

THE MODULATED GRATING HOLOGRAM

Thesis by
David R. MacQuigg

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1975

(Submitted January 9, 1975)

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Nicholas George for his initial encouragement of my study in optics and for his helpful advice along the way.

My friend and colleague, Alan Mickelson has contributed much in his help with the experiments and computer programming and in many stimulating discussions.

I am most grateful to Jack Lindsley and Dave Haas for allowing me the use of the JPL Image Processing Laboratory facilities and especially to Fred Staudhammer for help with the programming.

The excellent typing of text and figures was done by Ruth Stratton.

The financial assistance of the National Science Foundation, IBM Corporation and Caltech was greatly appreciated.

My deepest thanks go to my wife Carol for postponement of her career while I was in school and for her love and support through hard times.

ABSTRACT

A phase and amplitude, off-axis hologram has been synthesized from three computer-generated transmission masks, using a multiple-exposure holographic recording method. Each of the masks controls one fixed-phase component of the complex hologram transmittance. The basic grating is generated optically, relieving the computer of the burden of drawing details the size of each fringe. The maximum information capacity of the computer plotting device can then be applied to the generation of the grating modulation function. By this method large digital holograms (25 mm by 25 mm) have been synthesized in dichromated gelatin. The recording method is applicable to virtually any holographic medium.

The modulated grating hologram was designed primarily for the application of spatial filtering, in which the requirement is a hologram with large dynamic range and large free spectral range. Choice of a low-noise, high-efficiency medium such as dichromated gelatin will allow exceptionally large dynamic range. Independence of the optically-generated carrier grating from the computer-generated modulation functions allows arbitrarily large free spectral range.

The performance of a holographic spatial filter will be limited ultimately by noise originating from imperfections in the holographic medium. The characteristics of this noise are analyzed, and in the case of a high diffraction efficiency hologram are shown to differ significantly from previous analyses. The dominant noise source in holograms of high diffraction efficiency will be scattering of the first-

order or imaging wave by deformations in the hologram surface or other effects of low spatial frequency. Experimental measurements in various low-noise holographic media verify these predictions.

TABLE OF CONTENTS

I	INTRODUCTION	1
II	BASIC TOPICS IN HOLOGRAPHY AND SPATIAL FILTERING	4
	2.1 Holographic Diffraction Gratings	5
	2.2 Digital Holograms	24
	2.2.1 Thin Amplitude Holograms	24
	2.2.2 Thin Phase Holograms	27
	2.2.3 Phase and Amplitude Holograms	28
	2.2.4 Space-Bandwidth Limitations	31
	2.3 Holographic Spatial Filters	35
	2.3.1 A Simple One-Lens Processor	35
	2.3.2 Deblurring Filters	38
	2.3.3 Synthesis of Deblurring Filters	44
	2.3.4 Noise Problems in Coherent Optical Systems	48
	2.4 Computer Deblurring Experiment	50
	References	56
III	THE MODULATED GRATING HOLOGRAM	61
	3.1 Basic Idea--Multiple Exposure Wavefront Synthesis	61
	3.2 System for Generation of Modulated Grating Holograms	68
	3.2.1 Electronically Stabilized Interferometer	68
	3.2.2 Mask Generation System	72
	3.2.3 Fringe Phase Control System	73
	3.2.4 Phase Accuracy of the Interferometer	75
	3.2.5 Mask Phase Errors	78
	3.2.6 Suggested Improvements in the System	81
	3.3 Computer-Generated Holograms in Dichromated Gelatin	84
	3.3.1 Noise and Diffraction Efficiency in a Modulated Grating Hologram	84
	3.3.2 A Simple Three-Phase Hologram	90

3.4	Summary	98
	References	101
IV	NOISE LIMITS ON THE DYNAMIC RANGE OF HOLOGRAPHIC SPATIAL FILTERS	103
4.1	Dynamic Range and Signal-to-Noise Ratios	105
4.2	Noise in Holographic Media	107
4.3	Measurements of Diffraction Efficiency and Noise in Several Media	116
4.4	Summary	127
	References	130
APPENDICES		
A.	Complex Wavefront Synthesis by Multiple Exposure Holography (Reprint)	132
B.	Computer Programs for Analysis of Thick Holographic Diffraction Gratings	135
C.	Computer Programs for Image Processing and Generation of Holograms	139
D.	Interferograms of Various Hologram and Mask Substrates	150
F.	Measurement of Noise and Diffraction Efficiency in Holographic Media	155
G.	Hologram Processing Schedules	164
H.	Nonlinear Distortions in Holographic Media	171
K.	Summary of the Results of Kogelnik's Coupled Wave Theory for a Phase Transmission Hologram	190
N.	A Simple One-Lens Optical Processor Using Holographic Spatial Filters	194
P.	Relation of Phase Shift to Density in an Optical Medium	198

CHAPTER I
INTRODUCTION

The generation of holograms by computer and the application of holograms as spatial filters for image deblurring have been subjects of intensive research in recent years. In this thesis I have concentrated on two aspects of these subjects, a method which I have developed for making digital holograms, and the ultimate limitations on the performance of a holographic spatial filter imposed by noise in the holographic medium.

Chapter II presents several basic topics directly related to the subjects of Chapters III and IV in order to acquaint the reader who is not a specialist in holography with these topics, and to place the research in the context of what has been done before. The first section discusses the diffraction behavior of thick holographic gratings. The discussion of holographic diffraction gratings found in most texts treats only the limiting cases of very thin and very thick holograms. Many holograms, including the ones in the experiments of this thesis, are of an intermediate thickness, and neither limit adequately describes the behavior of these gratings. For this reason I have decided to present, in some detail, the theory for holographic gratings of arbitrary thickness. The theory for this general case does not provide the number of simple analytic relationships that derive from the theory for the thick limit, but it does offer more insight into the physics of the diffraction process and provides equations which may be easily evaluated with a computer.

The section on digital holograms presents a brief review of some of the methods which have been developed to generate holograms using a computer. The space-bandwidth limitations inherent in any digital hologram are discussed.

The section on holographic spatial filters presents a brief review of the work that has been done in image deblurring with coherent optical processors, and a discussion of the analog and digital methods which have been developed for the synthesis of deblurring filters. Most of the research in this field has concentrated on achieving filters of large dynamic range. The real limitations of coherent optical processing may not be in the dynamic range of the filter, but in noise and nonlinearities in the input medium. A computer simulation of a deblurring experiment shows the effect of these imperfections.

The modulated grating hologram was designed with the spatial filtering application in mind. In principle, it offers a digital hologram with large dynamic range and large free-spectral range, which should be ideal for spatial filtering. Chapter III presents first a discussion of the basic idea and its inherent advantages over previous methods, followed by a detailed description of the experimental system to synthesize such holograms. The performance of the system is evaluated and judged to be inadequate for the synthesis of precision spatial filters. Suggestions are made for future improvements. The synthesis of digital holograms in dichromated gelatin is demonstrated.

