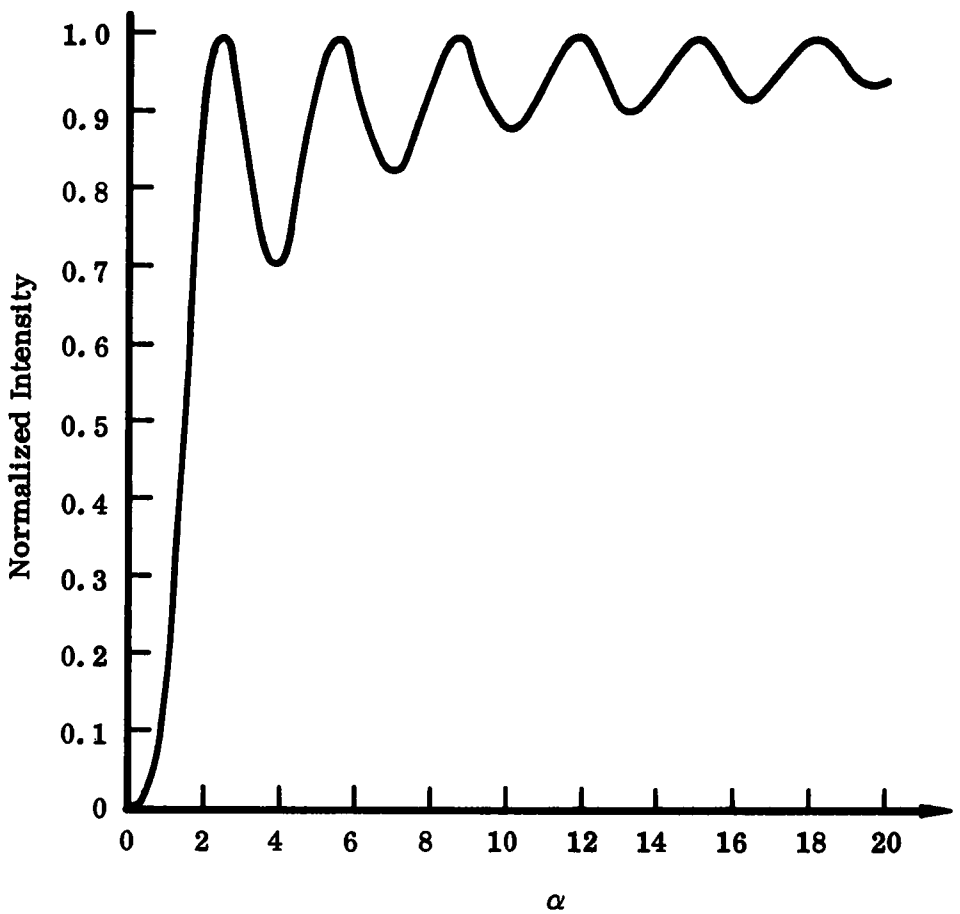


Figure 11. The normalized intensity at the center of the reconstructed image of a particle.

term, when considered from the viewpoint that it describes a grating, causes the diffractive process of the reconstruction. The intensity in the reconstruction is a function of the modulation amplitude. The information term resulting from an object, as exemplified by a particle, has the property that its spatial frequency increases and the amplitude decreases with increasing x. These important effects will be discussed further.

The effectiveness of the film in contributing to the definition of the reconstructed image is dependent on: (1) the ability of the film to record and resolve the amplitude and spatial frequency variation of the bias intensity modulated by the information intensity and (2) the smallest ratio β which gives significant intensity contribution to the reconstructed image. The hologram aperture is defined as that dimension beyond which either item 1 or 2 are not adequately satisfied.

Experimental demonstration will now be presented. All the results of this chapter were obtained using a He-Ne laser (λ = .6328μ) and Agfa Gevaert 10E75 film plates developed in Kodak D-19 developer.

Holograms were made of the two plane wave configuration discussed above. The observable exposure on the film plates was examined according to the parameters of Equation 5-8. For small values of ψ where the spatial frequency of the fringes was only about 2 lines per millimeter, fringes could be observed for β less than 10^{-5}. This figure was the

measurable intensity resolution available from the combined effects of the laser power, the optical components, and a Tektronics J16 photometer. It may be emphasized that this result is the principle reason for the successful experimental reality of holography. Because of the coherent nature of the light necessary for holography, the film is not required to resolve intensity ratios of β but rather--from Equation 5-8--of $2\sqrt{\beta}$. For the case at hand, this value was then only about 1/160. By visual estimation, Agfa Gevaert 10E75 was judged to have an intensity resolution of about 1/250 which corresponds to $\beta = 4 \times 10^{-6}$. These examples point out how weak the diffracted intensity field may be relative to the bias intensity field (the reference wave) and still yield a successful hologram. For example, a hologram was made of a 5 μ diameter aperture at a distance of 35 cm. The reconstruction gave adequate definition for unambiguous recognition of the image. The reference beam and the beam incident on the aperture were equally intense. From Equation 4-12

$$(a/r_2)^2 \, J_1^2(k_1 a r_2/z_1) \approx 4 \times 10^{-6} \qquad (5\text{-}9)$$

Using the asymptotic approximation of the Bessel function and replacing the cosine term within the Bessel function approximation by unity, then

$$r_2^3 \approx ((a\lambda z)/(4\pi^2)) \times 10^6 \qquad (5\text{-}10)$$

This gives $\max(r_2) \approx 2.4$ mm, which is by this criterion the effective maximum hologram radius R.

The upper limits of the spatial frequency of the intensity variations that the film can resolve were approximated by recording a "white light hologram" (14). In this hologram the object and reference beams are incident on the film plane from opposite directions. The interference fringes become exposed planes within the depth of the emulsion in a manner very similar to the recording of a Lippmann color photograph. The minimum spacing between the planes is $\lambda/2$. The fact that reconstruction was observed unambiguously indicates that 10E75 could resolve lines separated by $\lambda/2$. This corresponds to a spatial frequency of over 3000 lines/mm. Referring again to Equation 4-12, the spatial frequency of the hologram of an aperture is determined by the $\sin(kr_2^2/2z_1)$ term. Determining the r_2 value such that the fringe separation is only $\lambda/2$ gives $\max(r_2) = z_1$. Using the same parameter values as above, then effectively $R = z_1 = 35$ cm, a value which rather greatly exceeds normal hologram sizes. Conversely, the film spatial resolution is not exceeded when z_1 is less than the radial size of the hologram. Comparing this result with the intensity resolution of the film shows clearly that spatial resolution is not the limiting factor in the holographic recording process.

The results relating to the intensity resolution of the film imply that beyond a certain areal size of the film

no holographic interference fringes will be recorded. This, however, does not indicate whether all of the effective area of the hologram contributes significantly to the reconstructed image. For example, from Equation 5-9 it is apparent that as r_2 increases, the intensity diffracted to the points on the circle of r_2 radius decreases, i.e., that the object-to-reference beam intensity ratio decreases with increasing r_2.

To isolate individual parameters of the reconstruction process, the effects of the object-to-reference intensity beam ratio β, the film density, and the angle ψ between the object and reference beams were studied. For this purpose, holograms of two intersecting plane wave beams were made. The geometry was such as implied by Equation 5-8. Holograms were made with β = 1, 1/3, 1/10, 1/30, 1/100, and 1/300 for ψ = 3° and 90°. The holograms were exposed to give densities between 0.1 and about 2.0. In the reconstruction, the intensities in the undiffracted transmitted beam and the first-order diffraction images were measured, and the diffraction efficiencies were calculated. Diffracted efficiency is defined as the average first-order image intensity divided by the average incident beam intensity. When applying this definition, the effects of the absorption of the glass plate and the emulsion are specifically excluded. Also excluded, as part of the definition, are the higher order diffraction modes which result from the film nonlinearity.

Figure 12 shows the diffraction efficiency versus film density curves for the different values of β when the angle separating the two beams was 3 degrees. Since this angle is relatively small, it is reasonable to assume that a similar set of curves applies also for diffraction from in-line holograms. The greatest intensity in the reconstruction was obtained for $\beta = 1$. The maximum diffraction efficiency for $\beta = 1$ and $\beta = 1/3$ was obtained for a film density between approximately 0.5 and 0.6; for the smaller values of β, the optimum density shifted to about 0.7. For 10E75 the 0.5 density was obtained with an incident energy = (intensity) $\times$ (time) = 40(mw/m^2) (sec) when $\beta = 1$. From the roughly equal spacing between the curves, it may be observed that the log of the diffraction efficiency is a somewhat linear function of log β.

The fact that $\beta = 1$ gave the greatest diffraction efficiency is an immediate indication that the optimum hologram image reconstructions are obtained for a nonlinear holographic recording (14). The energy fluctuation incident on the film varies between zero and 4A (times exposure time), and the resulting densities are necessarily partially outside the linear region of the Hurter-Driffield curves. The amplitude transmittance curves of Figure 13, which were obtained from the data of Figure 12, reflect this nonlinearity. It may be seen that the curves for the different values of β have different slopes although the linear portion of the curves is approximately located in the same

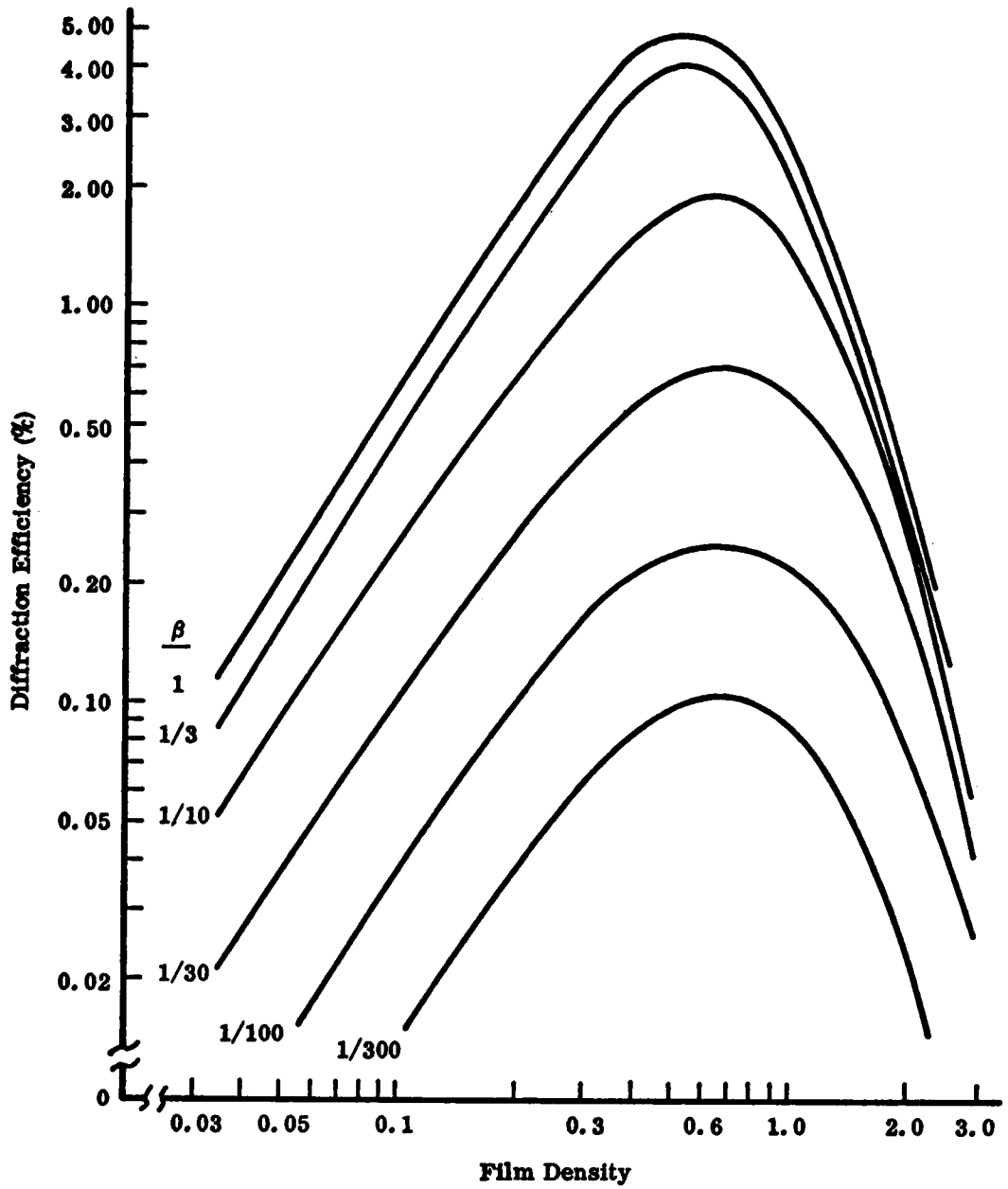


Figure 12. Diffraction efficiency as a function of exposed film density and the object-to-reference beam intensity ratio β obtained from holograms for which $\psi = 3°$.

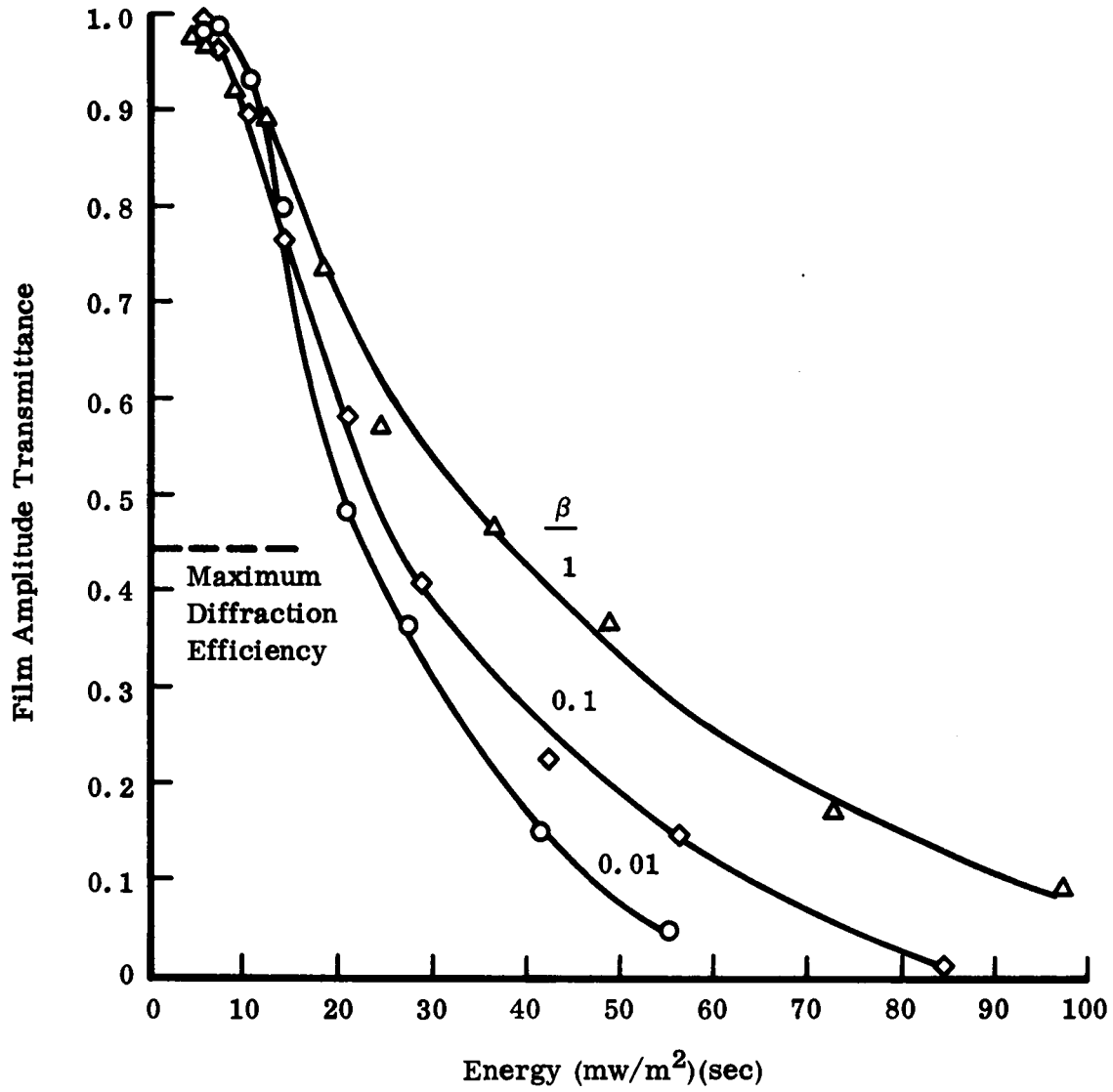


Figure 13. The amplitude transmission of the holograms as a function of the exposing energy and the object-to-reference beam intensity ratio β.

region. All the curves are based on the same bias intensity. If the film density response were linear, the intensity transmittance would be the average of the combination of the bias density modulated by the cosine function (Equation 5-8). The transmittance would then be proportional to the bias density, and all the curves of Figure 13 would overlap.

It may be emphasized that film nonlinearities in the off-axis holographic process do not cause serious deficiencies in the reconstruction. As already mentioned in Chapter III, the nonlinearities do not affect the phase information of the first-order diffraction. With phase information being preserved, the amplitude distortion will have relatively minor effects on the reconstruction. The higher order diffraction modes generated are nondetrimental since they physically propagate in different directions from the first-order image.

The curves of Figure 14 show the diffraction efficiencies as a function of β for different film densities. It is apparent that the curves tend to cluster for the film densities between approximately 0.4 and 1.2. Therefore, the decrease in diffraction efficiency with decreasing β is rather independent of the film density. The diffraction efficiency can be considered to be a function primarily of the object-to-reference beam intensity ratio. This important effect will be discussed further below.

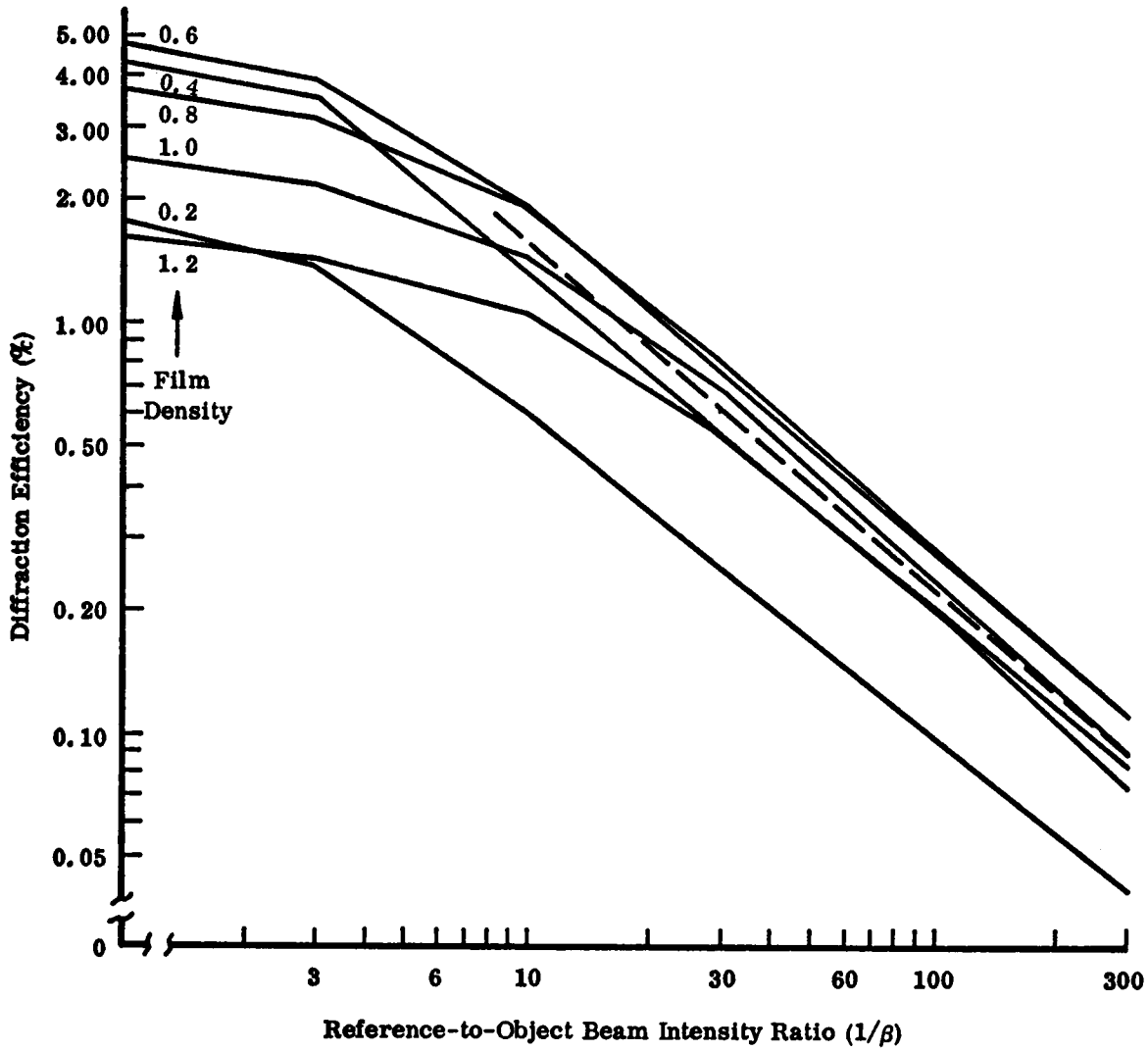


Figure 14. The diffraction efficiency versus β curves obtained from the $\psi = 3°$ holograms, which show the relative independence of the diffraction efficiency from the exposed film density.

To this point, the analysis has been based on the assumption that the film emulsion could be considered to have a negligible thickness. For such "thin" holograms, the reconstruction is entirely a diffraction process. As the angle ψ increases, the spatial frequency which the film must record also increases. The spacing between fringes is given by $d = \lambda/(2 \sin(\psi/2))$ which shows that the fringe spacing is comparable in size to the light wavelength for even relatively small ψ. Since emulsion thicknesses are generally greater than 5μ--10E75 being typically 7μ-- the holograms, to some degree, must be considered thick when ψ is increased. The lines of constructive interference of the intersecting object and reference beams then trace out exposure planes throughout the depth of the emulsion. The reconstruction process becomes a problem of diffraction from a thick grating.

Smith (26) has presented the first-order solution of the diffraction from a three-dimensional sinusoidal grating. A somewhat special case will be used for the purpose of this discussion. Assume that the angle bisector of the incident plane object and reference beams is perpendicular to the film emulsion plane. Then the exposed fringes construct planes within the emulsion which are perpendicular to the plane defined by the two beams and parallel to the bisector of the beams. Figure 15 defines the parameters to be used. The angles refer to the wave propagation within the emulsion. The reconstruction field varies as

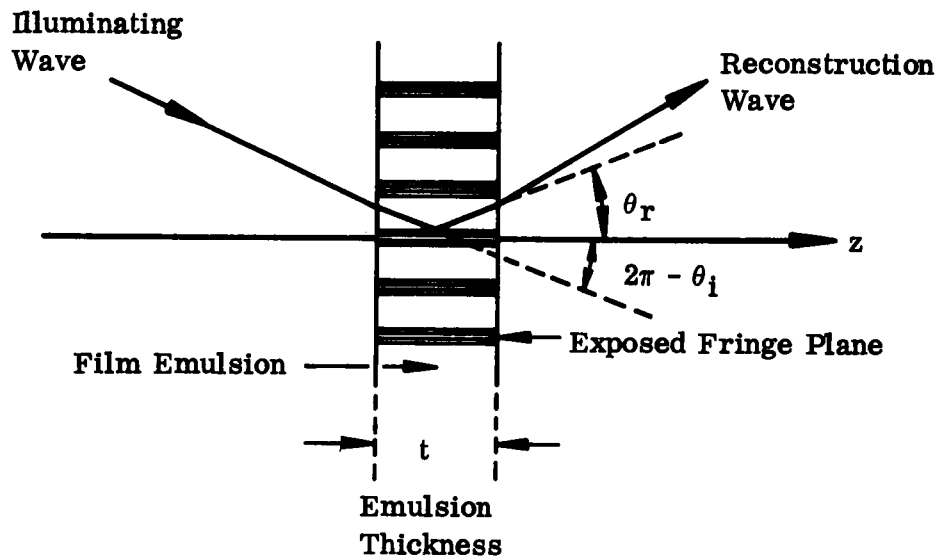


Figure 15. Reconstruction from a thick hologram.

$$U(\theta_r) \approx KM\{t\ \mathrm{sinc}[k_2(\cos\theta_r - \cos\theta_i)(t/2)]\} \cdot \{H\ \mathrm{sinc}[k_2(\sin\theta_r - \sin\theta_i) \pm 2k_1\sin(\psi/2)(H/2)]\} \quad (5\text{-}11)$$

where $M = 2\sqrt{\beta}/(1+\beta)$, K is a constant, t is the thickness of the emulsion, and 2H is the length of the recording medium. The plus/minus sign indicates the real and virtual images. It is a characteristic of thick holograms that there exists no θ_i which allows Equation 5-11 to have a maximum value simultaneously for both the plus and minus sign case. As this implies, only one of the images is observed experimentally. The maximum diffracted field exists when both

$$\cos\theta_r = \cos\theta_i \quad (5\text{-}12)$$

and

$$\sin\theta_r - \sin\theta_i = (k_1/k_2)\ 2\ \sin(\psi/2) = \lambda_2/d \quad (5\text{-}13)$$

are satisfied. These conditions represent respectively the law of reflection and the grating equation.

Additionally, it may be seen from Equation 5-11 that the intensity of the reconstructed image is dependent on M squared, which for small β is proportional to β. However, it is not apparent that the intensity of the reconstructed image decreases with increasing ψ, as reported by others (27, 28). The importance of this effect on the definition of the reconstructed images is subjective; if the intensity

in the reconstructed image is so low that details of the image are below the sensitivity threshold of the detector, the image definition is effectively degraded. The implication is that no impairment of the physical process of reconstruction causes the loss of image detail, but rather that the detail simply cannot be detected adequately. For this reason, a small ψ with its attendant higher diffraction efficiency is desirable.

Figure 16 shows the diffraction efficiencies for conditions similar to Figure 12, page 67, except that the angle between the object and reference beams was 90 degrees. Compared with the curves of Figure 12, the maximum diffraction efficiencies again occur for exposure densities of about 0.7. However, the diffraction efficiencies are not as large, and with smaller β they decrease more rapidly. Like Figure 14, page 70, Figure 17 shows the independence from film density of the β dependent diffraction efficiencies. The diffraction efficiency decreases more rapidly with smaller β than for the $\psi = 3°$ case.

The correlation of the variable β with diffraction information from an object is made by noting that the amplitude of the oscillating terms of the diffracted intensity decreases proportionally with the distance from the center of the diffraction pattern. If, for example, the diffraction rings resulting from a particle are considered according to radial zones, the outer zones will have an effective lower β than the inner zones. Consequently, the

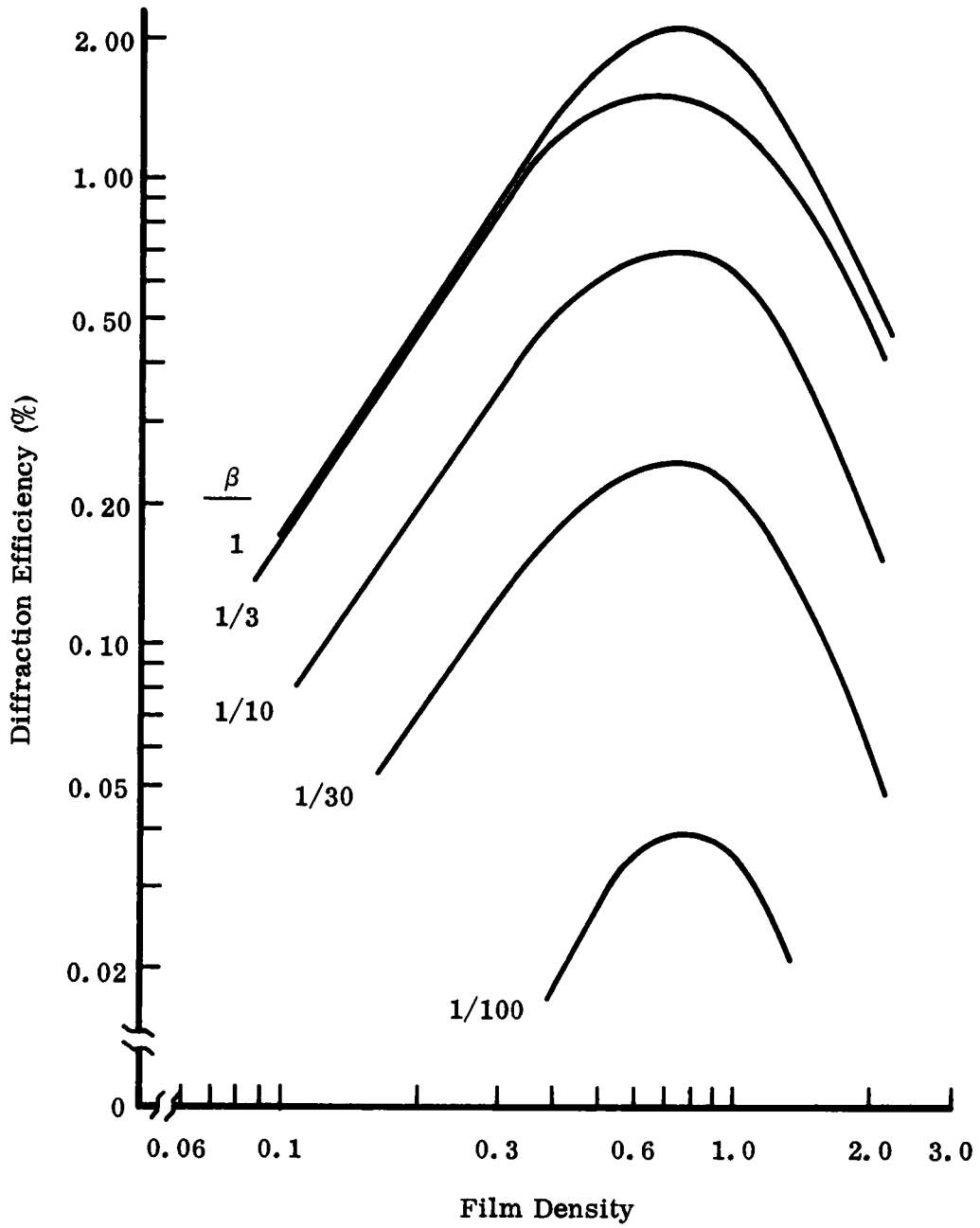


Figure 16. Diffraction efficiency as a function of exposed film density and the object-to-reference beam intensity ratio β obtained from holograms for which $\psi = 90°$.

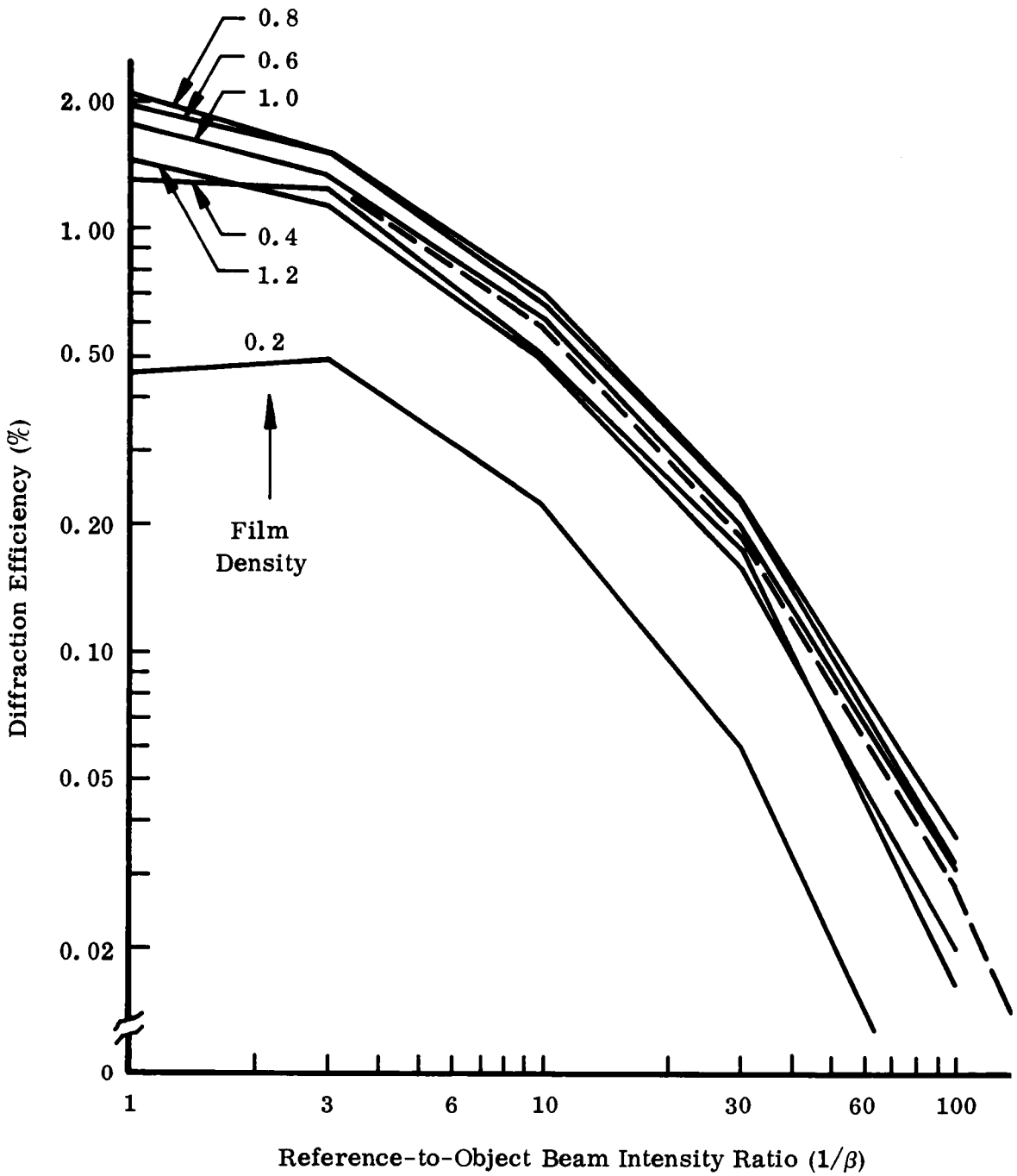


Figure 17. The diffraction efficiency versus β curves obtained from the $\psi = 90°$ holograms, which show the relative independence of the diffraction efficiency from the exposed film density.

outer zones can be expected to contribute less to the reconstructed intensity than the inner zones.

The most significant influence on the areal size of the hologram comes from the interpretation of the slopes of the curves in Figure 14, page 70, and Figure 17. The dashed lines represent the average of the slopes of the curves within the cluster. Since the $\psi = 90°$ case slope is steeper than the $\psi = 3°$ case, the difference in β necessary for the diffraction efficiency to decrease to some particular value is not as large for the $\psi = 90°$ case as for the $\psi = 3°$ case. For example, for the diffraction efficiency to decrease to one percent of the maximum available requires $\beta \approx 1/100$ when $\psi = 90°$ and $\beta \approx 1/600$ when $\psi = 3°$. Thus, in the reconstruction, a $\psi = 3°$ hologram will have a greater number of outer diffraction zones contributing effectively to the image intensity and clarity than will a $\psi = 90°$ hologram. For this reason, the small angle ψ holograms can be expected to give better reconstruction results than the larger angle holograms.

It remains to be shown that these expectations regarding β are valid. The approach is to calculate the dependence of the energy within the reconstructed image as a function of the size of the hologram. When considering the reconstruction of an aperture (which, by the results of Chapter IV, implies similar results for the reconstruction of a particle) the energy within the in-focus real image is given by

$$E = \int_{0}^{2\pi} \int_{w}^{w'} \{F_2\}^2 \, r_3 dr_3 d\theta \qquad (5\text{-}14)$$

where F_2 is obtained from Equation 4-17. The limits of integration are from some point w on the radius of the aperture image to a point w' further out on the radius. By this notation, integration over the entire image is represented by the limits w = 0 and w' = a. The normalized energy within the image from w to the edge is

$$E_N = (2\alpha^2/(1 - w/a)) \int_{w/a}^{1} \{\int_{0}^{1} J_0(\alpha\varepsilon\rho) \, J_1(\alpha\rho) d\rho\}^2 \, \varepsilon d\varepsilon \qquad (5\text{-}15)$$

where $\alpha = k_1 aR/z_1$. This integral was numerically calculated for w/a = 0 and .9 with the results shown in Figure 18. The energy within the entire image (w/a = 0) increases, as expected, with increasing α. Local maxima are obtained for values of α which are the values of the zeros of $J_1(\alpha)$. The energy within the first, second, and third zeros are respectively 78.2%, 87.3%, and 91.0% of the maximum theoretical energy. It may be observed that increasing the hologram size beyond the value represented by $\alpha = 7.016$, which is the second J_1 zero, contributes relatively little extra to the overall intensity within the image. By doubling the radius of the film size ($\alpha \approx 14$), the energy is increased by only 5.7%, a change to which the eye is not particularly sensitive.