Chapter IV discusses the ultimate limits on the performance of holographic spatial filters, assuming all problems such as noise and

nonlinearities in the input medium and optical system noise have been eliminated. The ultimate limit is assumed to arise from noise in the holographic medium. The characteristics of this noise are analyzed and, in the case of a high diffraction efficiency hologram, are shown to differ significantly from previous analyses. The dominant noise source for high efficiency holograms will often be scattering from the imaging wave, an effect which has not previously been analyzed. Experimental measurements in various low-noise holographic media verify this effect.

I have tried to make this thesis a thorough, yet readable presentation of research. In many places in the main text, I have simply stated results and left the derivations to appendices. Included in the appendices are details of derivations, details of experiments, and several topics which turned out to be peripheral to the main subject of the thesis.

The thesis is a little longer than initially planned, primarily because of the inclusion of some basic topics for the benefit of the reader who is not a specialist in holography, and the inclusion of detailed appendices, which may be useful for the continuation of this research.

The synthesis of computer-generated holograms, and spatial filters in particular, is an area of research I believe worthy of further study. Limitations of time and resources in this study have prevented the synthesis of a precision spatial filter using a modulated grating hologram. It is my hope that this thesis will provide a good foundation for further development of these holograms.

CHAPTER II

BASIC TOPICS IN HOLOGRAPHY AND SPATIAL FILTERING

In this chapter several topics are presented which are basic to an understanding of the research presented in later chapters. Before moving on to these topics a brief discussion of units and notation is in order.

The electric and magnetic fields in this thesis are assumed to be time harmonic:

$$\begin{aligned}\bar{E}(\bar{r},t) &= \text{Re}\{\bar{E}_0(\bar{r}) e^{-i\omega t}\} \\ \bar{H}(\bar{r},t) &= \text{Re}\{\bar{H}_0(\bar{r}) e^{-i\omega t}\}\end{aligned}\quad (2.1)$$

\bar{E}_0 and \bar{H}_0 are complex vector functions of position \bar{r} . Holographic media, being dependent on photo-ionization and excitation mechanisms, respond to the time-averaged energy density in the electric field,⁽¹⁾

$$\langle w_e \rangle = \frac{1}{2} \epsilon \langle \bar{E} \cdot \bar{E} \rangle = \frac{1}{4} \epsilon \bar{E}_0 \cdot \bar{E}_0^* \quad (2.2)$$

Sensitivity data and other related experimental measurements are commonly given in terms of exposure $\mathcal{E}(x,y)$, (e.g., millijoules/cm²), rather than electric or magnetic fields. In this thesis the optical field amplitude at a point (x,y) in a given plane is represented by a complex scalar function, $\tilde{a}(x,y) = a(x,y) e^{i\phi(x,y)}$, with the property that

$$\mathcal{E}(x,y) = \tau |\tilde{a}(x,y)|^2 \quad \tau: \text{exposure time} \quad (2.3)$$

and ϕ is the phase of the electric field. This representation is valid when the electric field is linearly polarized and perpendicular to the plane of incidence defined by the two waves exposing the hologram. In

situations of more complex polarization, the vectorial nature of the field must be considered. For a simple, linearly polarized plane wave

$$\tilde{a}(x,y) = \left[\frac{1}{2}(\epsilon/\mu)^{1/2} \right]^{1/2} \tilde{E}_0(x,y) \quad (2.4)$$

2.1 Holographic Diffraction Gratings

The recording and reconstruction processes in most holograms can be analyzed in terms of elementary periodic gratings. If the recording wavefronts are sufficiently smooth, ray directions may be defined at each locality of the hologram, as shown in Figure 1. Interference fringes are formed along planes which bisect the angle between the two rays. These fringes are recorded as a periodic variation of the refractive index or absorption constant of the medium.

If the object wavefronts have a large angular bandwidth, however, the ray directions may vary considerably over several fringe periods,

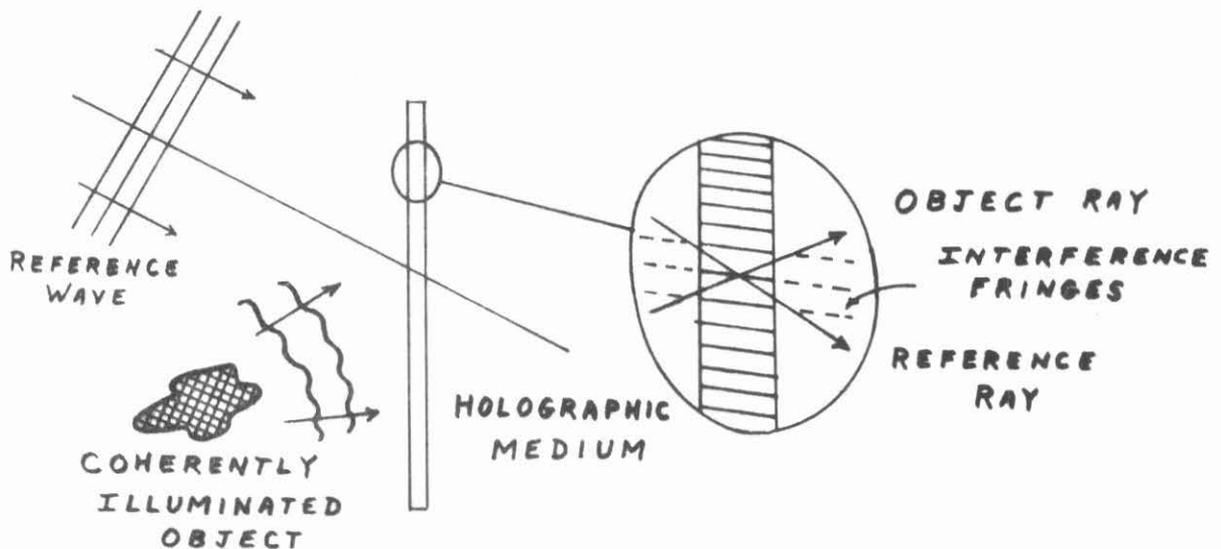


Fig. 2.1 Illustration of the formation of an elementary periodic grating at each point in the hologram.

and the grating may therefore have no local periodicity.

An alternative view, which is valid also for large angular bandwidths, is to consider the hologram as a linear superposition of elementary periodic gratings, one for each plane wave component of the incident object wavefront. This view requires that the hologram exposure be sufficiently low that the linear range of the recording medium is not exceeded.

In the extreme situation of large exposure, large angular bandwidth holograms, the analysis of the hologram in terms of periodic gratings will be very difficult. The recorded pattern in this situation will have no local periodicity, and the fringes may not even be continuous through the finite thickness of the medium. Since most holograms are not of this extreme category, an understanding of the diffraction behavior of periodic gratings will be sufficient to predict the behavior of more complex holograms in most situations.

In the discussion to follow, we will consider the hologram as an infinite periodic grating, with fringes perpendicular to the surface as shown in Figure 2. An incident wave of unit amplitude forms an angle θ with the z axis. The amplitude of this wave at the xy plane is given by

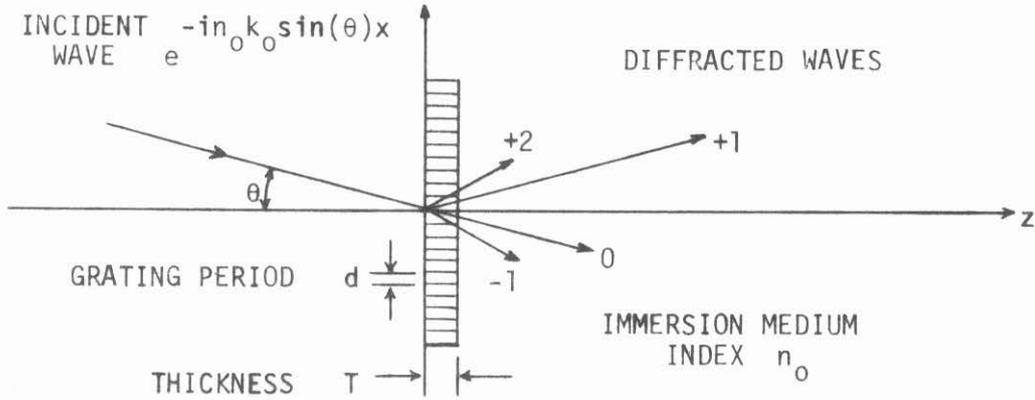
$$\tilde{a}(x,y) = e^{-in_0 k_0 \sin(\theta)x} \quad (2.5)$$

$$k_0 = 2\pi/\lambda_a$$

$$\lambda_a = \text{wavelength in air}$$

$$n_0 = \text{refractive index of external medium}$$

Fig. 2.2 Diffraction of an incident wave into various orders by a periodic grating of finite thickness.



In order to make the analysis simple, the refractive index of the immersion medium, n_0 , will be the same as the average refractive index of the grating. This assumption will eliminate multiple reflections and changes in ray directions at interfaces, but will preserve the essential physics of the diffraction process.

The simplest model considers the hologram thickness T to be much smaller than the fringe period d . For a simple sinusoidal absorption grating the amplitude transmittance may be written as

$$t(x) = t_{\max} \left(\frac{1}{2} + \frac{1}{2} m \cos K_g x \right) \quad (2.6)$$

$$K_g = 2\pi/d$$

The modulation parameter m may vary from zero to one. For a unit amplitude incident wave, as shown in Figure 2, the transmitted field

will consist of only three diffracted waves:

$$e^{-in_0 k_0 \sin(\theta)x} \left[\frac{t_{\max}}{2} + \frac{mt_{\max}}{4} e^{iK_g x} + \frac{mt_{\max}}{4} e^{-iK_g x} \right] \quad (2.7)$$

For $t_{\max} = m = 1$, the maximum amplitude of the first order wave is 1/4. The diffraction efficiency is defined as the square of the relative amplitude, and for a thin, sinusoidal, amplitude hologram:

$$\eta_{\max} = 1/16 = 6.25\% \quad (2.8)$$

For a thin sinusoidal phase grating, the transmittance may be modeled as a pure phase shift:

$$t(x) = e^{i\phi \cos(K_g x)}$$

$$\phi = k_0 n_1 T / \cos(\theta) \quad (2.9)$$

The grating medium in this case is assumed to have no absorption, and a refractive index given by

$$n(x) = n_0 + n_1 \cos(K_g x) \quad (2.10)$$

The amplitudes of the diffracted orders may be expressed as Bessel functions

$$e^{i\phi \cos(K_g x)} = \sum_{\ell} i^{\ell} J_{\ell}(\phi) e^{i\ell K_g x}$$

This well known formula may be derived from the generating function for Bessel functions⁽²⁾ or from an integral representation of the Bessel functions.⁽³⁾

The diffraction efficiency for the first order wave in this case is

$$\eta_1(\phi) = J_1^2(\phi)$$

η_1 has a maximum of 33.9% at $\phi = 1.84$ radians.

The thin grating models are adequate to explain the basic holographic recording and wavefront reconstruction principles and many of the properties of specific media. Thin hologram models have been used successfully to analyze the effects of noise⁽⁴⁾ and nonlinearities⁽⁵⁾ in holographic recording. These simple models fail to explain, however, the high diffraction efficiencies and observed sensitivity of thick gratings to small changes in incident wave angle⁽⁶⁾.

The diffraction of light by a thick holographic grating is similar to the diffraction of light by acoustical waves, a problem which has been studied extensively for many years.^(7,8,9)

Holographic gratings differ from acoustooptic devices mainly in the fact that in holographic gratings absorption may play a prominent role in the diffraction process, and the holographic fringes are not necessarily normal to the surface of the grating. Studies of holographic gratings include those of Burckhardt^(10,11) who derived a solution for a nonabsorbing sinusoidal grating. Burckhardt's solution uses a matrix eigenvector method to solve a finite set of coupled differential equations. Boundary conditions are applied at both sides of the medium, allowing for multiple reflections within the grating. Unfortunately these details tend to obscure the essential simplicity of the method and make the resulting numerical results difficult to interpret.

Kogelnik⁽¹²⁾ has derived some very general solutions for the case of very thick holograms, in which only two coupled waves are present. This reduction in the number of coupled equations allows the derivation of several simple analytic relations which would never be apparent from inspection of numerical calculations. Kogelnik's solutions apply to sinusoidal gratings which may have any degree of slant (including the case of reflection holograms) and may be phase type or absorbing type holograms. For the phase transmission hologram of Figure 2 (with T very large), Kogelnik's theory would predict a first order diffraction efficiency (see Appendix K)

$$\eta_1 = \sin^2(v^2 + \xi^2)^{1/2} / (1 + \xi^2/v^2) \quad (2.11)$$

$$v = \frac{\pi n_1 T}{\lambda_a \cos \theta} = \phi/2$$

$$\xi = \pi \delta / \delta_0 \quad (\delta_0 = d/T)$$

$$\delta = \theta - \theta_B \ll 1 \text{ radian}$$

$$\theta_B = \arcsin(\lambda_a / 2n_0 d) \quad \text{Bragg angle} \quad (2.12)$$

The parameter v represents the strength of the wave coupling constant in this two-wave theory, and ξ the deviation from the Bragg angle θ_B . At the Bragg angle ($\delta = 0$) the equation reduces to

$$\eta_1 = \sin^2(\phi/2) \quad (2.13)$$

which is the form first derived by Phariseau. For weak holograms ($\phi \ll 1$) the diffraction efficiency will be a simple function of θ ,

with a maximum at θ_B and the first null at $\delta = \delta_0$.

An extension of Burckhardt's solution to gratings of arbitrary profile (nonsinusoidal) and arbitrary absorption has been presented by Kaspar.⁽¹³⁾ Numerical results are compared with Kogelnik's theory in this paper, and good agreement is shown for thick gratings.

Holograms in a real experimental situation often exhibit behavior somewhere between the limiting cases of thin and thick gratings. This is the case for the holograms used in the experiments of this thesis. For this reason it is worthwhile to explore this intermediate region in more detail. The equations for an arbitrary thickness grating may be programmed and endless solutions plotted for various specific cases. It is possible, however, to reduce these solutions to a few convenient parameters in order that the results be generally applicable to gratings in any range.

A numerical solution for an unslanted dielectric grating was implemented using Burckhardt's matrix eigenvector method. The essentials of this method are most clearly presented in a paper by Fillmore and Tynan.⁽¹⁴⁾ The derivation here follows a similar line of reasoning.