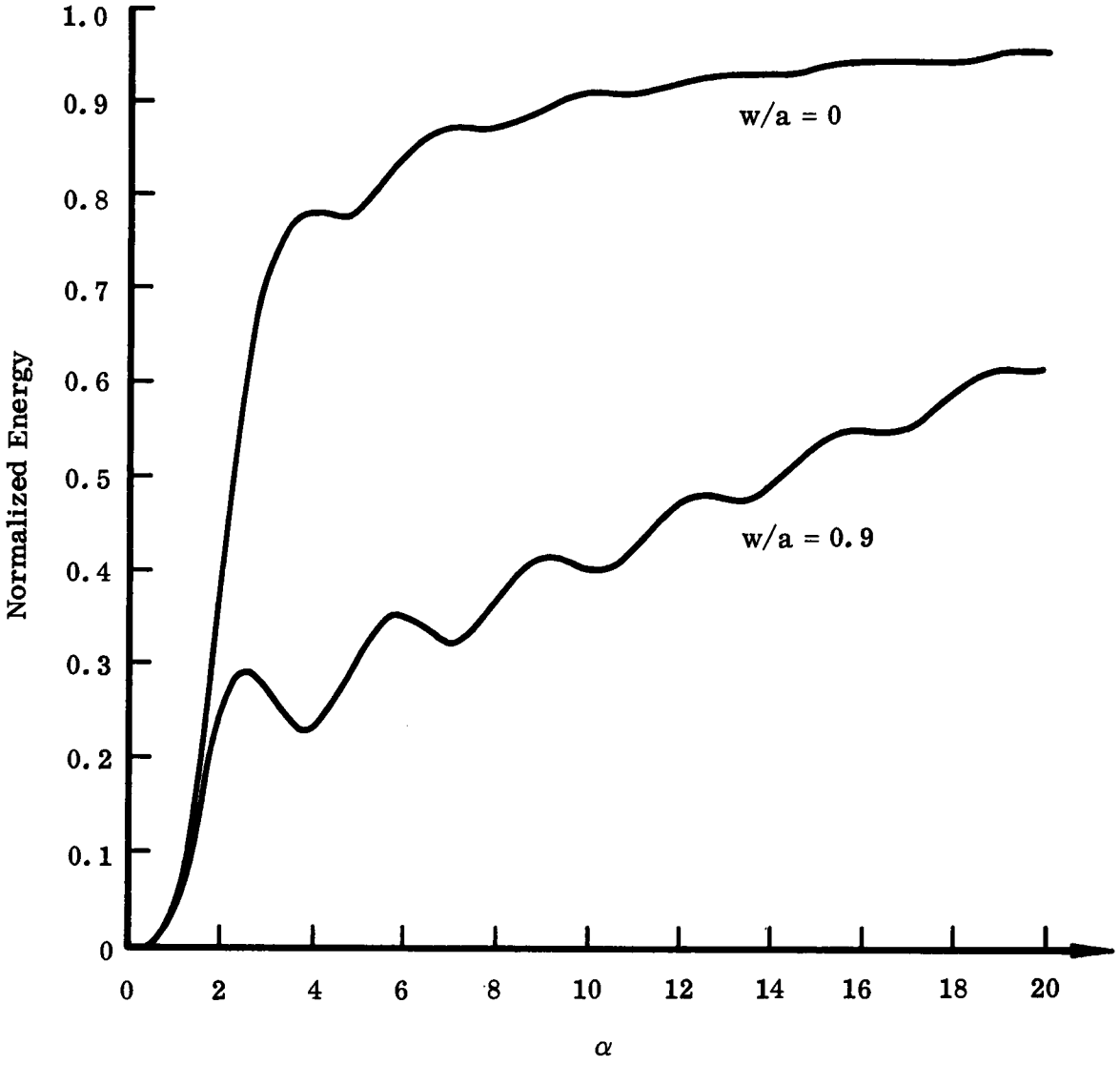


Figure 18. The normalized energy within the entire image (w/a = 0) and near the edge of the image (w/a = .9) of a generalized particle reconstruction.

On the other hand, the w/a = .9 curve represents the energy (again normalized) in the region near the edge of the image, from $r_3 = 0.9a$ to the actual edge $r_3 = a$. It is apparent that the energy is still increasing when the increase in energy of the entire image has already leveled off. Therefore, increasing the hologram size contributes primarily to the detail near the edge of the image. Likewise, the outer portion of the hologram where β is small contributes principally to the definition of the image edge; they do not contribute substantially to the energy in the image.

Thus, the following qualified remarks can be made about the reconstruction of a particle image:

1. When an energy detector is used to "recognize" the presence of an image, a hologram size sufficient to record only the diffraction between the first and second zeros of $J_1(\alpha)$ will be adequate. Those portions of the hologram where β is smaller than some arbitrary value contribute rather insignificantly to the image.
2. When the definition of the edge detail is important in the reconstruction, the smaller values of β cannot be neglected. During the holographic recording, care must be exercised that the higher spatial frequency fringes will be recorded as free of noise as possible.

3. In either case, the hologram obtained with a large ψ will, as β decreases, make less contribution to the reconstructed image than a hologram made with a smaller ψ.
4. Using $\alpha = kaR/z$, it can be seen that as a becomes larger, the hologram radius R necessary to record any particular normalized energy percentage becomes smaller. This is because a larger diffracting object will produce a more compact diffraction fringe distribution than will a smaller object.

The holographic systems discussed thus far have displayed circular symmetry and could be described by cylindrical coordinates. From Equation 3-27, it is expected that systems which have linear symmetry--such as the hologram of a straight wire--have different reconstruction characteristics. To make a comparison of the two systems, the energy in the reconstructed image of a wire with cross-sectional radius a is shown in Figure 19. Like Equation 5-15, the normalized energy in the wire image can be shown to be

$$E(\text{wire}) = 4/(\pi^2(1 - w/a)) \int_{w/a}^{1} \{ \int_{0}^{1} (1/\rho) \sin(\alpha\rho) \cos(\alpha\rho\varepsilon) d\rho \}^2 d\varepsilon \qquad (5\text{-}16)$$

where the meaning of α, w, and a are the same as before. Compared with Figure 18, page 79, it is apparent that the wire image energy has local maxima for α equal to multiples

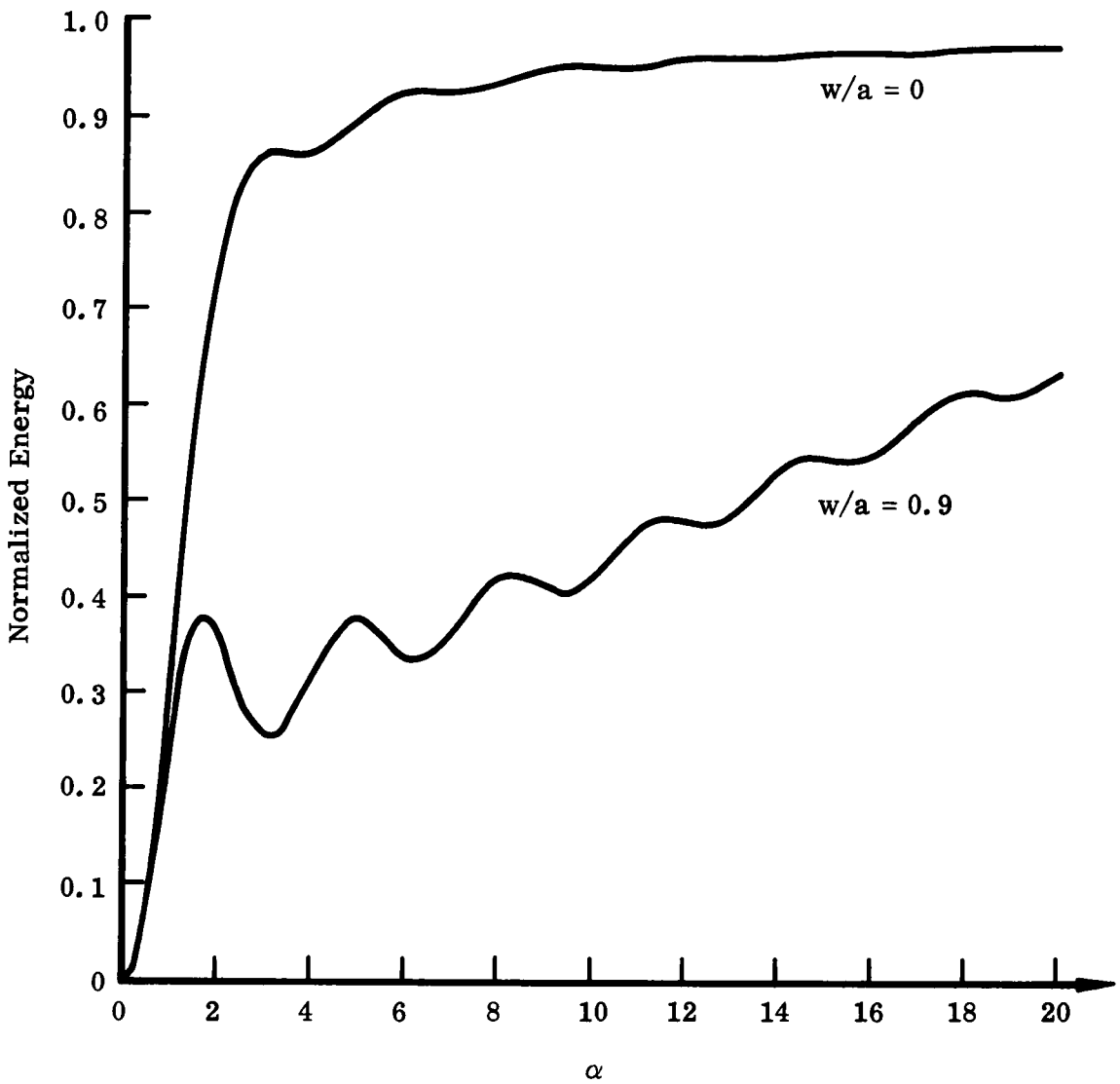


Figure 19. The normalized energy within the entire image (w/a = 0) and near the edge of the image (w/a = .9) of a generalized wire reconstruction.

of π. Since $\alpha = n\pi$ is smaller than the nth zero of $J_1(\alpha)$, the energies at the local maxima are obtained for smaller values of α than for the reconstruction of the aperture. Furthermore, the energy values of the local maxima are greater in the wire reconstruction as compared with the respective local maxima of the aperture reconstruction. The net effect is that in obtaining a given percentage of the normalized reconstructed energy, the hologram of a wire may be smaller than the hologram of an aperture. The steepest portions of the slopes of the curves of Figure 18, page 79, and Figure 19 can be made to overlap when the wire image energy is displaced by $+\Delta\alpha = 0.9$; then the local maxima are also approximately overlapping. Similarly, the $w/a = 0.9$ curves overlap with this $\Delta\alpha$ displacement. Additionally, their rate of increase is virtually identical. These facts imply that the similarities of the reconstruction of circularly symmetrical objects and linearly symmetrical objects (as shown in the figures of Chapter IV) are obtained only when the size of the hologram of the circularly symmetrical objects is larger by $\Delta\alpha \cong 0.9$. Even then, the intensity under the image of the linearly symmetrical objects is greater.

In short, reconstructions of linearly symmetrical objects generally give the impression of having more definition in the images than reconstruction of circularly symmetrical objects. There are three main reasons for this phenomenon:

1. The shape and length of linear objects such as wires, straight edges, or sharp corners contribute substantially to the ease of their recognition as opposed to the more "shapelessness" of circular objects such as particles or aerosol sprays.
2. For a given hologram size, the energy in the reconstructed image is greater for the linear objects than for the circular objects.
3. The diffraction intensity variations from circular objects vary as the first-order Bessel function divided by its argument, whereas, from a linear object, they vary as the sinc function. Because $J_1(\alpha)$ decreases in amplitude with increasing α, $J_1(\alpha)/\alpha$ will be smaller than sinc (α). Consequently, β becomes smaller more rapidly for circular objects than for linear objects. Therefore, the effective contribution to the size of the hologram of a circular object will be less than for the linear object even though the object, cross section, incident intensity, and distance from the film plane are equal.

These effects have been repeatedly observed by the author. In fact, the observations preceded and thus suggested the above analysis. Because of the difficulty of quantitatively describing image quality, the validity of the analysis must, of necessity, remain somewhat subjective.

CHAPTER VI

IMAGE CLARITY AND RESOLUTION

In support of the discussion of Chapter V, this chapter presents experimentally obtained reconstructed image characteristics relating to the intensities in the object and reference beam fields which formed the hologram. The second part of this chapter presents an experimental discussion of resolution.

The intensity in the reconstructed image of a hologram is by Equation 3-14 proportional to β and the hologram area squared. Additionally, from Equations 3-13 and 4-24 the intensity is also dependent on the film sensitivity γ, the reference beam intensity $B_1^{-2\gamma}$, and the reconstruction beam intensity B_2^2. For the results of this chapter, no attempt was made to isolate or evaluate parameters other than β and the film size R since that would have added to the experimental difficulties by requiring precise film processing control. Rather, the emphasis was in obtaining qualitative results using standard film processing techniques. All the results obtained are from the use of Agfa Gevaert 10E75 processed in Kodak D-19. All holograms were made using plane wave fronts in both the object and reference beams. In each case the normal to the film plane bisected the angle formed by the object and reference wave propagation vectors. The reconstruction wave fronts were

plane and the reconstructed real images were observed visually and photographically on Kodak Tri X through a microscope.

For the purpose of optimizing linearity in the recording film, the object and reference intensities must be chosen such that β is less than unity (17). These conditions are frequently implied in discussions of holographic recording (15, 16, 29, 30). Since recording linearity is not in itself a necessary condition for optimizing the reconstructed image intensity (14), the requirement $\beta < 1$ needs reexamination. Consider the intensity incident on the film plane due to a uniform reference wave B and to an object whose amplitude in the film plane is $A(r_2)$,

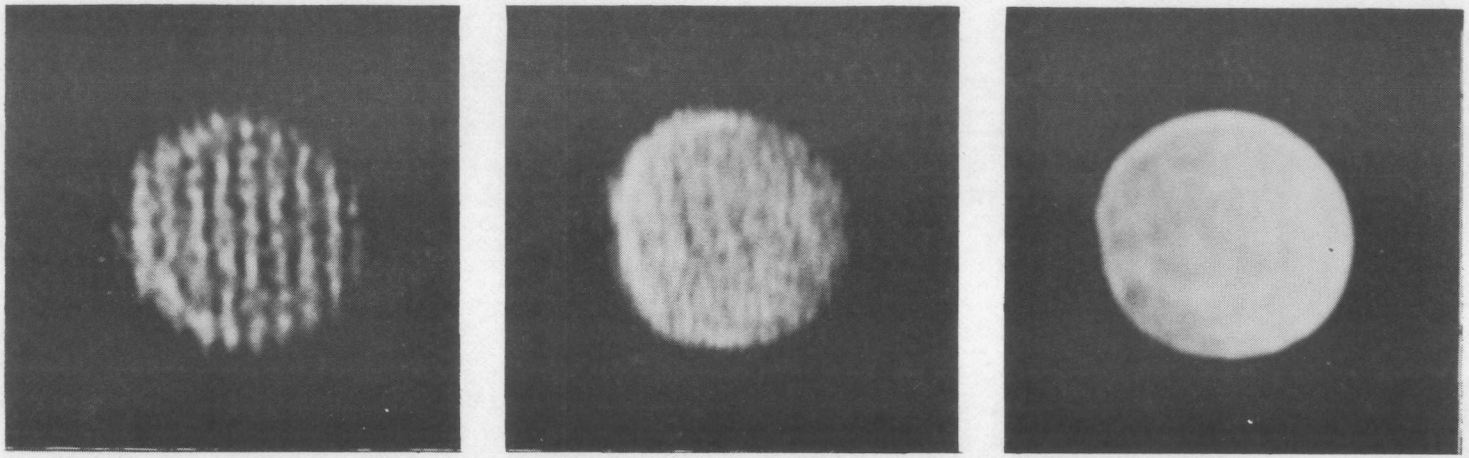
$$I = |A(r_2)|^2 + B^2 + 2BA(r_2) \cos \psi \qquad (6\text{-}1)$$

This equation is similar to Equation 5-8. The holographic information term is proportional to the product $BA(r_2)$ which can be satisfied by a combination of beam amplitudes resulting in $\beta > 1$. This condition can be beneficial when recording small object fields such as particles since the diffracted information resulting from a strong object beam will be recorded over a larger area. The reconstructed images will be brighter and have better definition. However, a detrimental effect is that intermodulation between diffraction from the various objects will increase with the result that the background field of the reconstructed scene will exhibit greater noise.

The three reconstructed images of Figure 20 are of a 165 μ diameter aperture from a $\psi = 7°$ off-axis hologram. The β is indicated for each image. Images a and b were recorded such that the sum of the incident object and reference waves were the same, yielding a film density of about 0.5. Image c was recorded with the same reference beam intensity as image b; however, the increased object beam resulted in fringes on the hologram that were rather heavily overexposed. In other words, a and b were exposed with equal beam energies causing a to have an underexposed bias level and b the near optimum recording density. Image c has the correct bias level, but the higher spatial frequency fringes have much greater contrast because of the increased object beam intensity. The relative improvement in the definition of the reconstructed images is readily apparent. The intensities of the reconstructed image are implied by the exposures which were necessary to record the images.

Another example of effects of varying β was obtained by making off-axis holograms of apertures of 5μ, 7μ, 10μ, and 15μ diameters with $\beta = 30, 10, 3, 1, 1/3, 1/10$, and $1/30$. The reference beam intensity was the same in each case; the exposure level on the film was chosen to give a density near the optimal 0.5. The results of the reconstructions are summarized in Table IV. The significance of the results applies to recording small particle fields: as β decreases, small objects either become lost in noise, or their image intensities are too weak to be detected.

a b c



β = 3, Exposure = 45 sec
Weak Reference Beam

β = 1/3, Exposure = 6 sec
Normal Reference Beam

β = 3, Exposure = 1/4 sec
Normal Reference Beam

Note: The exposure times indicate the relative intensities of the images.

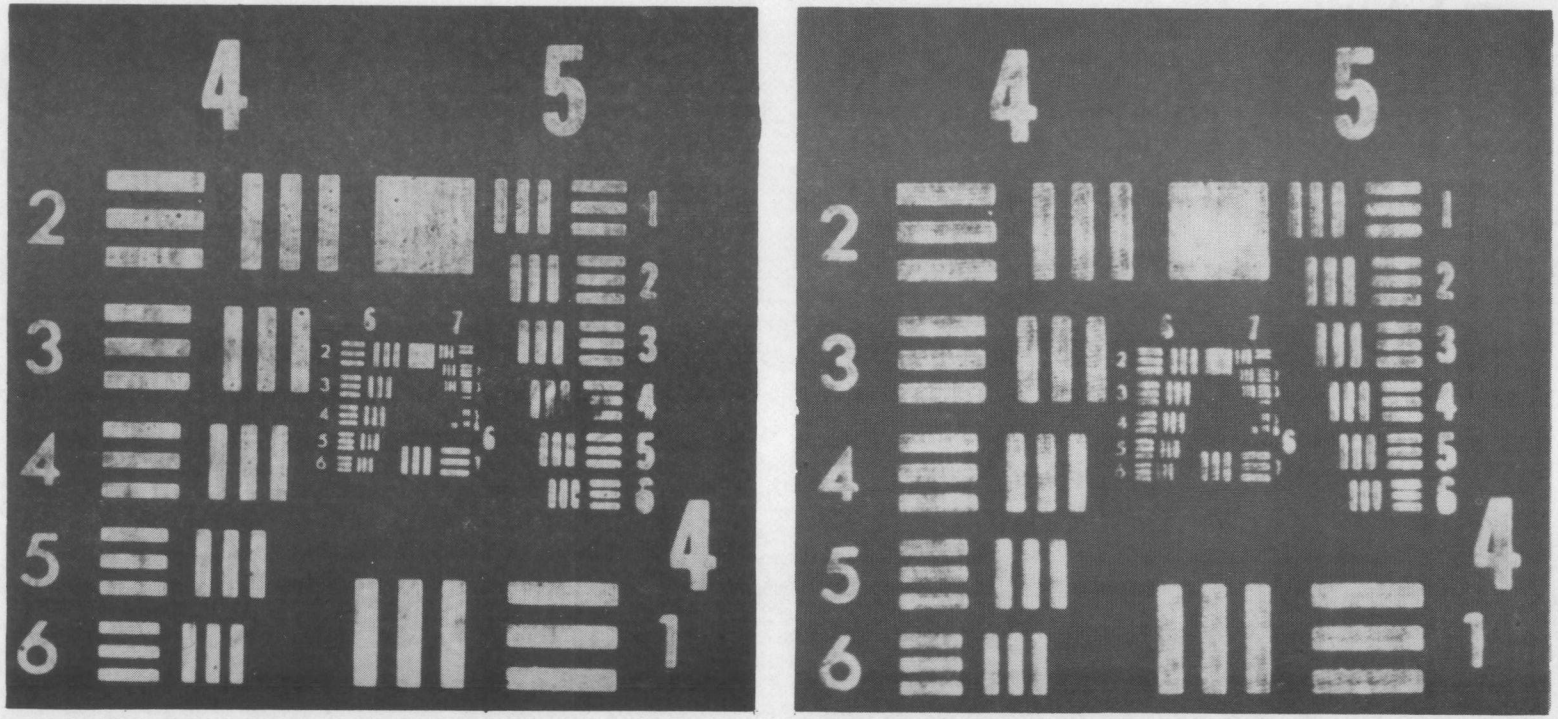
Figure 20. The in-focus, off-axis reconstructions of a circular aperture for different reference beam intensities and β.