A wave equation for the electric field in the holographic medium of Figure 2 can be derived from Maxwell's equations:

$$\begin{aligned} (\nabla^2 + k^2) \vec{E}(x,z) &= -\nabla[(1/\epsilon) \vec{E} \cdot \nabla\epsilon] \\ k^2 &= k_0^2 \tilde{\epsilon}_r \end{aligned} \tag{2.14}$$

Variations in both refractive index and absorption constant are included in the complex relative dielectric constant $\tilde{\epsilon}_r$. Assuming the grating is periodic in x , this constant can be written as a Fourier series:

$$\tilde{\epsilon}_r(x) = \sum_m \tilde{\epsilon}_{rm} e^{imK_g x} \quad (2.15)$$

For fields polarized in the y direction, the term $\vec{E} \cdot \nabla \epsilon$ is zero, and equation 2.14 reduces to a simple scalar Helmholtz equation:

$$(\nabla^2 + k^2) E_y(x, z) = 0 \quad (2.16)$$

We will assume a solution of the form

$$E_y(x, z) = e^{-in_0 k_0 \sin(\theta)x} \sum_{\ell=-\infty}^{\infty} V_{\ell}(z) e^{i\ell K_g x} \quad (2.17)$$

This assumption can be justified by considering the fields entering and leaving each thin section Δz of the grating. If the fields entering each section consist of a set of discrete spatial frequencies separated by multiples of the grating frequency, then the transmitted fields will contain only those same spatial frequencies. Only the relative amplitudes of the component waves $V_{\ell}(z)$ change as the waves propagate through the medium. A more rigorous mathematical solution can be obtained by solving the differential equation 2.16 using separation of variables. (10,15)

Equations 2.15, 2.16, and 2.17 may be combined to give:

$$(\nabla^2 + k_0^2 \sum_m \tilde{\epsilon}_{rm} e^{imK_g x}) \sum_{\ell} V_{\ell}(z) e^{i\alpha_{\ell} x} = 0 \quad (2.18)$$

$$\alpha_{\ell} = \ell K_g - n_0 k_0 \sin \theta$$

If we carry out the indicated differential operations and combine terms of the same spatial frequency, equation 2.18 can be written as:

$$\sum_{\ell} [(d^2/dz^2) V_{\ell}(z) - \alpha_{\ell}^2 V_{\ell}(z) + k_0^2 \sum_m \tilde{\epsilon}_{rm} V_{\ell-m}(z)] e^{i\alpha_{\ell}x} = 0 \quad (2.19)$$

The coefficient of each $e^{i\alpha_{\ell}x}$ must be zero for this equation to be true. This gives an infinite set of coupled differential equations for $V_{\ell}(z)$. Each wave ℓ is coupled to its nearest neighbors $\ell-1, \ell+1$ by $\tilde{\epsilon}_{r,\pm 1}$, the coefficients of the fundamental frequency of the periodic grating, and to its neighbors $\ell \pm m$ by coefficients of the m^{th} harmonic.

The set of coupled differential equations in equation 2.19 represents an exact solution for a periodic grating of arbitrary thickness and can even be extended to slanted gratings by allowing $\tilde{\epsilon}_{rm}$ to be a function of z . These equations may be truncated to represent a finite set of waves and solved by a straightforward numerical integration.⁽¹⁶⁾ A sufficient number of equations must be retained in the analysis to include all waves in which there is any significant fraction of the incident power. Usually, most of the power will be diffracted into orders clustered about the central, or zero order, and very few equations are needed.

A much faster procedure for solution of the coupled wave equations is the matrix eigenvector method. The truncated set of equations is written as a matrix differential equation:

$$(d^2/dz^2) V_{\ell}(z) + \sum_p A_{\ell p} V_p(z) = 0 \quad (2.20)$$

$$\ell = -L, \dots, L$$

$$A_{\ell p} = (k_0^2 \tilde{\epsilon}_{r0} - \alpha_\ell^2) \delta_{\ell p} + k_0^2 \sum_{m=1}^M \tilde{\epsilon}_{rm} (\delta_{\ell-m, p} + \delta_{\ell+m, p}) \quad (2.21)$$

$$\delta_{\ell p} = \begin{cases} 1 & \ell = p \\ 0 & \ell \neq p \end{cases}$$

In equation 2.21 we have assumed $\tilde{\epsilon}_{rm} = \tilde{\epsilon}_{r, -m}$. This will allow a simpler solution for the matrix eigenvectors. Since the grating was presumably formed by the interference of two plane waves, we may choose the origin of the coordinate system so that the exposure distribution, and hence the resulting grating, has even symmetry in x .

The matrix differential equation can be solved if we can find a complete set of eigenvectors $U_\ell^{(j)}$ such that

$$A_{\ell p} U_p^{(j)} = \gamma_j U_\ell^{(j)} \quad (2.22)$$

$$\ell = -L, \dots, L$$

These eigenvectors represent steady state solutions of the form

$$V_\ell^{(j)}(z) = U_\ell^{(j)} e^{\pm i\sqrt{\gamma_j} z} \quad (2.23)$$

The eigenvalue γ_j will be positive for propagating solutions and negative for exponentially decaying solutions (evanescent waves).

The complete solution for waves propagating to the right is then:

$$V_\ell(z) = \sum_{j=-L}^L c_j U_\ell^{(j)} e^{i\sqrt{\gamma_j} z} \quad (2.24)$$

The $2L+1$ coefficients c_j are determined by the conditions at the left boundary of the medium.

$$V_{\ell}(0) = \sum_{j=-L}^L c_j U_{\ell}^{(j)} = B_{\ell}$$

$$c_j = [U_{\ell j}]^{-1} B_{\ell} \quad (2.25)$$

Inclusion of boundary conditions at the right side of the medium will allow solution for the left traveling waves also.⁽¹⁰⁾

Equation 2.24 is an exact solution of the scalar diffraction problem (assuming only the validity of the truncation to a finite number of orders). Solutions for the alternate polarization are given by Burckhardt⁽¹⁰⁾ and by Kogelnik.⁽¹²⁾

A FORTRAN subroutine has been written using the above methods to solve for the diffracted wave amplitudes of an arbitrary thick phase grating. (WAVES subroutine--Appendix B) Given the incident wavelength and angle, grating period d , and thickness T , and the Fourier coefficients of the grating profile ϵ_{rm} , the subroutine returns the complex amplitudes V_{ℓ} of the diffracted waves. The restriction to non-absorbing gratings is necessary only to simplify the calculation of eigenvectors. For phase only gratings, the matrix will be real-symmetric, and the Householder method may be used to find the eigenvectors.⁽¹⁷⁾ Use of a more general routine at this step would allow consideration of absorptive gratings as well.⁽¹⁸⁾

Figures 3 through 5 show some results of computations using the WAVES subroutine. The gratings for these calculations were assumed to be purely sinusoidal phase gratings:

$$\epsilon_r(x) = \epsilon_{r0} + \epsilon_{r1} 2 \cos(K_g x) \quad (2.26)$$

$$\epsilon_{rm} = 0 \quad \text{for } m > 1$$

From the relation $n^2 = \epsilon_r$ and the assumption $\epsilon_{r1} \ll \epsilon_{r0}$, relations between the n's of equation 2.10 and ϵ 's of equation 2.26 are easily derived:

$$\begin{aligned} \epsilon_{r0} &= n_0^2 \quad \text{for index matching} \\ \epsilon_{r1} &= n_0 n_1 \end{aligned} \quad (2.27)$$

The parameter ϕ , defined for a thin grating in equation 2.9, may be similarly defined for the general case even though the physical reasoning leading to the definition of ϕ as a phase shift does not apply for thick gratings:

$$\phi = \frac{2\pi T}{n_0 \lambda_a \cos \theta} \epsilon_{r1} \quad (2.28)$$

A very thin grating (Figure 3) follows the predicted Bessel function behavior well past the first peak in the diffraction efficiency η_1 . The almost complete absence of Bragg effects is seen in the overlapping of the plotted points for waves of equal plus and minus orders.