TABLE IV

OFF-AXIS RECONSTRUCTIONS OF 5μ, 7μ, 10μ, AND 15μ DIAMETER APERTURES

β	Reconstruction Exposure Time (sec)	Reconstruction Observations
30	1/4	All images clearly visible
10	1/4	All images clearly visible
3	1/2	All images clearly visible
1	5	5μ image very weak
1/3	20	5μ image in noise, 7μ image very weak
1/10	40	5μ image not visible, 7μ image in noise, 10μ image very weak
1/30	80	5μ, 7μ, 10μ images not visible, increased exposure time only increases noise

To test the definition and resolution of a more dense object field, the Air Force Resolving Power Test Target 1951 (abbreviated to TT) was used. For off-axis holograms the TT used had an opaque field in which the details of the TT were like apertures. For in-line holograms a complimentary TT was used, one in which the groups of elements are opaque. The spatial frequency of the elements in each group has been given in Table I, page 40. Figure 21 presents both a photographic image of the TT and a high resolution off-axis reconstruction. The reconstruction (for which z_2 = 114 mm) is capable of resolving with good definition the smallest lines on the TT. The hologram was recorded and developed without undue attention given to any particular steps in the procedure. Higher resolution has been presented by Champagne (31). However, in this work extraordinary care was placed on every detail of the holographic process. For example, Champagne used microflat glass plates for mounting the film emulsions in order to minimize phase errors. Furthermore, he found it necessary to index-match the glass base of the photographic plate during the recording step to eliminate multiple reflections. The resolution he obtained was approximately 600 lines/mm using a recording and reconstruction distance of 42 mm. Equally significant, to realize such high resolution, an alignment tolerance of better than ±15 seconds of arc had to be satisfied. The resolution of elements 7/6 on the TT in Figure 21 required an alignment of an extimated ±2 minutes



Photographic Image

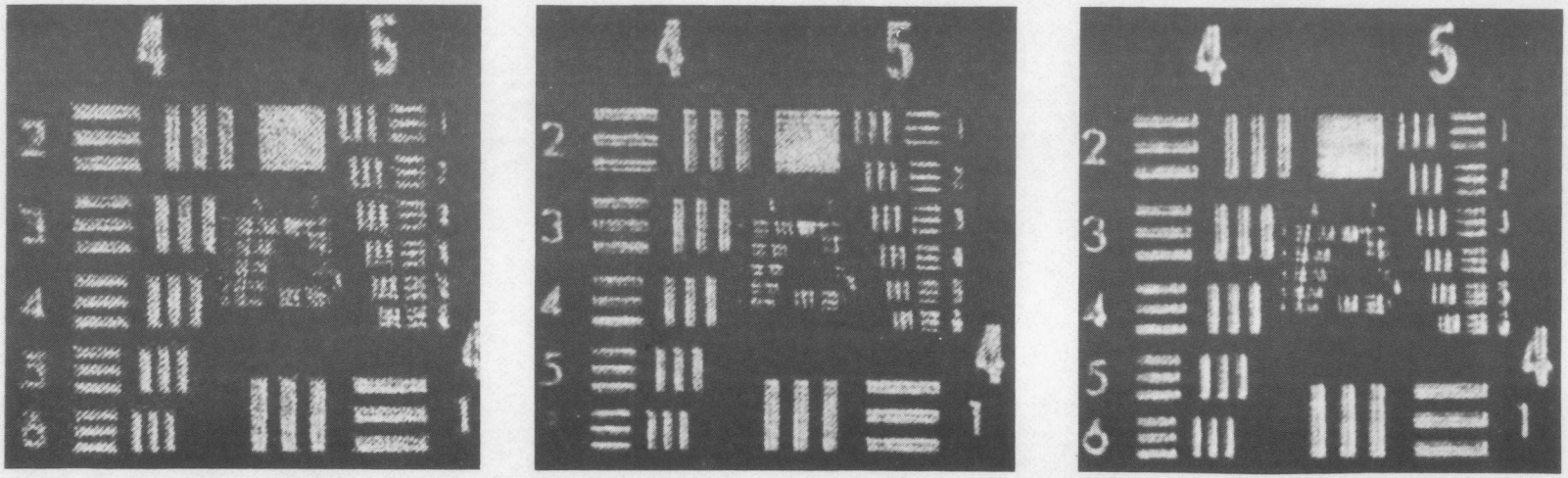
Off-axis Reconstructed Image

Figure 21. Comparison of the photographic and off-axis reconstructed image of the Resolving Power Test Target.

of arc. (The notation group/element refers to the description of the TT in Table I, page 40.) It may be emphasized that obtaining such alignment accuracy is a difficult trial-and-error, time-consuming task, one that is not amenable to large quantity data reduction nor to automation.

Nevertheless, Figure 21 serves to illustrate the imaging quality of holography. Invoking the parallels drawn in Chapter III between conventional lens imaging and holographic reconstruction imaging, Figure 21 reinforces the contention that holography may be considered as one of a class of image projection techniques rather than as unique unto itself. As a member of the class, it exhibits the property that the image characteristics are, to a large degree, dependent on the aperture function of the optical system under consideration. It is worth noting that the mechanisms of wave front shaping (reflective, refractive, or defractive) are in themselves not uniquely deciding factors of image quality.

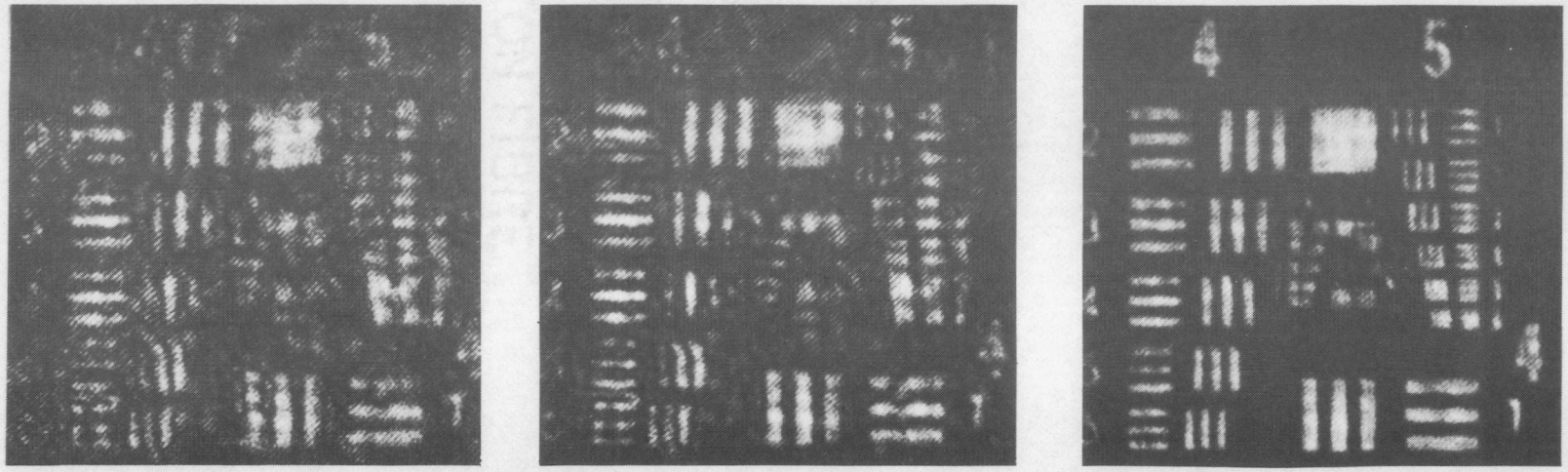
The influence of β on the definition and resolution of the reconstructed images of the TT can be observed by comparing Figures 22, 23, and 24. The focal distance for each image in these figures was 35 cm. Parenthetically, it should be mentioned that some of the image differences are marked by subtleties which are more apparent by direct observation than when presented after several generations of reproduction as is necessary in publications.



β = 1/30, Resolution: 5/4 β = 1/10, Resolution: 6/1 β = 1/3, Resolution: 6/1

Note: The sum of the object and reference beam intensities was the same in each case.

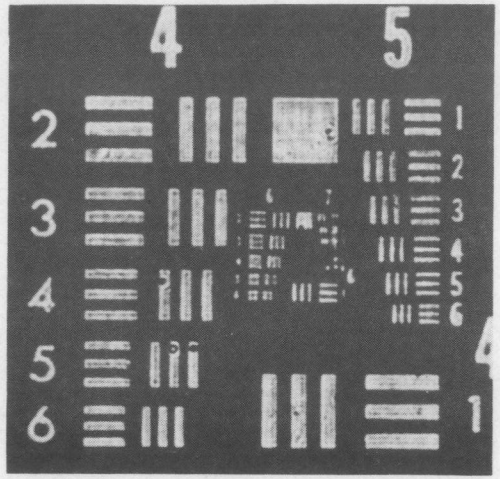
Figure 22. The off-axis reconstruction of the Resolving Power Test Target for different β and normal exposure of the object and reference beams.



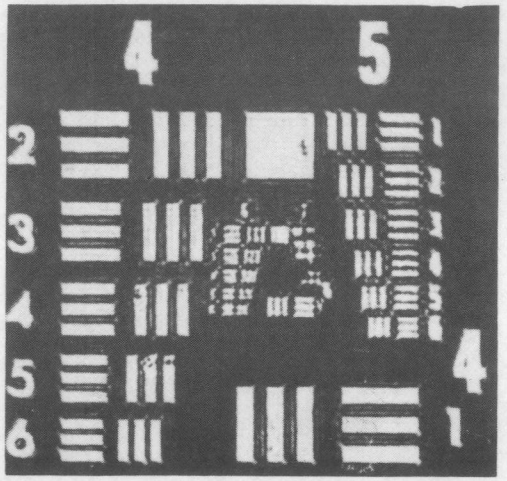
β = 30, Resolution: 4/3 β = 10, Resolution: 4/6 β = 3, Resolution: 5/3

Note: The sum of the object and reference beams was the same as in Figure 22.

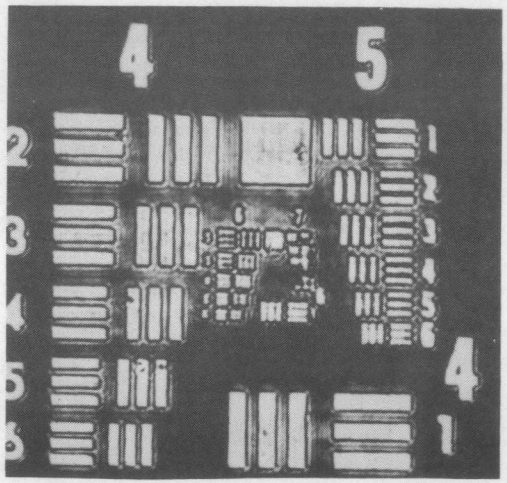
Figure 23. The off-axis reconstructions of the Resolving Power Test Target for different β, under-exposed reference beams, and normally exposed object beams.



β = 3, Resolution: 6/3



β = 10, Resolution: 6/2



β = 30, Resolution: 6/2

Figure 24. The off-axis reconstructions of the Resolving Power Test Target for different β, normally exposed reference beams, and over-exposed object beams.

The intensities of the object and reference beams used during the recording of the holograms which produced the images of Figure 22 were adjusted so that their sum was the same for each β. This was done to insure that the illuminating energy of the holographic system was the same in each case. For the three values of β in Figure 22, the resolution and definition are only slightly different; however, it was found that as β decreased the intensity of the reconstructed image decreased, and the apparent resolution became poorer. The decline in resolution may be subjective; it was not apparent whether or not the resolution and definition could have been improved had a more intense reconstructing beam been available.

The summed intensities of the object and reference beams for the holograms yielding the images of Figure 23 were the same as in Figure 22. In this case, as β increased the reference beams became weaker. The clarity of the reconstructions are obviously decreasing with increasing β. Furthermore, compared one-to-one with the corresponding reciprocal β images of Figure 22, the images are decidedly poorer. Thus, an underexposed reference beam will cause the reconstruction to be degraded in quality. This would seem to be contradictory to the implications of Equation 6-1 and to the film intensity resolution discussed in Chapter V. However, the intensity resolution was described in terms of observing the modulating effect of a weak "signal" beam on a normal "reference" beam. In the present case, the intensity

is a weak reference beam combined with the diffracted signal which decreases as the spatial frequency increases. The result is simply a combined intensity which beyond some spatial frequency cannot satisfy the effective threshold of the film sensitivity.

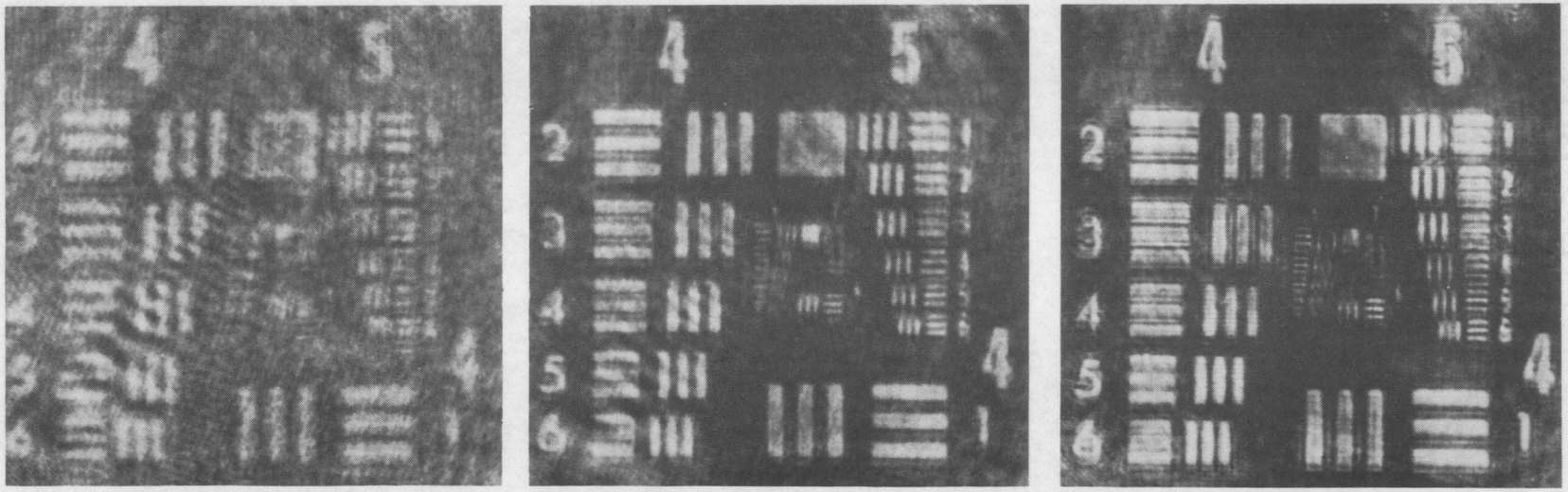
If, on the other hand, the reference beam is chosen correctly and the object beam is increased to give a large β, the reconstructed images appear as in Figure 24, page 95. Here the resolution has increased, and the contrast has improved. Additionally, it may be seen that as β increases the intermodulation distortion arising from the interference between the diffraction from various portions of the TT also increases. Some of this distortion proves useful in that it tends to give a contrasting outline to the image. However, in the sense that the fidelity of the image is degraded, the intermodulation distortion is, of course, undesirable. The overall effect of Figure 24 is to show that when the intensity in the reference beam is sufficient to bias the film properly, then an object-to-reference beam ratio β of some reasonable value greater than one will improve the definition and resolution of the reconstructed images.

The TT which was the holographic object of Figures 20 through 24, pages 88, 91, 93, 94, and 95, was, in each case, a mask in which the diffracting elements were apertures. The mask blocks the dc portion of the object beam; otherwise, for $\beta > 1$, the holographic film would be overexposed. To utilize a $\beta > 1$ when making holograms of more

conventional object fields requires that the object beam have the dc component removed by suitable filtering.

For comparison purposes, Figure 25 presents the reconstructed TT images from in-line holograms which were exposed to different densities. As the recorded density increases, the contrast in the reconstruction of the image to the background also increases. When the density is increased past the optimum value (about 0.5), the entire image becomes progressively less intense. The resolution of the overexposed holograms remains good; however, image degradation similar to the intermodulation distortion of Figure 24, page 95 becomes more pronounced. Although the resolution displayed by the in-line reconstruction is similar to that of the off-axis reconstruction, the general clarity or "sense of quality" is not as good as for the off-axis case. These general observations are used as experimental justification of the statement made at the end of Chapter III to the effect that no inherent difference is expected between the resolutions of the reconstructed images from in-line and off-axis holograms.

One of the chief advantages of in-line over off-axis holography is considered to be the experimental simplicity and the rather uncritical hologram alignment necessary in the reconstruction process. However, while recording the reconstructed images of Figure 25, it was found that to obtain the best definition and resolution required about as careful an alignment of the hologram as for the off-axis



Film Density = 0.25
Resolution: 5/3

Film Density = 0.5
Resolution: 6/2

Film Density = 1.2
Resolution: 6/1

Figure 25. The in-line reconstruction of the Resolving Power Test Target for different exposed film densities.

cases of Figure 22, page 93. This simply states that alignment requirements are dictated by the desired accuracy of the reconstructed phase front combination and not by a distinction of in-line or off-axis holography.

Resolution of optical systems is often expressed in terms of Rayleigh's criterion. The simple parametric requirements of the criterion make its application useful primarily only for resolving point or line objects. When the object field in question consists of extended objects such as the group of bars on the TT, direct application of the Rayleigh criterion can give rather unsatisfying results. This will be shown by the following discussion of experimentally obtained resolution parameters.

The spacing of the bars on the TT are such that the amplitude of its image intensity would ideally form a square wave as in Figure 26. By dc filtering the image the square wave is reduced to a set of delta functions which are located where the leading and trailing edges of the square wave were located. An actual reconstructed image is, of course, degraded from a square wave; dc filtering of an actual image will, in turn, produce spread functions at the image edges. Assume that resolution of the bars on the TT is obtained when their edges are sufficiently separated that the spread functions are resolved. The resolution of the spread function can be considered in the same spirit as the Rayleigh criterion. To quote Rayleigh (32):

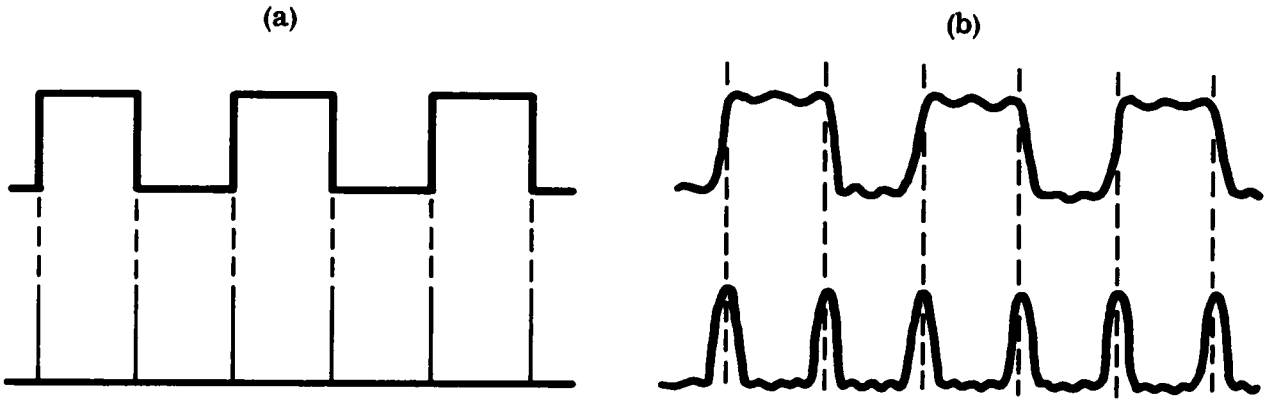


Figure 26. The image intensity of a bar chart and its dc filtered image for (a) an idealized case with infinite resolution and (b) a realistic case with finite resolution.

> This rule is convenient on account of its simplicity and it is sufficiently accurate in view of the necessary uncertainty as to what exactly is meant by resolution.

The spread function associated with the dc filtered bar image is the line spread function described in Equation 3-26. By letting the argument of the sinc function be η, then

$$\eta = k_2 x_3 L / z_2 \tag{6-2}$$

Since the resolution of the TT is conveniently given by the spatial frequency f_s (lines/mm) of the group/element number, x_3 (in mm) is better expressed by $x_3 = 1/2\ f_s$. Then

$$\eta = \pi L / \lambda z_2 f_s \tag{6-3}$$

The value of η serves to define the resolution: given values of η, L, λ, and z_2 determine f_s which, because it is expressed in lines per millimeter, gives the spacing of the resolved lines.

The value of η was found experimentally by the following procedure. A hologram of the TT was aligned to yield the best reconstruction. The parametric values of λ and z_2 were known as part of the experimental conditions. An adjustable iris was placed in the reconstructing beam so that its diameter could be continuously adjusted from the maximum diameter of 5 cm down to about 2 mm. The hologram was centered on the axis of the reconstructing beam. While viewing the reconstructed image through a microscope, the

diameter of the reconstructing beam was adjusted to allow a particular group/element on the TT just to be resolved. This resolution was completely judgemental; however, the results were found to be quite repeatable. Measuring the diameter of the beam that allowed the group/element to be resolved then gave sufficient parameters to calculate η using Equation 6-3. The same procedure was repeated for the other group/element units. Typical results where z_2 = 35 cm and λ = .6328μ are shown in Figure 27. The error trees indicate the range of the data while the circles indicate the average value for each spatial frequency. The average over all of the spatial frequencies plus-minus the standard deviation is $\eta = 2.78 \pm .10$. The experiment was repeated for holograms of the TT in which the object-to-reference beam ratio, the focus distance, the angle between the object and reference beams ψ, and the recorded size of the hologram were varied. The averages within their standard deviations were all found to overlap.

The importance of the results of Figure 27 is that the resolving ability of the hologram reconstruction is better than indicated by the Rayleigh criterion. By that standard, the maximum of one spread function should overlap the first minimum of the adjacent spread function; the separation of the maxima for the line spread function should be given by $\eta = \pi$. Figure 28 shows the resulting intensity distribution for an $\eta = 2.78$ separation of two line spread functions. The shallow dip in the center region indicates

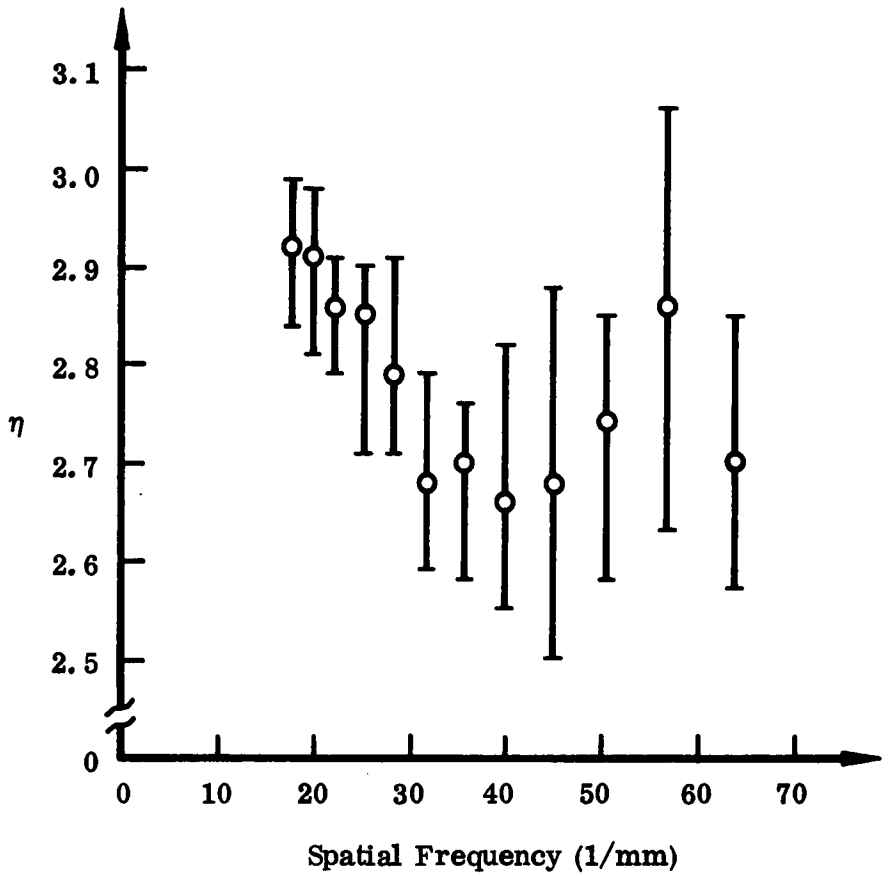


Figure 27. Experimental determination of the resolution defining parameter η.

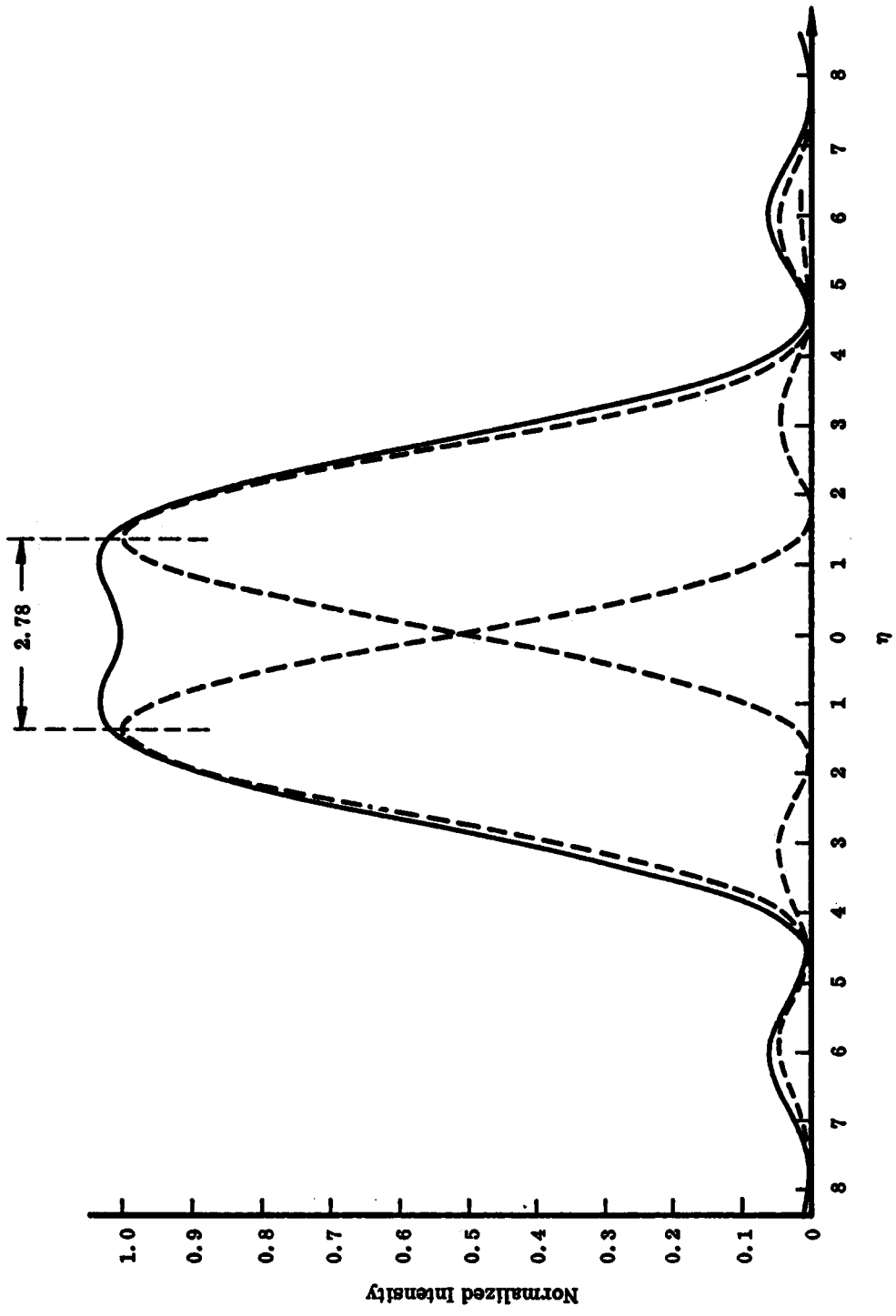


Figure 28. The intensity distribution resulting from two line spread functions separated by $\eta = 2.78$.

that the resolving capability closely follows the Sparrow criterion of resolution (33). By this criterion, resolution is defined to be that separation of the spread function for which the central minimum of the combined intensity ceases to exist. For line spread functions, this separation (33) is given by $\eta = 2.606$ which is quite close to the experimentally obtained 2.78 value. The application of the Sparrow criterion must, in the author's opinion, be considered from a pragmatic viewpoint as demonstrated by the case at hand. The TT images were considered resolved when the bars of the group/element could just be distinguished. Yet by the Sparrow criterion the intensities would have merged to give a broad image without defining center features. Clearly, this discrepancy between physical observation and the resolution's mathematical definition serves to illustrate the elusiveness of obtaining a generally adequate definition of resolution.

For the sake of comparison, a similar resolution test was made of a 5 cm diameter f/4 diffraction limited lens. The distance between the TT and the lens was 40 cm. A variable aperture was placed directly in front of the lens, and a He-Ne laser illumination was used. Again, the individual group/element images were rendered barely resolvable by adjusting the aperture which was then measured. Application of Equation 6-3 gave $\eta = 2.41 \pm .09$. Although this value is somewhat smaller than that obtained holographically, the actual appearance of the microscopically

observed images was in no manner better defined, more distinct, or more easily recognizable than from the holographic reconstruction. Again, the η obtained implies resolution conditions which meet the Sparrow criterion better than the Rayleigh criterion. The smallest group/element that could be resolved was 5/6. Using this together with the lens aperture of 5 cm diameter and z = 40 cm, then by Equation 6-3 $\eta = 2.72 = (.866)\pi$. In other words, less than one Airy zone of the diffraction was necessary for the lens imaging system to achieve the 5/6 resolution. Indeed, this is precisely what is implied by all of the above results. No more direct comments need be made other than it appears that such resolution is readily realizable as can be expected by having the lens diameter and hologram diameter equal in size to the first Airy diffraction zone.

The focal depth resolution of holographic reconstruction was also determined experimentally. An object field of equally spaced 100μ diameter wires was made. The object is pictorially like a ladder in which the wires serve as rungs. The wire spacing was 3.048 ± .008 cm. The wires were positioned in a 5 cm collimated He-Ne laser beam such that the longitudinal axis of the "ladder" was displaced from the illumination propagation direction sufficiently for all the wires to be distinctly illuminated. The effective spacing between the wires along the direction of the beam propagation was then 3.043 ± .008 cm. The distance from the nearest wire relative to the film plane was 32 cm.

Reconstruction was by the same collimated He-Ne laser beam. In the reconstruction the distances between successive wire images (judged to be at their best focus) were measured. Averaging and calculating the standard deviation, the distance between the reconstructed images was 3.04 ± .15 cm. From multiple repetition of the experiment, the standard deviation was found to vary by about ±.03 cm. The total spread of the position uncertainty is twice the standard deviation. Assuming that each wire contributes an equal amount to this uncertainty, one can conclude that the positioned uncertainty of each wire image is half the distance spread. Thus, the focused image is located with an uncertainty of 0.15 ± .03 cm. It should be emphasized that this is an average resulting from many data points. It is not intended to imply that each wire reconstruction can be focused satisfactorily within such a tolerance. Rather, the average result indicates that a particular focused image position cannot be considered more accurately located than allowed by the standard deviation.

The focus uncertainty, called Δz, can be used in several ways to obtain an estimate of the effective size of the hologram of each wire. The hologram size is obtained using Equation 3-21 in the form

$$L = z_2/(\Delta z/\lambda)^{1/2} \qquad (6\text{-}4)$$

If Δz and λ can be satisfactorily determined for any single object located at z_2, the hologram size L follows. Likewise,

consider the hologram size in terms of the number n of Airy zones of the diffraction pattern recorded. Again, assuming linear symmetry

$$kaL/z_2 = n\pi \tag{6-5}$$

and, in combination with Equation 6-4,

$$n = 2a/(\Delta z\lambda)^{1/2} \tag{6-6}$$

Equation 6-6 simply points out once more the interdependence of the hologram size (represented here by n), the object dimension a, and the focal uncertainty Δz. Applying Equation 6-6 to the above experimental results indicates $6 \lesssim n \lesssim 7$. In a somehwat less direct manner Equation 6-6 also illustrates the effectiveness of the $\beta > 1$ holographic technique. For $\beta > 1$, the effective size of the hologram is increased because higher spatial frequency fringes are recorded. Correspondingly, the focal uncertainty is smaller than would be the case for $\beta \leq 1$. This is particularly important when a is small since the energy diffracted is directly proportional to the cross-sectional area of the diffracting object. Quite apparently, small values of a will have attendant small Δz, and with $\beta > 1$ the effect of having a very shallow focal uncertainty is additionally enhanced.

Equation 6-4 cannot be applied directly to finding the effective hologram size in the above experiment because z_2 varies for each wire. However, analogous to the averaging

procedure used in evaluating the experimental data, an average $\overline{\Delta z}$ spread resulting from the simultaneous consideration of all the wire image positions can be found. Let

$$z_2 = \zeta + md \tag{6-7}$$

where ζ is the distance from the film plane to the nearest wire (32 cm), d is the spacing between the wires (3.043 cm), and m ranges from zero to N, the total number of wires in the object field (which was 17). Now

$$\overline{\Delta z} = (1/N) \sum_{m=0}^{N-1} \Delta z_m = (1/N)(\lambda/L^2) \sum_{m=0}^{N-1} (\zeta + md)^2 \tag{6-8}$$

Solving for L results in

$$L = [(\lambda/\overline{\Delta z})(\zeta^2 + \zeta d(N - 1) + (1/6)d^2(N - 1)(2N - 1))]^{1/2} \tag{6-9}$$

Using $\overline{\Delta z} = (.15 \pm .03)$ cm, then Equation 6-9 yields $(1.09 \lesssim L \lesssim 1.34)$ cm. It must again be emphasized that this is an average result. Also, recalling the definition of L, the hologram size is actually given by 2L. Regarding the composite results of this experiment, it appears the hologram size of the wire at its average location was only about half the size of a maximum value determined by the beam diameter.

CHAPTER VII

EXPERIMENTAL DETERMINATION OF HOLOGRAPHIC LINE SPREAD FUNCTIONS

In Chapter III, the characteristics of the point and line spread functions of the holographic system were derived, and its similarity with lens imaging optics was stressed. In this chapter, the importance of the spread function to fundamental descriptions of optical systems is discussed, and the experimentally determined line spread functions of a holographic system is presented.

The theoretical and experimental foundations of the applicability of the line spread function to the analysis of photographic systems has been developed extensively in an excellent series of articles by Lamberts, et al. (34, 35), Perrin (23, 36), and Higgins, et al. (37). The Fourier transform technique of optical image evaluation, that is to say, the transfer function, is described at length by Smith (38).

The applicability of the spread function and transfer function methods of image analysis to a holographic system can be viewed from a linear systems approach to the problem (15). Thus, the entire holographic system--i.e., the diffraction phenomenon, the holographic recording, the reconstructing diffraction, and the detection of the reconstructed image--is considered as a "black box" optical

device that has the useful property of transmitting an image from one place to another. For any given input to the black box, there exists a corresponding output whose facsimile fidelity has been degraded by the transmission process. For instance, the energy in the transmitted image of a step function input such as an illuminated knife edge does not have a step function distribution. Rather, the energy is "spread" so that a portion falls into the geometrical shadow of the step function. The black box viewpoint of a holographic system, in addition to being a vehicle for a systems approach, is somewhat intuitively justified by again referring to the similarity of a point image obtained by a lens system and a holographic system. By the results of Chapter III, the physical properties of either image (under appropriate conditions) do not determine which system was used in making the image. The black box internal components can equally well be a lens system, a mirror system, or a holographic system. Indeed, from this viewpoint, holography is "merely" another imaging system; its unique properties are characterized by the image storage capabilities but not by the image characteristics themselves.

It is appropriate to draw attention to several properties of linear systems. An image (the output) function $i(x_2, y_2)$ is correlated with the object (the input) function $0(x_1, y_1)$ by the superposition integral (15),

$$i(x_2, y_2) = \iint_{-\infty}^{\infty} 0(x_1, y_1)\ \sigma(x_2, y_2; x_1, y_1)dx_1dy_1 \qquad (7\text{-}1)$$

where $\sigma(x_2, y_2; x_1, y_1)$ is the response of the system to an impulse function. With unit magnification, the space described by (x_1, y_1) is the same as the (x_2, y_2) space, and the object and image can be compared point for point. When σ is dependent only on the coordinate differences $(x_2 - x_1)$ and $(y_2 - y_1)$, the linear system is space invariant. Then,

$$i(x_2, y_2) = \iint_{-\infty}^{\infty} 0(x_1, y_1)\ \sigma(x_2 - x_1, y_2 - y_1)dx_1dy_1 \qquad (7\text{-}2)$$

This equation can be equivalently expressed by

$$i(x_2, y_2) = \iint_{-\infty}^{\infty} \sigma(x_1, y_1)\ 0(x_2 - x_1, y_2 - y_1)dx_1dy_1 \qquad (7\text{-}3)$$

although the (x_1, y_1) coordinates are changed in sign from before. These equations are each a two-dimensional convolution of the object function with the impulse response function. Space invariance requires that the system operating on a moving object point result in moving the image point without changing its functional form. Optically this requirement demands an aberration free system. The convenience of the convolution integral makes it desirable to be able to apply this technique to more general cases. Thus, the object field domain is restricted sufficiently so

that the image field is satisfactorily space invariant. The object field may be composed of several adjacent domains, each domain having associated with it a different impulse response. The paraxial condition used in Chapter III describes the conditions for a satisfactory domain.