For a very thick grating (Figure 4) the calculated points follow very nicely the sine waves predicted by Kogelnik's theory.⁽¹²⁾ The number of diffracted orders retained in the calculation, NL, was set to 2 for this plot. Similar calculations with NL = 12 showed the same results for power in the 0 and 1 orders, and almost no power in any other orders.

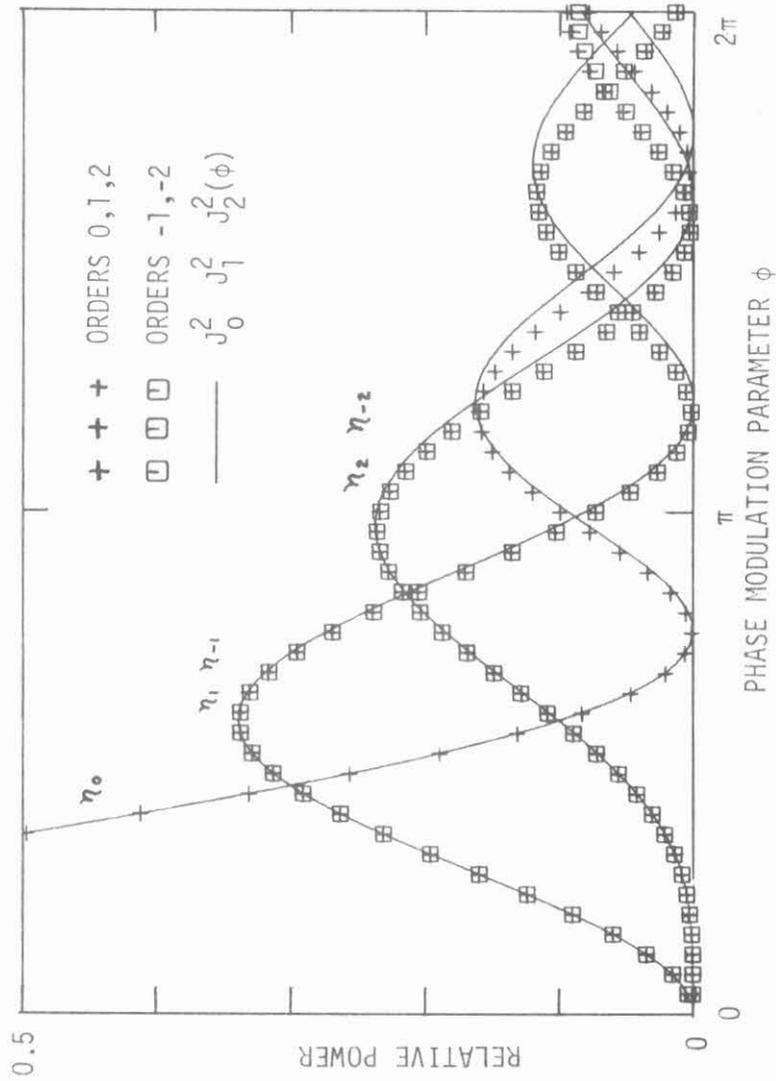


Fig. 2.3 Calculated power in the diffracted waves of a thin phase grating compared to Bessel functions. $\eta_0 = 1.54, \lambda_a = 0.5, d = 10, T = 1, NL = 12, Q = 0.02$.

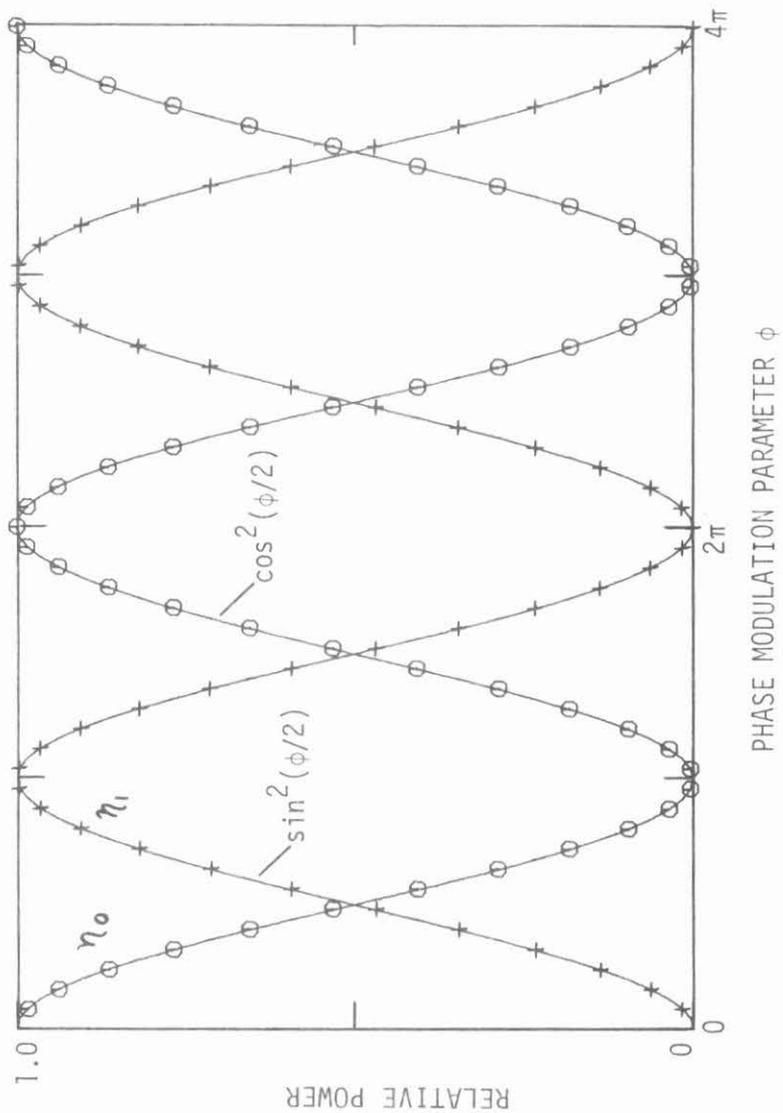


Fig. 2.4 Calculated power in the diffracted waves of a very thick phase grating compared to sine functions. $n_0 = 1.54$, $\lambda_a = 0.5$, $d = 2$, $T = 100$, $NL = 2$, $Q = 51$.