The point spread function of an optical system is the physical equivalent of the impulse response. When the spread function of an optical system is known, the entire transformation characteristics of the system can be obtained from Equation 7-2 or 7-3. When the system is composed of multiple subsystems, the spread function of the entire process is obtained by convolving the spread functions associated with each subsystem (23),

$$S(x, y) = \int_{-\infty}^{\infty} S_1(x', y')\, S_2(x - x', y - y')dx'dy' \qquad (7\text{-}4)$$

Conversely, the spread function of a subsystem can be found by a deconvoluting process. Since this is usually very difficult to perform mathematically, it is advantageous to use the Convolution Theorem by which the Fourier transform of the convolution of two functions in distance space is equivalent to the multiplication of the Fourier transforms of the two functions in spatial frequency space (39). Thus, spatial deconvolution is the equivalent operation to division by the correct Fourier transform and finding the inverse Fourier transform of the result,

$$S_1(x, y) = \mathcal{F}^{-1}\{ [S(x, y)]/\mathcal{F}[S_2(x, y)]\} \quad (7\text{-}5)$$

The Fourier transform of the spread function of an optical system is known as the optical transfer function (OTF) of the system. Because the Fourier transform pair--that is, the Fourier transform and the inverse Fourier transform--have the reciprocal relationship that knowledge of one implies knowledge of the other, the OTF can, like the spread function, be used to obtain all the imaging characteristics of the optical system. The process of obtaining the Fourier transform of the real spread function results in both a real and an imaginary component. When plotting these in a complex space, the magnitude (the square root of the sum of the real and imaginary components squared) is called the modulation transfer function (MTF), whereas the argument is the phase transfer function (PTF). The MTF represents the relative amplitude by which different spatial frequencies are transmitted through the system (37). The PTF describes the displacement of wave fronts which sinusoidal distribution of various spatial frequencies suffer in transmission through the system (38).

If the MTF and PTF of a system are known, the spread function can be obtained by finding the real and imaginary parts of the OTF and using these to calculate the inverse Fourier transform. It is typical of good optical systems that phase variations are significant only when the MTF is small, which is when the spatial frequency is large. Then

the imaginary component of the OTF is quite small, and the inverse Fourier transform may be satisfactorily determined using only the MTF. Because the MTF is the magnitude of the OTF, it is always positive. The spread function obtained from the inverse Fourier transform of the MTF must be a symmetrical function. Conversely, an experimentally determined asymmetry in the spread function indicates the presence of non-negligible phase distortion.

A more visual representation of the definition of the MTF can be obtained by considering the optical system to be operating on the image projection of a system of parallel lines whose intensity varies sinusoidally with an adjustable spatial frequency. The transfer function is defined (38) as the ratio of the contrast of the image at each spatial frequency to the contrast of the object at the same spatial frequency. By contrast is meant the classical Michelson visibility--the difference of the maximum and minimum of the function variations divided by their sum. The utility of the transfer function definition is a direct result of the fact that the image of a sine wave object illumination will also have a sine wave illumination distribution (38).

This result is sufficiently important to be included in the body of this text. Consider the system of parallel lines to be adequately described in the one dimension which is perpendicular to the length of the lines. The object illumination is

$$O(x_1) = 1 + \Omega \cos(2\pi x_1 f_s) \qquad (7\text{-}6)$$

where $\Omega \leq 1$, and f_s is the spatial frequency which is the reciprocal of the period of the oscillation. The image may be obtained using Equation 7-3 in an analogous one-dimensional form,

$$i(x_2) = \int_{-\infty}^{\infty} S(x_2 - x_1)\ O(x_1)dx_1 \qquad (7\text{-}7)$$

where $S(x_2 - x_1)$ is the line spread function. Using a change of variables $\chi = x_2 - x_1$ and expanding the cosine term results in

$$\begin{aligned} i(x_2) = & \int_{-\infty}^{\infty} S(\chi)d\chi \\ & + \Omega \cos(2\pi x_2 f_s) \int_{-\infty}^{\infty} S(\chi) \cos(2\pi\chi f_s)d\chi \\ & + \Omega \sin(2\pi x_2 f_s) \int_{-\infty}^{\infty} S(\chi) \sin(2\pi\chi f_s)d\chi \end{aligned} \qquad (7\text{-}8)$$

The integrals each have numerical values which will be designated respectively by M_o, M_c, and M_s. Then

$$i(x_2) = M_o + \Omega\ M_c \cos(2\pi x_2 f_s) + \Omega\ M_s \sin(2\pi x_2 f_s) \qquad (7\text{-}9)$$

Clearly, the M_c and M_s integrals are the cosine and sine components of the Fourier transform of the line spread function. Considering these as two independent variables,

they may be written as

$$M_c = M \cos \theta$$

$$M_s = M \sin \theta \tag{7-10}$$

and the image reduces to

$$i(x_2) = M_o + \Omega M \cos(2\pi x_2 f_s - \theta) \tag{7-11}$$

This shows the sine wave illumination distribution of the image with a modulation M and a phase change θ. If the spread function is symmetrical, the sine integral is zero, $M_s = 0$, and the phase shift is also zero. Then, the application of the transfer function definition yields a transfer function equal to M/M_o.

Generalizing this procedure to a complex object function,

$$O(x_1) = 1 + \Omega e^{i\phi(x_1)} \tag{7-12}$$

gives

$$i(x_2) = \int_{-\infty}^{\infty} S(\chi) d\chi + \Omega \int_{-\infty}^{\infty} S(\chi)\, e^{i\phi(x_2-\chi)} d\chi \tag{7-13}$$

When the phase function has the linear property, $\phi(\alpha + \beta) = \phi(\alpha) + \phi(\beta)$, the image becomes

$$i(x_2) = M_o + \Omega e^{i\phi(x_2)} \int_{-\infty}^{\infty} S(\chi) e^{-i\phi(\chi)} d\chi \tag{7-14}$$

If the contrast of the complex object function is considered to be the amplitude Ω of the modulation, then, by definition, the transfer function becomes

$$OTF = \left[\int_{-\infty}^{\infty} S(\chi)e^{-i\phi(\chi)}d\chi\right]/\left[\int_{-\infty}^{\infty} S(\chi)d\chi\right] \qquad (7\text{-}15)$$

which is the Fourier transform of the spread function normalized by the area under the spread function.

Experimentally, it is difficult to obtain the point spread function. Consider an optical system which images a point source of radiation, i.e., an illuminated pinhole. In principle, the point spread function is obtained from the trace resulting from scanning the image with a suitably small pinhole. However, the energy passing through the pinhole is insufficiently small for accurate measurement. On the other hand, if a line source of radiation is imaged, the image may be scanned with an elongated slit parallel to the image axis. This light distribution is referred to as the line spread function (LSF). Clearly, the LSF is the integration of the point spread function $S(x_1, y_1)$ along one of its coordinates,

$$S(x_1) = \int_{-\infty}^{\infty} S(x_1, y_1)dy_1 \qquad (7\text{-}16)$$

The simplest one-dimensional object is an illuminated knife edge; the illumination function is simply a step

function. The step function changes the limits of integration of Equation 7-3, and the image becomes

$$i(x_2) = \int_{-\infty}^{x'} S(x_1)dx_1 \qquad (7\text{-}17)$$

where x' is the boundary of the step function. When the system has unity magnification so that the coordinates x_1 and x_2 are dimensionally the same, differentiation of the image

$$\frac{di(x)}{dx} = S(x) \qquad (7\text{-}18)$$

indicates that the slope of the trace of an image of a knife is equal to the LSF.

Experimentally, this result has only limited direct application because when a knife edge image is scanned, the edge trace incorporates the LSF of both the optical system and the scanning system. Separation of the two LSF would require the difficult deconvolution process. However, the edge trace slope obtained when the scanning system traces an illuminated knife edge directly is the LSF of the scanning system.

To circumvent the necessity of the deconvolution in obtaining the experimental LSF of a holographic system, the transform function analysis technique was used. Briefly, the methodology used is as follows. A holographic system is used to obtain an image of a knife edge which is recorded on

film. By the black box analogy, the entire system yielding the experimentally usable image consists of the initial holographic diffraction process, the holographic recording and development, the reconstruction diffraction process, and, finally, the film recording of the reconstructed image. A microdensitometer is used to obtain the edge traces of both the knife edge image on the film plate and of the knife edge itself. From the edge traces, the MTF and PTF of the two systems are obtained. The MTF of the holographic system without the microdensitometer effects is obtained by dividing the MTF obtained from the image edge trace by the MTF of the knife edge trace. The PTF of the holographic system minus the microdensitometer contribution is obtained by subtracting the PTF of the knife edge trace from the PTF of the image trace. The derived MTF and PTF of the holographic system are used to obtain the real and imaginary components of the OTF. These, in turn, are used to find the inverse Fourier transform of the OTF which is, of course, the LSF.

The method for obtaining the MTF and PTF is a graphic technique described in detail by Scott, et al. (40). The principles on which it is based are described by Shack (41). Since the description of the technique is relatively involved, it will not be discussed in this text other than for the observations that it is applicable to all optical systems for which a microdensitometer edge trace may be obtained, and that it is theoretically exact.

The hologram of the knife edge was made using the off-axis recording mode. The angle between the beams was about seven degrees, and the knife edge (a razor blade) was 35 cm from the film plate which was Agfa Gevaert 10E75. The object and reference beams were collimated plane waves, and their intensity ratio was unity. In order to facilitate the correct alignment of the hologram for the reconstruction, the hologram also contained the holographic recording of a pinhole which was precisely in the same plane as the knife edge. During the reconstruction, the angular alignment of the hologram was adjusted until the reconstructed pinhole image showed the least amount of aberration. This was found to be a much more sensitive method of aligning the holograms than by using the knife edge image to judge the alignment. To insure the best focus during the recording of the reconstruction, both the recording film plate and the focal position of the real image were precisely located through a microscope. The pinhole image was used to determine the focal plane.

A Jarrell-Ash Console Comparator Microdensitometer was used to trace the film recording of the reconstructed knife edge image. The tracing slit was adjusted by the instrument dials to an indicated 2μ width and 1.2 mm length; the tracing speed was 1 mm/min. Although the optics of the microdensitometer do not require critical focusing (42), the focus was adjusted to give the best square wave response.

This adjustment was entirely judgemental and allowed for some focusing error which will be further discussed below.

Two holographic cases were investigated which differed only in the hologram size. One hologram (to be referred to as the A hologram) was 2.4 cm in diameter, this value being chosen as representative of the nominal hologram size requirements. In anticipation of a broader spread function, the other hologram (the B hologram) was 0.75 cm in diameter. In both cases, the geometrical image of the knife edge was near the center of the hologram.

Figure 29 shows the microdensitometer traces of a knife edge and the reconstructed knife edge image traces of the A and B holograms. The traces have been normalized to remove the effects of the unexposed film emulsion absorption. It is immediately apparent that none of the traces are symmetrical, thus indicating that there was phase distortion present. Some sense of the degradation due to the holographic system is obtained by noting that the slope of the reconstructed image traces is not as steep as that of the knife edge trace. Likewise, the contrast between the clear and opaque fields is not as strong. Another feature worth noting is that the slope of the traces is not always positive. By Equation 7-17, this implies that there exist regions in which the spread function has negative values.

Figure 30 shows the MTF and PTF of the microdensitometer obtained from the knife edge trace. Some characteristics pertinent to the technique of edge trace

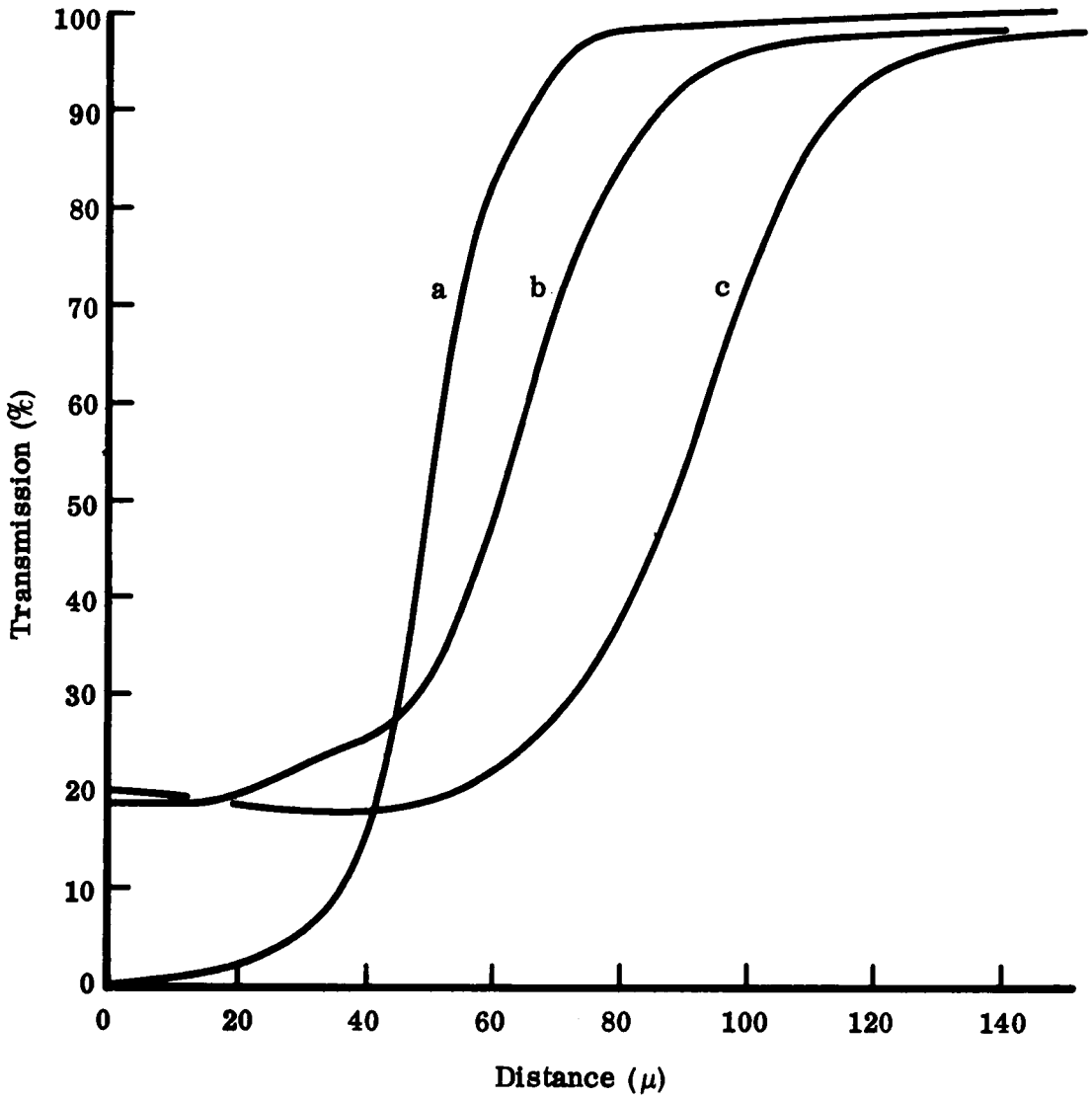


Figure 29. Microdensitometer traces of (a) a knife edge and (b) and (c) the reconstructed image of a knife edge from the A and B holograms respectively.

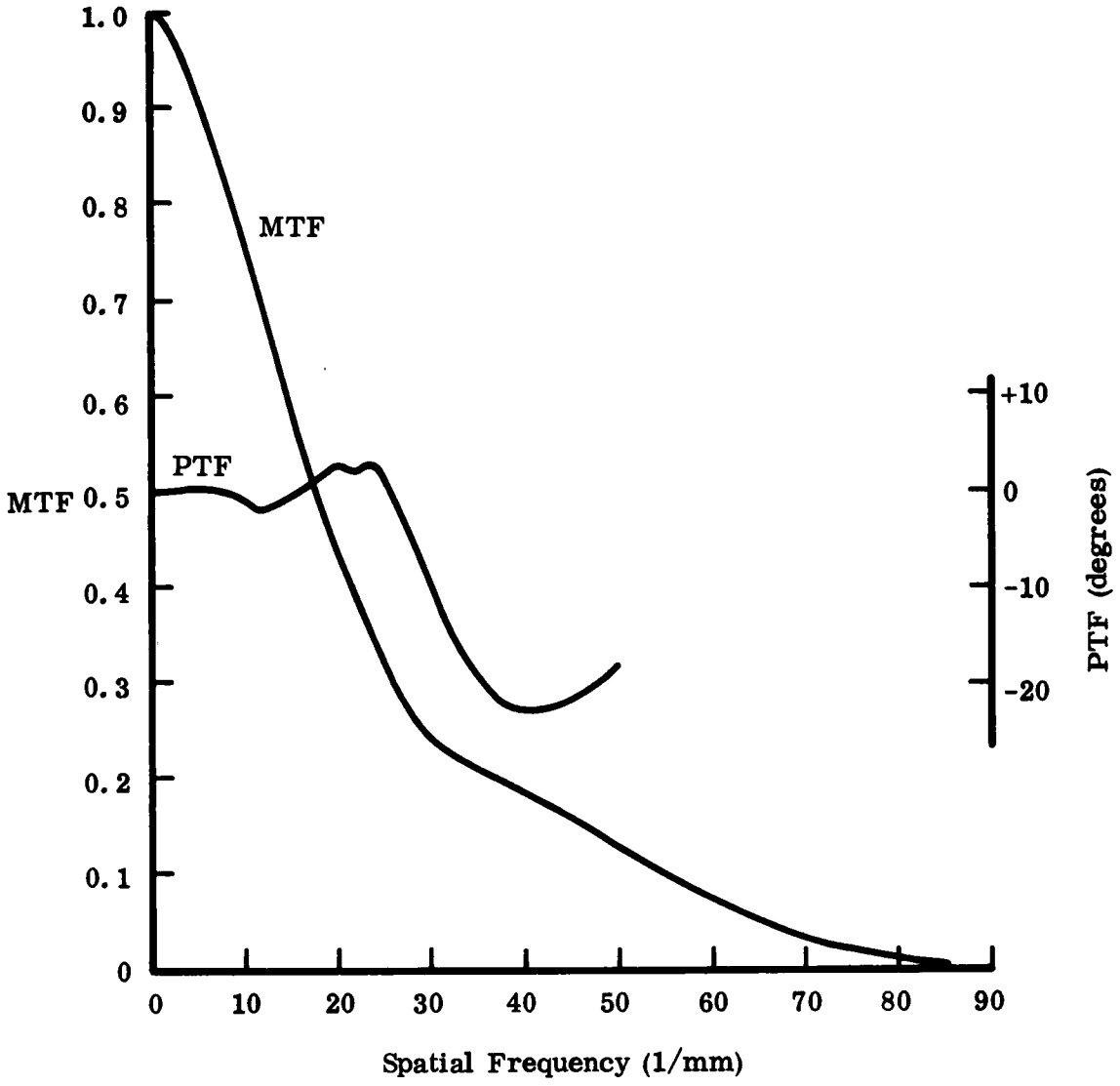


Figure 30. The modulation and phase transfer functions of the microdensitometer.

scanning can be discussed relative to these curves. The variation of the PTF significantly away from zero values indicates the presence of phase distortion in the microdensitometer system. This distortion, caused by aberrations, may be resulting from defocusing and/or aberrations of the microdensitometer optics, angular misalignment in the slit orientations, or damaged edges of the slit. The microdensitometer was not disassembled to check for physical damage; edge integrity of the slit jaws was assumed. Geometrical alignment of the slit-trace geometry is visual on the Jarrell-Ash microdensitometer and orientation better than within one degree is easily attained. The probable cause for the PTF variations is a slight defocusing of the microdensitometer optics.

The MTF decreases rather gradually so that a clear cut-off frequency is not apparent. The transfer function of a microdensitometer slit is approximately given (40) by $\mathrm{sinc}(\pi w f_s)$ where w is the effective width of the slit. Using an estimated cut-off spatial frequency of 85 to 100 (1/mm) suggests an effective slit width of 10μ to 12μ. When compared with the instrument setting of 2μ, the effective slit width implies a discrepancy that could undermine the confidence in the use of this experimental method of determining the transfer function characteristics. This problem was addressed by Jones and Coughlin (39) who found that correction for the microdensitometer scanning of photographic images by dividing the MTF of the photographic image

by the microdensitometer MTF resulted in an "overcorrection" at the higher spatial frequencies. Their approach was to develop the more exact but also more difficult technique of incorporating in the data reduction analysis a microdensitometer correction function. As part of their technique, the experimentally determined microdensitometer spatial frequency is truncated at a value corresponding to its 0.1 MTF value. Smaller values are not considered sufficiently reliable. This accounts for the apparent discrepancy discussed above.

Because of the large number of arithmetic computations involved, the use of a microdensitometer correction function cannot be used successfully without extensive computer programming. For that reason, it was not incorporated in this work. In defense of this omission, it may be pointed out that the work by Jones and Coughlin shows that inclusion of the microdensitometer correction function improves the detail of the results but does not significantly change the overall characteristics of the results obtained.

This observation is backed up by the curves of Figure 31. The two sets of data are the LSF of the microdensitometer derived from the transfer function (the solid lines) and from a direct calculation of the slope of the knife edge trace (the circles). The data was normalized for unity at the peak of the curve. The high level of correlation justifies the use of the transfer function method of

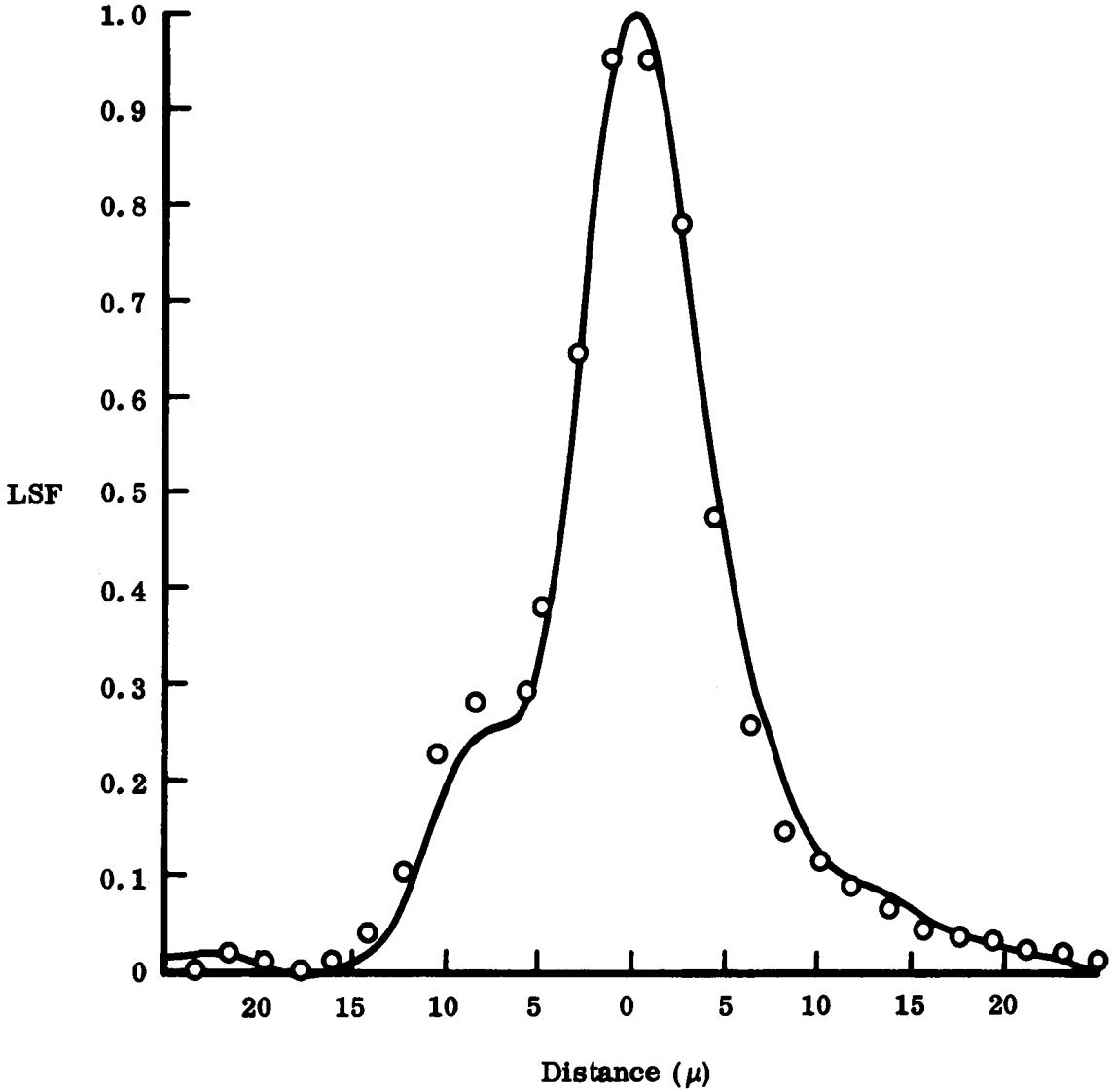


Figure 31. The normalized line spread function of the microdensitometer; the solid line is the inverse Fourier transform of the system's OTF, and the circles are a direct measurement of the slope of the knife edge trace.

obtaining the LSF and increases the confidence in the results to be obtained for the holographic systems.

Figures 32 and 33 are the MTF and PTF of the A and B holographic systems respectively. The B system shows phase distortion over a larger spatial frequency range than the A system. The cut-off frequency of the MTF curves cannot be accurately compared partly because the A system MTF was extrapolated (the dashed line portion) from the trend in the curve and partly because of the system uncertainty at higher f_s. The modulation of the A system is considerably larger past about f_s = 10 (1/mm) than the B system; the greater area under the MTF curve in the A system somewhat loosely implies a better quality of that imaging system (36).

Figures 34 and 35 are the experimental LSF minus microdensitometer effects of the holographic systems A and B respectively. The intent in calculating these experimental spread functions was to obtain a qualitative representation of actual image characteristics. As already stated, some aspects of the details can be improved by the proper inclusion of a microdensitometer correction function. Specifically, when the MTF of the image traces is divided by the uncorrected microdensitometer MTF, the resulting MTF curves are "overcorrected" showing an erroneous contribution to the MTF at the higher spatial frequencies. Likewise, the PTF at higher f_s is not reliable. When the LSF is obtained from the inverse Fourier transformation of the OTF, the erroneous contributions of the higher f_s will have the

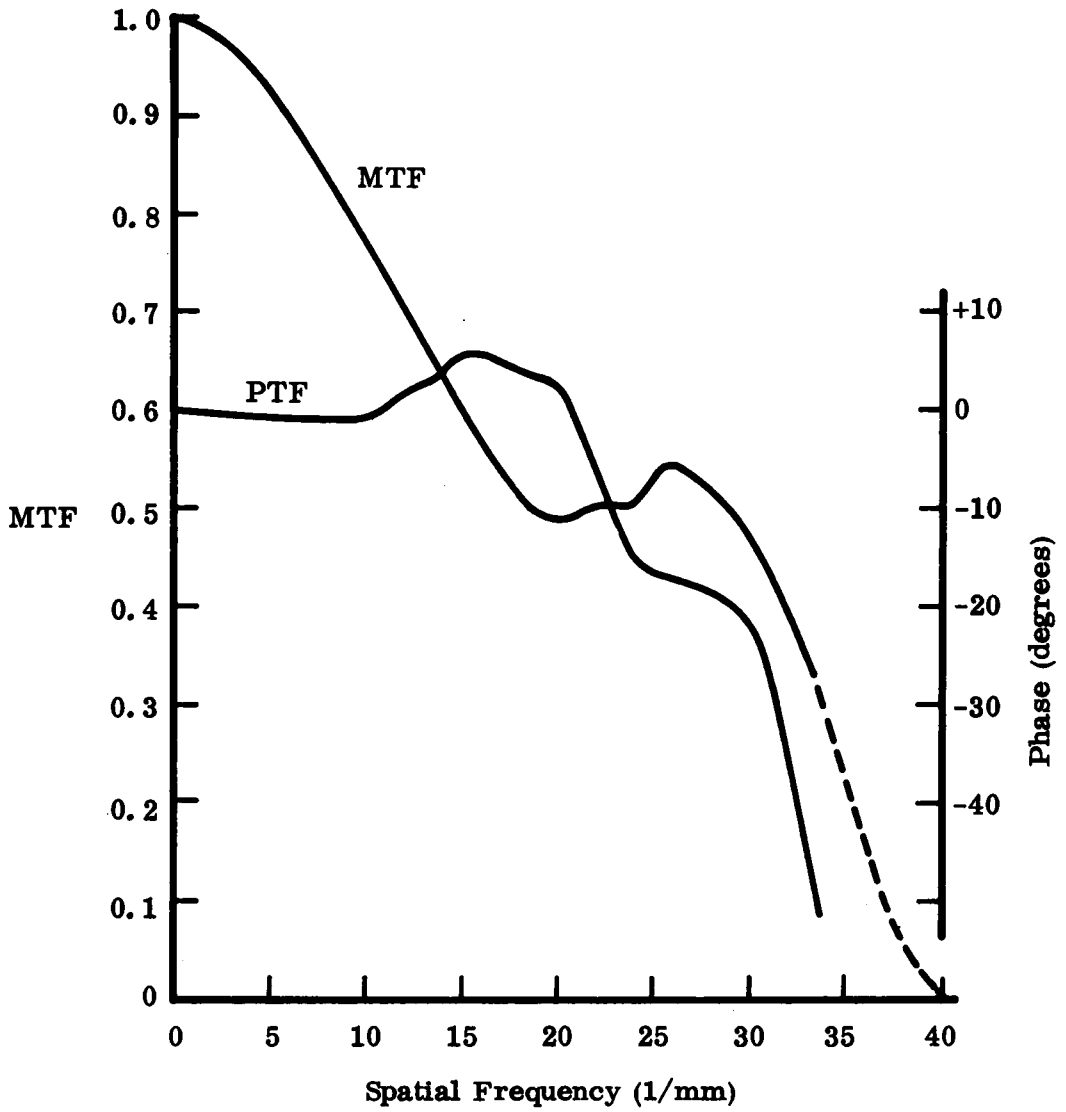


Figure 32. The modulation and phase transfer functions of the A hologram minus the microdensitometer effects.

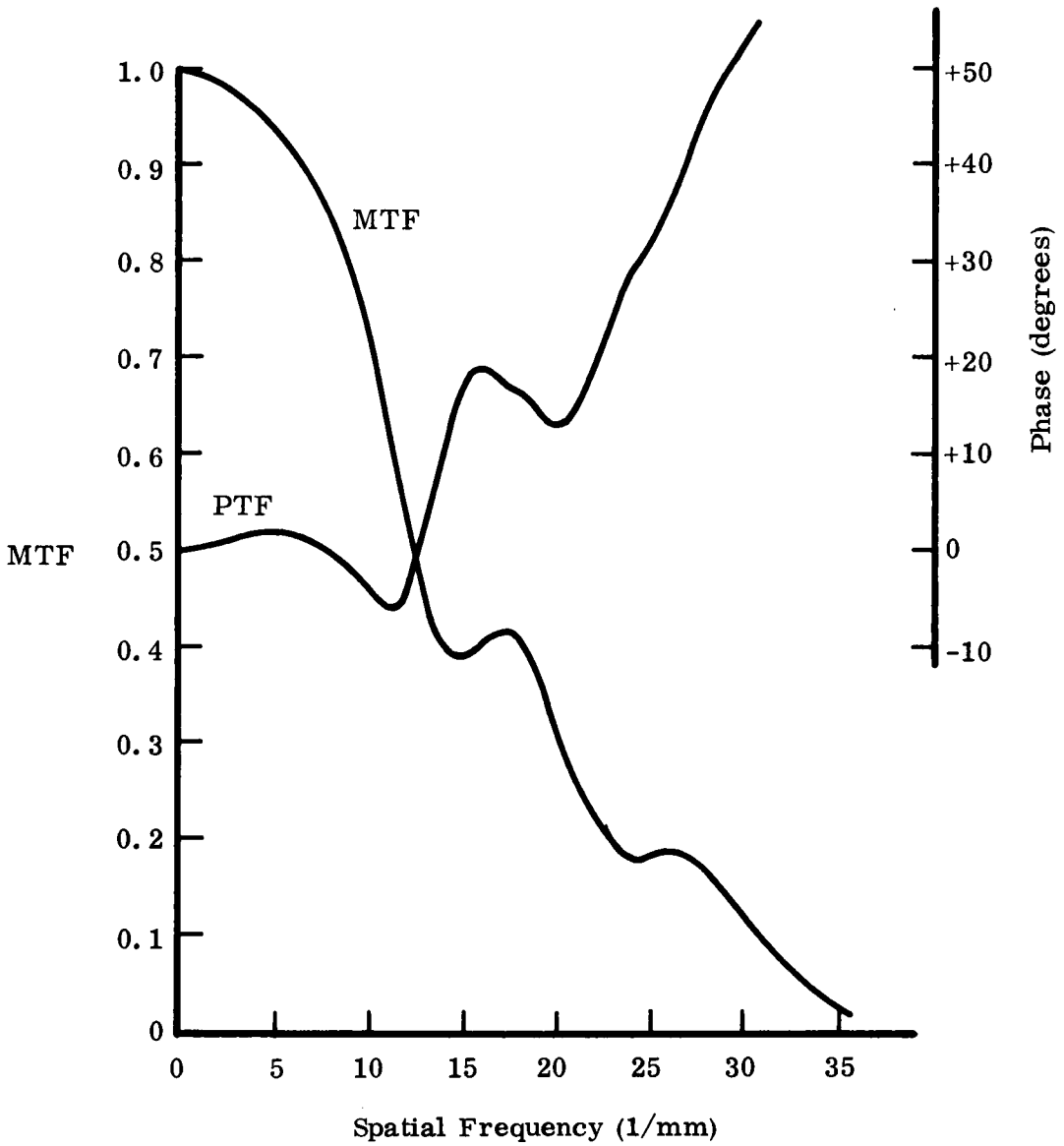


Figure 33. The modulation and phase transfer functions of the B hologram minus the microdensitometer effects.

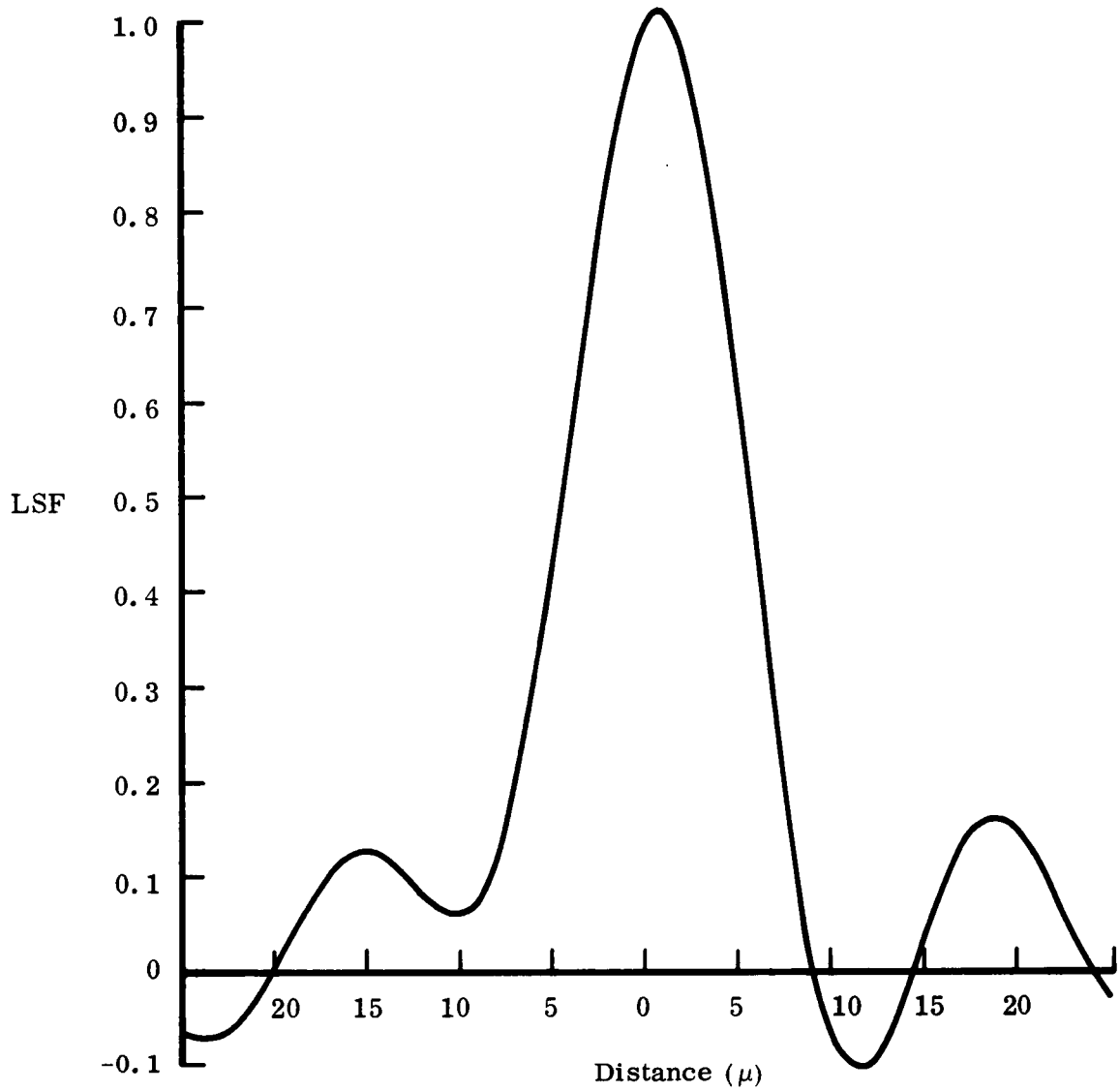


Figure 34. The normalized line spread function of the A hologram minus the microdensitometer effects.