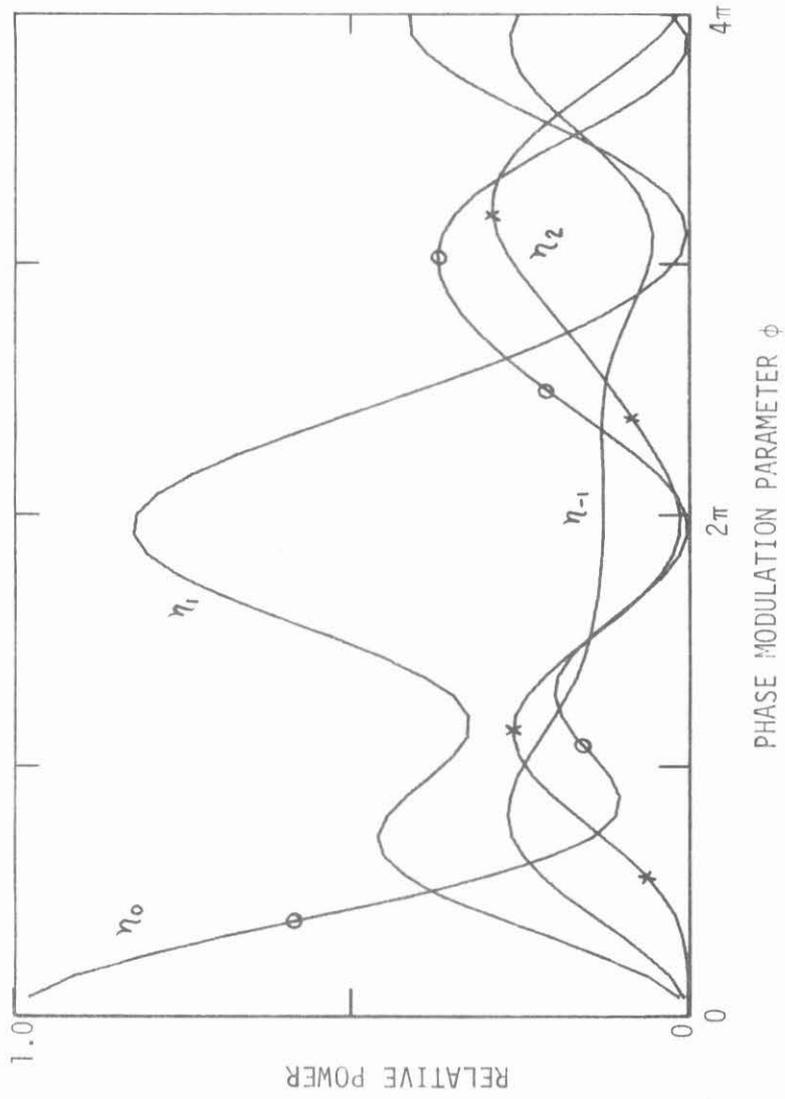


Fig. 2.5 Calculated power in the waves diffracted from a typical grating of intermediate thickness, $\eta_0 = 1.54$, $\lambda_a = 0.488$, $d = 2.86$, $T = 12$, $NL = 12$, $Q = 2.9$.

Figure 5 shows the predicted behavior for gratings with parameters typical of the holograms used in the experiments of Chapters 3 and 4. These gratings were formed in gelatin layers from Kodak 649F plates ($T = 12 \mu\text{m}$), using an argon laser ($\lambda_a = 0.488 \mu\text{m}$) with an interbeam angle of 9.8° (in air). This results in a grating with a period $2.86 \mu\text{m}$, and a Bragg angle 3.18° (internal to the medium, index $n_o = 1.54$).

The self-consistency of these calculations was checked by monitoring the power in each diffracted order and the total power of all the waves. As the modulation parameter ϕ is increased on the thinner gratings, power spreads into higher orders. Eventually a significant fraction of power reaches the edge orders and "spills over". The total power at this point drops abruptly from 1.0 and the calculations become very erratic.

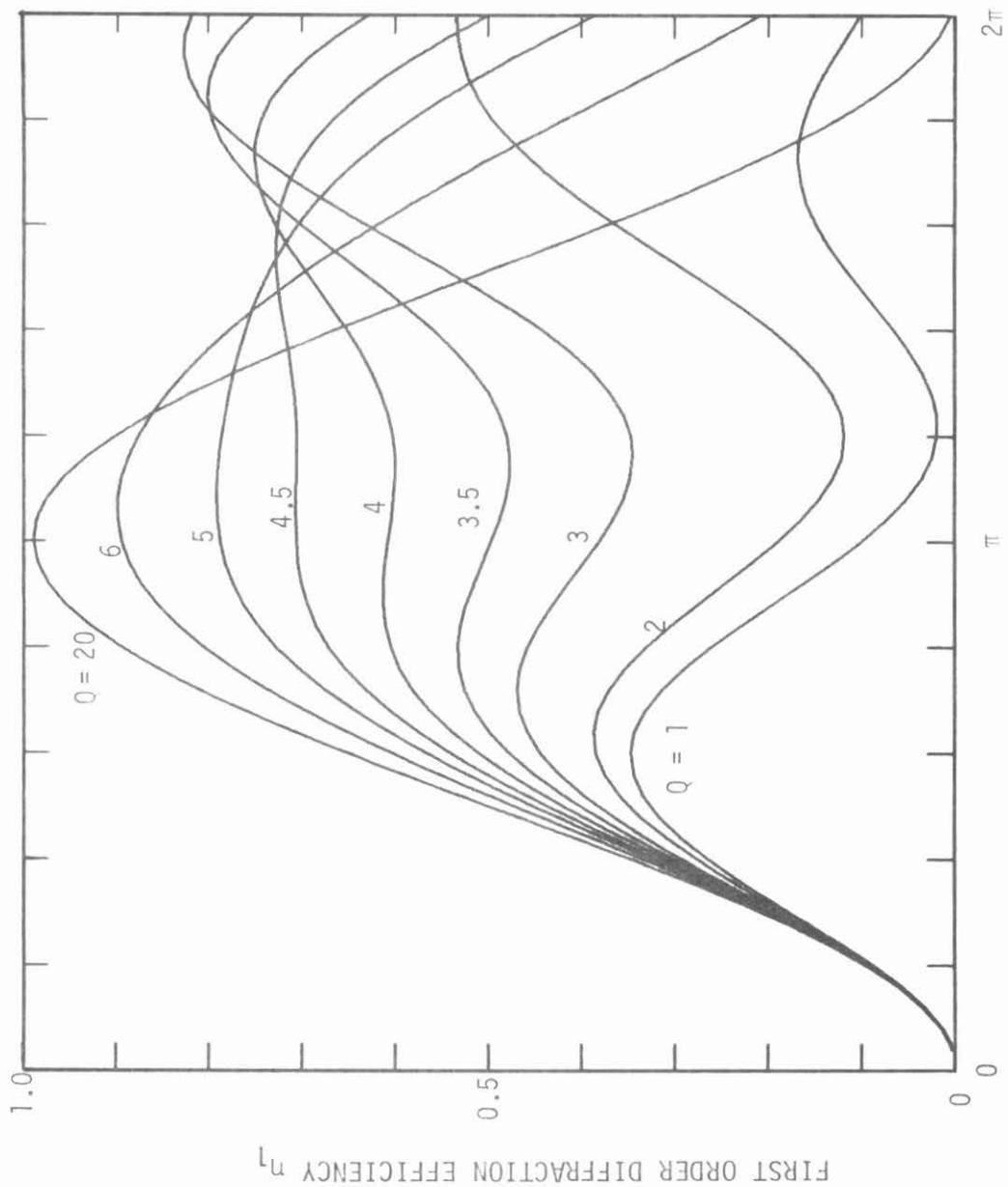
A parameter which is commonly used to characterize holographic gratings is

$$Q = 2\pi \lambda_a T/n_o d^2 \quad (2.29)$$

For $Q > 10$, Kogelnik's theory is said to give good results.⁽¹²⁾ The plots of Figures 3 through 5 are consistent with these observations.