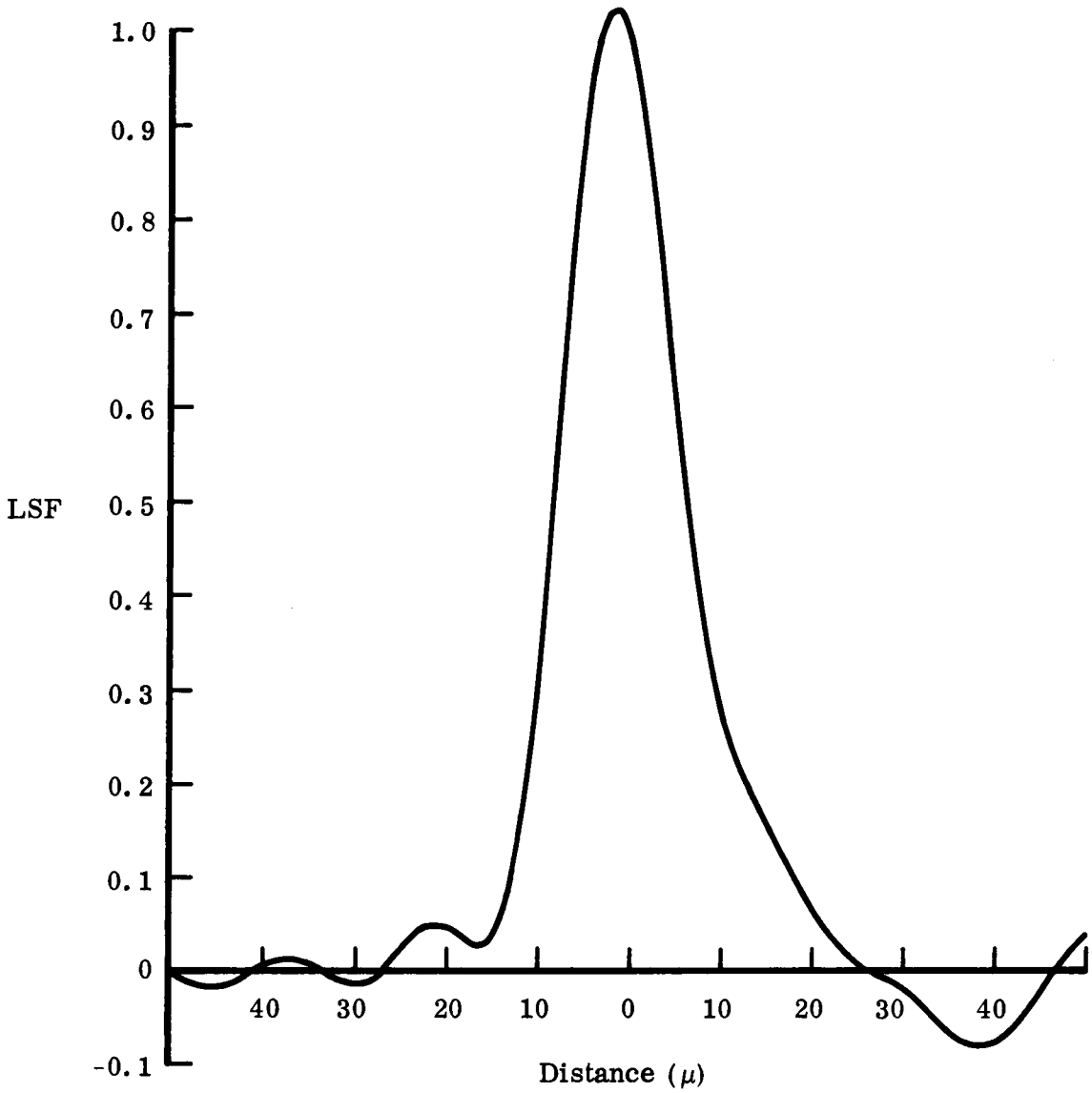


Figure 35. The normalized line spread function of the B hologram minus the microdensitometer effects.

effect of making the LSF curve width more narrow than for the correct case. Thus the LSF of Figures 34 and 35 imply an image resolving capability of the respective systems that is better than might be expected.

As stated before, the holographic system includes the film effects during both the holographic recording and while making a permanent recording of the reconstructed image. These steps introduce nonlinearities into the system because of the nonlinear nature of film recording parameters. While the nonlinearities do not contribute greatly to the reconstructed image, that is not the case when recording the image. The image field consists of the illuminated region which, in imitation of the original step function, rather abruptly goes to zero. The greatest demand on the film is in recording the transition region; yet it is precisely at the transition region that the nonlinearities are most prominent. Very low levels of illumination in the shadow region of the step function either will be below the threshold of the film sensitivity or will nevertheless be attenuated by the slow threshold response of the film's Hurter-Driffield sensitivity curves. Either way, the effect is that the recorded image shows a steeper response curve than would be expected of the holographic system (actually, any optical system) without the final recording step.

This synergistic improvement of the image is analogous to the case of lens correction through compensating aberrations. It also points out one of the limitations of

the transfer function approach to optical analysis. The cascading effect of simply multiplying the component MTF of a system to obtain the overall system MTF is not sufficiently versatile to provide descriptions of compensating effects.

Consequently, the LSF of Figures 34 and 35, pages 132 and 133, cannot be expected to correlate to the theoretical model expressed by Equation 3-26. For example, the inverse relationship which relates the spread function width with the hologram size is not demonstrated with any greater accuracy than the observation that the spread function width did increase as the hologram size decreased. Nevertheless, the line width obtained implies a resolution which is comparable to that observed visually. The half-widths of the A and B line spread functions are about 10μ and 15μ respectively; by an approximate application of the resolution criterion of Chapter VI, the systems should resolve line groups with line widths and separations of about the same size as the LSF half-widths. The line groups on the Air Force Resolution Test Targets observed from reconstruction where the hologram was apertured to the same size as holograms A and B were 6/1 and 4/2. The line widths inferred from these values are about 8μ and 28μ. Of course, during the visual observation, the resolution is degraded by the optics of the microscope and of the eye. For a comparison, for the parametric condition of the systems A and B,

Equation 3-26 predicts the width at half-height of the line spread functions to be about 5.6μ and 17.8μ respectively.

Another sense of the accuracy and actual applicability of the experimental spread function is to apply it to finding the image of an object using the convolution integral, Equation 7-3. The results of convolving the A hologram spread function with a 30μ wide step function is shown in Figure 36. The intensity distribution is remarkable for both its dimensions and its shape. At half-height the width is 30.2μ, which reflects on the accuracy of the size of the image. Also, two intensity maxima are observed within the image. For the parameters used, $(kaL/z_1) = 5.11$. By Equation 4-40 this result correctly indicates the number of intensity maxima which are observed. The negative region of the image intensity distribution in Figure 36 clearly indicates that the system has not been sufficiently corrected of experimental errors. The two main sources for error within the convolution process itself are errors within the spread function distribution and the finite width of the spread function used. By this last reason, the integration performed was not an accurate convolution because the limits of integration were not infinite.

The synergistic effect discussed above is in effect primarily for images that have sharp contrasts. Different resolving characteristics are to be expected if noise is present or if a background radiation is part of the reconstructed image as in an in-line holographic system. The

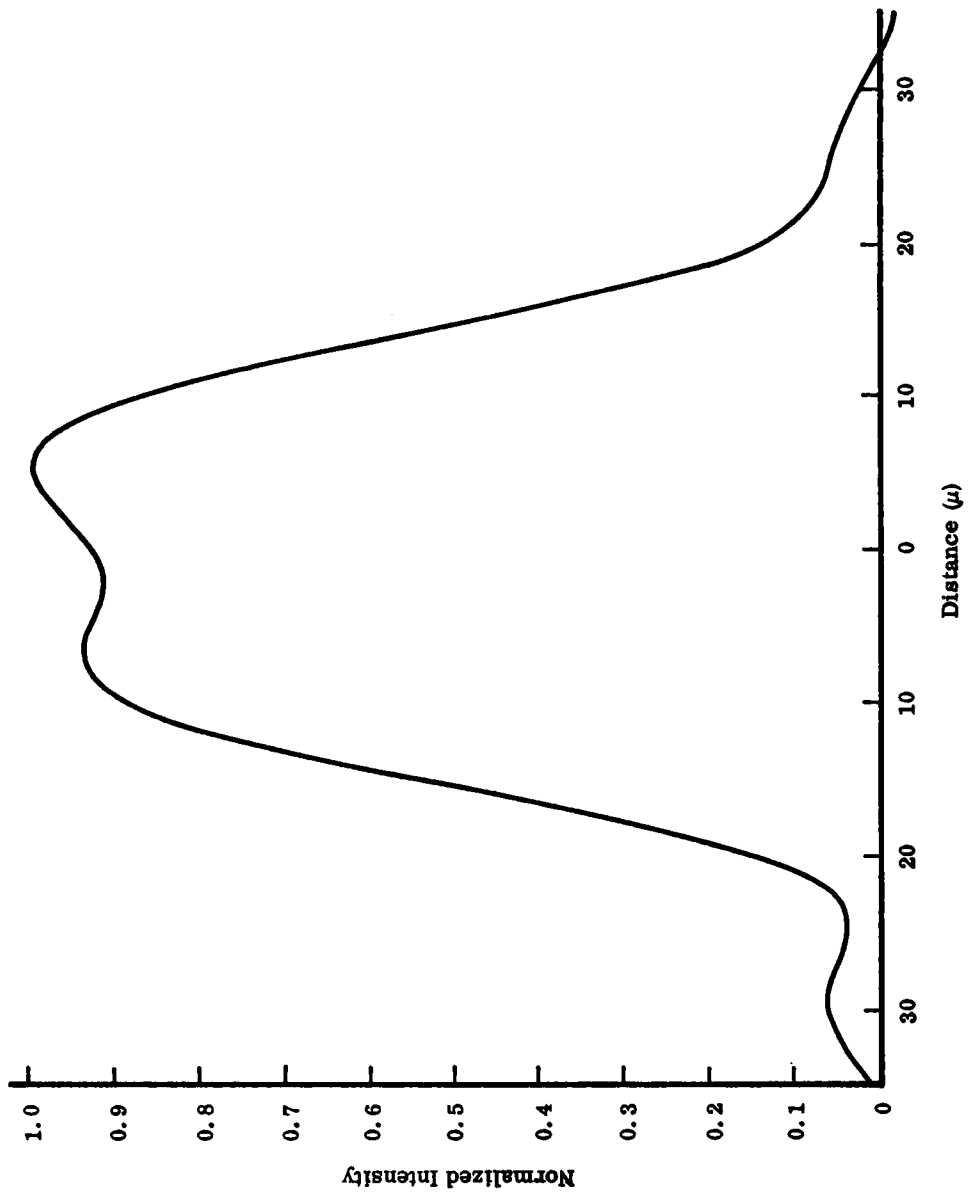


Figure 36. The image of a 30μ wide bar obtained from the convolution integral using the A hologram line spread function.

recording of the reconstructed image may then be entirely above the threshold region of the film sensitivity. In that sense the off-axis reconstruction of a high contrast object field may under suitable recording conditions have an improved resolution over a comparable in-line system.

The line spread function of an in-line holographic system was not successfully obtained. The method used for obtaining the transfer functions relies to a significant extent on a large contrast in the edge function which is traced by the microdensitometer. When the contrast is low, as in the in-line reconstruction, the uncorrected transfer function of the microdensitometer obscures the transfer function of the system being evaluated to the extent that the derived MTF indicates unrealistically large values for the higher spatial frequencies.

CHAPTER VIII

CONCLUSIONS AND DISCUSSIONS

The overall effect of this investigation is to present a compilation and synthesis of physical optics oriented to the exposition of holographic phenomena. As such, this work emphasizes the details of holographically obtained images and points out new correlations between holography and other optical imaging techniques.

Fundamental characteristics of holographic images are discussed both in terms of the most basic of the imaging forming concepts, the spread function, and by the images of a few experimentally important objects which have sufficient simplicity that an analytical derivation is partially possible. The properties which are, by experience, found to be the most important for unambiguous image recognition are the definition of details and the intensity. These, in turn, are shown to be multiparametric functions in which the effective hologram size is singled out as the parameter which changes most from case to case. The experimental conditions influencing the effective hologram size, particularly the object-to-reference beam intensity ratio, are described. Finally, by the linear systems theory, all of these effects are implicit in the spread function of the system under consideration; for that reason, some holographic system spread functions are experimentally determined.

The major contributions of this work may be enumerated as follows:

1. The solid theoretical foundations of scalar diffraction theory are emphasized, and a phenomenological description of the theory is given.
2. The spread functions of the holographic systems and the lens systems are shown to be derived from the same integral equation.
3. A mathematical description of in-line and off-axis holography shows that these two systems have similar imaging properties.
4. A description of the multiparametric nature of the holographic image characteristics is derived, and the particular importance of the hologram size is emphasized.
5. The improvements in image clarity and intensity distribution obtained by using object-to-reference beam intensity ratios greater than unity is shown both theoretically and experimentally.
6. Image resolution and hologram size are experimentally correlated and are shown to satisfy the Sparrow resolution criterion.
7. An experimentally determined line spread function is obtained, and the results are compared with the theoretical predictions.

These contributions are evolutionary in nature and, in most

cases, were obtained or suggested by concentrating on the details of the scalar diffraction theory model of the holographic process.

A few remarks concerning possible future work are in order. Improvements in the derivations of holographic reconstruction fields by including more realistic radiation fields and an improved film model would, in all likelihood, render the diffraction equations insoluble in an analytic form. To require the reference and object beam to have, for example, a Gaussian intensity envelope would necessitate better approximation techniques than are currently used. Likewise, it is desirable that a film model which is not mathematically restricted to linearity be developed by suitable approximation techniques. Finally, the mathematics of the holographic phenomena can be made more realistic by having the reconstruction diffraction model include effects which are due to both amplitude and phase variations within the hologram.

The feasibility of unambiguous image recognition in the reconstruction is strongly dependent on a foreknowledge by the observer of the image to be recognized. For the purposes of automating holographic data reduction, considerable work is still necessary to determine which criterion of reconstructed images can best be used for programming machine recognition. A variety of experimental conditions must be optimized in as many details as possible to render the data reduction to a manageable form. In this light, the

accurate geometrical alignment of the hologram during reconstruction is very important. Therefore, it can be recommended that, where possible, a recognizable object be part of the recorded holographic field so that reconstruction alignment can be made by optimizing the image of that object. The holographic recording of the radiation through a pinhole provides a good example because small misalignments cause easily observed aberrations.

The work involved in experimentally calculating the line spread function is considerable. For this reason, only a representative choice of the different holographic systems has been investigated. To obtain a thorough empirical characterization of the holographic process, it would be desirable to obtain the spread function evolution as a function of such parametric variation as tracing the image through the entire focal volume, hologram misalignment, film variations, and different recording media. The technique can also be used to evaluate holographic systems having various geometries or perhaps different noise environments. The utility of the experimental method used will depend, to a large degree, on the extent to which it is automated.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. J. D. Trolinger, and R. A. Belz, AEDC Technical Report, AEDC-TR-73-160.

2. W. M. Farmer, K. S. Burgess, and J. D. Trolinger, AEDC Technical Report, AEDC-TR-70-181.

3. R. A. Belz, and N. S. Dougherty, "In-Line Holography of Reacting Liquid Sprays," Published by SPIE in Proceedings of the Engineering Application of Holography Symposium, February, 1972.

4. A. E. Lennert, et al., "Electro-Optical Techniques for Diesel Engine Research," AEDC Technical Report, to be published.

5. D. I. Staselko, and V. A. Kosnikovskii, Opt. Spectrosc., 34, 365 (1973).

6. H. Royer, Nouv. Rev. Optique, 5, 87 (1974).

7. O. Theimer, G. D. Wasserman, and E. Wolf, Proc. Roy. Soc. A, 212, 426 (1952).

8. H. S. Green, and E. Wolf, Proc. Phys. Soc. A, 66, 1129 (1953).

9. E. Wolf, Proc. Phys. Soc., 74, 269 (1959).

10. E. Wolf, Proc. Roy. Soc. A, 253, 349 (1959).

11. B. Richards, and E. Wolf, Proc. Roy. Soc., 253, 358 (1959).

12. P. M. Morse, and H. Feshbeck, Methods of Theoretical Physics, McGraw-Hill Book Co., Inc., New York, 1953.

13. J. C. Heurtley, J. Opt. Soc. Am., 63, 1003 (1973).

14. F. T. S. Yu, Introduction to Diffraction, Information Processing and Holography, MIT Press, Cambridge, 1973.

15. J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill Book Co., Inc., San Francisco, 1968.

16. J. B. Develis, and G. O. Reynolds, Theory and Application of Holography, Addison-Wesley Publishing Co., Reading, Mass., 1967.

17. F. T. S. Yu, J. Opt. Soc. Am., 59, 490 (1969).

18. J. W. Goodman, and G. R. Knight, J. Opt. Soc. Am., 58, 1276 (1968).

19. E. B. Champagne, J. Opt. Soc. Am., 57, 51 (1967).

20. M. Born, and E. Wolf, Principles of Optics, Pergamon Press, Oxford, 1965.

21. E. H. Linfoot, and E. Wolf, Proc. Phys. Soc., B, 69, 823 (1956).

22. C. A. Taylor, and B. J. Thompson, J. Opt. Soc. Am., 48, 844 (1958).

23. F. H. Perrin, J. SMPTE, 69, 151 (1960).

24. R. A. Belz, "An Investigation of the Real Image Reconstructed by an In-Line Fraunhofer Hologram Aperture-Limited by Film Effects," Ph.D. dissertation, The University of Tennessee, Knoxville, 1971.

25. R. A. Belz, and F. M. Shofner, Appl. Opt., 11, 2215 (1972).

26. H. M. Smith, Principles of Holography, Wiley-Interscience, New York, 1969.

27. A. A. Friesem, A. Kozma, and G. F. Adams, Appl. Opt., 6, 851 (1967).

28. A. V. Alekseev-Popov, I. I. Komissarova, and G. V. Ostrovskaya, Opt. Spectrosc., 37, 656 (1974).

29. G. W. Stroke, An Introduction to Coherent Optics and Holography, Academic Press, New York, 1966.

30. K. Preston, PS and E, 11, 190 (1967).

31. E. B. Champagne, and N. G. Massey, Appl. Opt., 8, 1879 (1969).

32. Lord Rayleigh, Collected Papers, Cambridge University Press, Cambridge, 1902.

33. R. Barakat, J. Opt. Soc. Am., 52, 276 (1962).

34. R. L. Lamberts, G. C. Higgins, and R. N. Wolfe, J. Opt. Soc. Am., 48, 487 (1958).

35. R. L. Lamberts, J. Opt. Soc. Am., 48, 490 (1958).

36. F. H. Perrin, J. SMPTE, 69, 239 (1960).

37. G. C. Higgins, and F. H. Perrin, PS and E, 2, 66 (1958).

38. F. D. Smith, Appl. Opt., 2, 335 (1963).

39. R. A. Jones, and J. F. Coughlin, Appl. Opt., 5, 1411 (1966).

40. F. Scott, R. M. Scott, and R. V. Shack, PS and E, 7, 345 (1963).

41. R. V. Shack, J. Res. Natl. Bur. Std., 56, 245 (1956).

42. H. B. Beal, R. K. Brehm, and R. F. Jarrell, J. Opt. Soc. Am., 47, 306 (1957).

VITA

Reinhard Menzel was born in Germany and became a United States citizen in 1958. He received a B.S. degree in Physics from the University of Maryland in 1964. Between 1964 and 1966 he was employed as a geophysicist by the United States Coast and Geodetic Survey and spent the 1964-1965 geophysical Quiet Sun Year at Eights Station, Antarctica, recording geomagnetic data. In 1966 he entered The University of Tennessee Space Institute for graduate studies. He obtained the M.S. degree in Physics in 1968 and the Ph.D. degree in Physics in 1975. Since 1972 he has been an employee of Arnold Research Organization as a Research Engineer and his major work emphasis has been concentrated on a study of fundamental optical phenomena.