For application as holograms, the quantity of most interest in these plots is η_1 , the first-order diffraction efficiency. Figure 6 shows a plot of η_1 vs. ϕ for various values of the thickness parameter Q . This set of curves was generated assuming gratings of fixed thickness, $T = 12\mu$. A similar set was generated setting $d = 2.5\mu$ and varying T . The two sets of curves were found to coincide except for thin gratings (low Q) at very large modulations ($\phi \approx 2\pi$).

Although it is apparent, in the theory on which these calculations are based, that the set of curves could be characterized by two parameters, d and T (normalized to unit wavelength), the reduction to one parameter Q is not at all obvious. The parameter Q appears in several theories of ultrasonic diffraction.^(19,20) Typically, in these theories, approximations are made to reduce the coupled second-order differential equations to first-order equations. The effect of these neglected second order terms would explain the discrepancies observed for thin gratings of very high modulation. For most holographic applications, it appears that Q and ϕ are quite adequate to completely characterize the grating behavior (at Bragg angle incidence). To illustrate this point, the diffraction efficiencies at the first maximum are plotted in Figure 7 for gratings in which T and d vary independently over a wide range. A very similar curve has been derived from an ultrasonic diffraction theory.⁽²¹⁾



PHASE MODULATION PARAMETER ϕ

Fig. 2.6 Diffraction efficiency of the first order wave for gratings with various values of thickness parameter Q . $N_0 = 1.54$, $\lambda_a = 0.5$, $T = 12$, $NL = 12$.

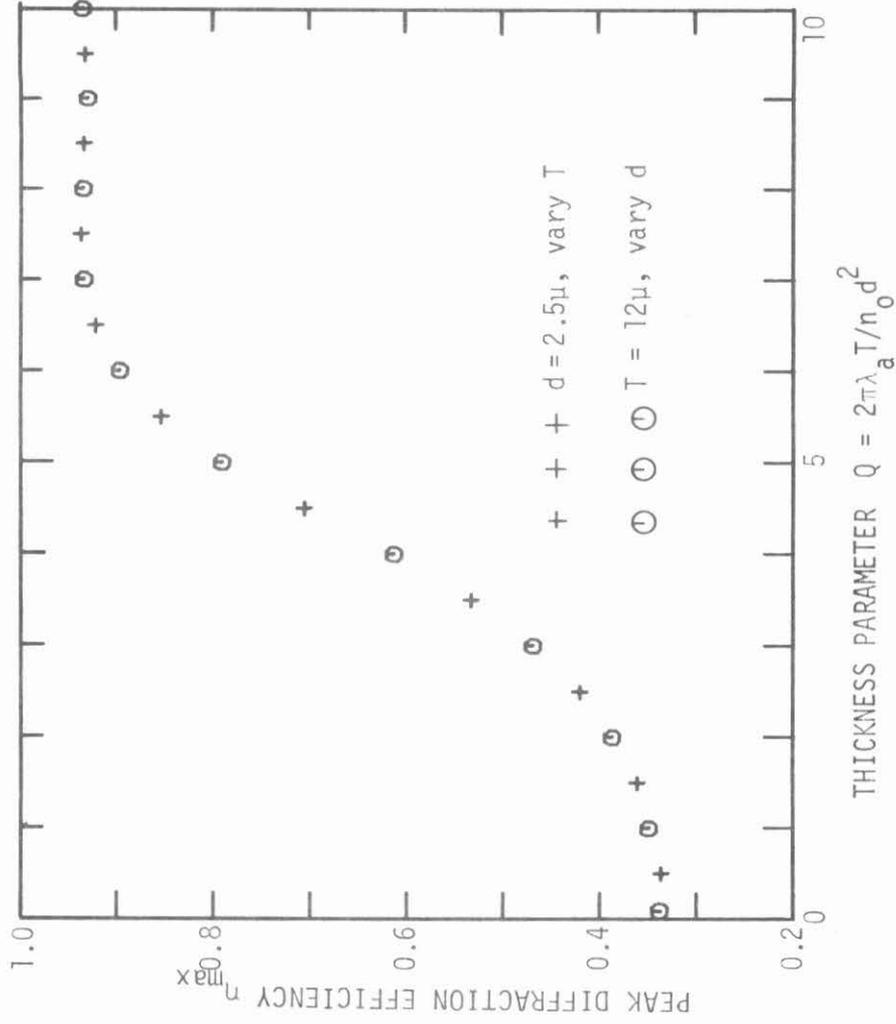


Fig. 2.7 Diffraction efficiency at the first maximum for gratings of various thicknesses and grating periods. $n_0 = 1.54$, $\lambda_a = 0.5$, $NL = 12$.

2.2 Digital Holograms

Many applications of holography require holograms which cannot be produced by the recording of readily available wavefronts. Examples of such applications are spatial filtering,⁽²²⁾ in which the desired filter function may be an arbitrary complex function, or interferometric testing of aspheric optical components,⁽²³⁾ in which the desired reference wavefront may be difficult to obtain with combinations of simple optical elements. For many of these applications, digital holograms are ideally suited. The term "digital hologram" is applied here to any hologram in which the recorded wavefront is obtained directly or indirectly from a computer. Many methods for producing such holograms have been invented over the last few years. A good survey is found in the review paper by Huang.⁽²⁴⁾ The following sections present a brief description of some of the basic types of digital holograms.

2.2.1 Thin Amplitude Holograms

An ordinary hologram recorded on a photographic plate contains a microscopic pattern of dark bands corresponding to the interference fringes in the recorded wavefronts. Such a pattern may be simulated by drawing millions of little lines on a computer plotting device, and photoreducing the resulting plot. A more efficient method was presented by Brown and Lohmann.^(25,26) An array of short lines is drawn with the computer and photoreduced. The resulting hologram consists of an array of transmitting apertures on a dark background, as shown in Figure 8.

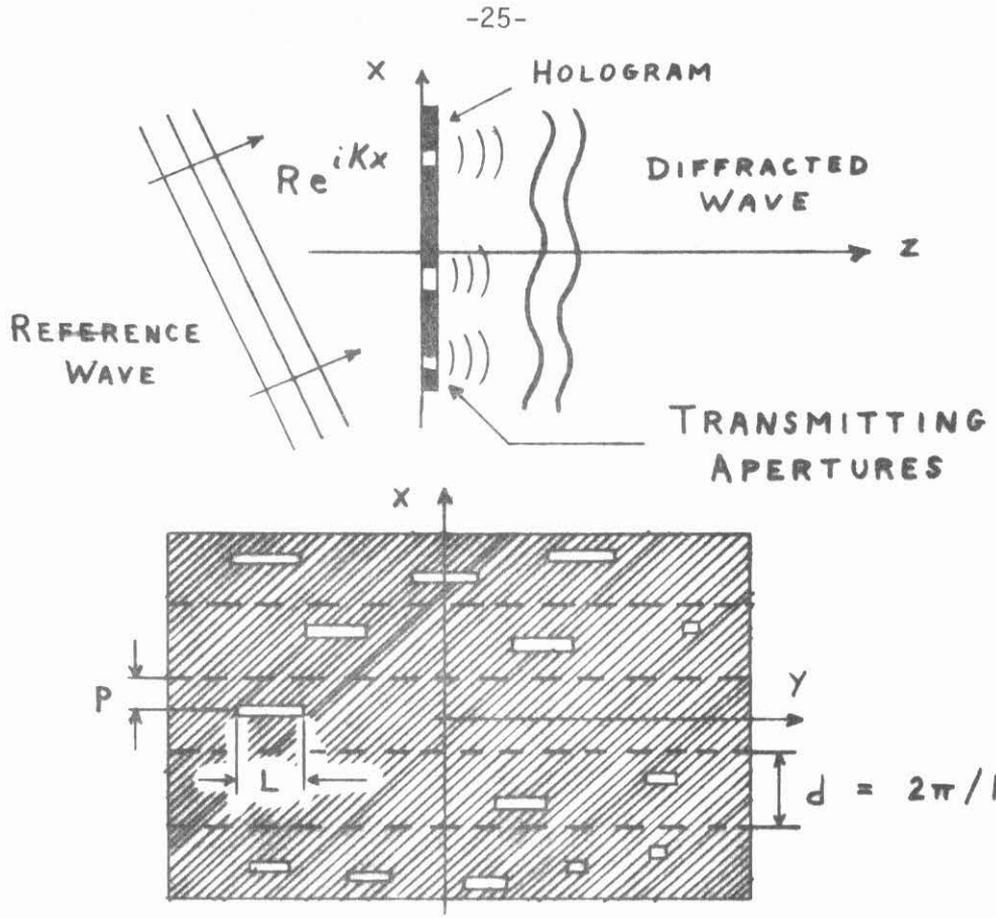


Fig. 2.8 Binary detour phase hologram. Aperture position p determines transmitted phase, and length L determines amplitude.

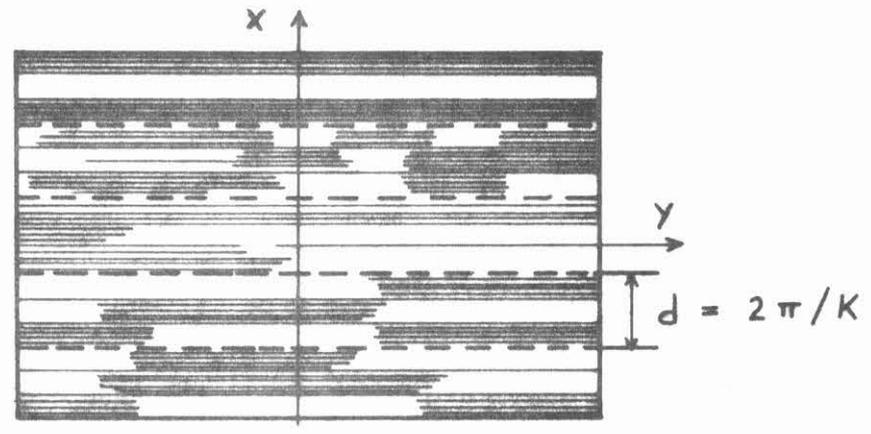


Fig. 2.9 Lee-Burckhardt method. Relative shading of the three bands within each zone determines amplitude and phase of the diffracted wave.

The array is illuminated obliquely, and the desired wavefront is obtained in the far field of the apertures. The hologram is divided into zones in the x direction of Figure 8, each zone corresponding to 2π phase shift of the incident reference wave. The position p of the apertures within each zone will determine the phase of the diffracted wave at that location, and the length L will determine the amplitude. This type of synthetic hologram is called a binary detour phase hologram. Variations on this method include modulation of the width of the apertures instead of the length, and splitting each aperture into a pair,⁽²⁷⁾ controlling the amplitude of the wave by the separation of the pair in the x -direction. The scheme illustrated in Figure 8 is particularly suited to mechanical plotting devices with a fixed pen width.

Another method for producing a digital hologram, which takes advantage of the capability of a CRT plotting device to produce variable density plots, is the method introduced by Lee⁽²⁸⁾ and later simplified by Burckhardt.⁽²⁹⁾ In this method the 0 - 2π phase zones in the hologram are divided into three bands as illustrated in Figure 9. Amplitude and phase of the diffracted wave are controlled by the relative shading of the bands in each zone. Photographic processing must be more carefully controlled than in a binary hologram, in order to insure repeatability of the density levels.

Both these methods involve photoreduction of a two-dimensional graphic pattern, and hence are inherently limited to thin holographic media. The maximum theoretical diffraction efficiency for thin binary

amplitude holograms is obtained when the width of the aperture is half the width of the phase zone. The relative amplitude of the diffracted wave will then be:

$$\frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{i\phi} d\phi = \frac{1}{\pi} \quad (2.30)$$

The maximum diffraction efficiency is then

$$\eta_{\max} = \left(\frac{1}{\pi}\right)^2 = 10.1\% \quad (2.31)$$

A similar result for the three-zone hologram is obtained by placing limits $\pm\pi/3$ on the integral (equation 30). The maximum diffraction efficiency for the Lee-Burckhardt hologram is

$$\eta_{\max} = \left[\frac{1}{\pi} \sin(\pi/3)\right]^2 = 7.6\% \quad (2.32)$$

2.2.2 Thin Phase Holograms

A thin phase hologram, called a kinoform,⁽³⁰⁾ which has high diffraction efficiency and makes good use of available plotter capacity, can be synthesized by recording only the phase part of the desired wavefront. The wavefront will, in general, have both amplitude and phase

$$\tilde{a}(x,y) = a(x,y) e^{i\phi(x,y)} \quad (2.33)$$

but in the case of a diffuse object, the average intensity at the hologram will be nearly constant.

The function $\phi(x,y)$ (modulo 2π) is generated on a suitable graphic device and recorded in a thin phase medium, such as a bleached photographic emulsion. The kinoform is illuminated with an on-axis reference wave, and the transmitted wavefront is then

$$\tilde{a}_K(x,y) = a_0 e^{i\phi(x,y)} \quad (2.34)$$

The discarding of amplitude information results in noise in the reconstructed image, but for applications such as displays, this may not be a problem.

Kermisch has presented an analysis of the noise in kinoform images.⁽³¹⁾ If there are no nonlinearities or imperfections in the recording of the phase function, 78% of the total energy goes into reconstruction of the original image, and the remaining 22% goes into noise in the form of multiple self-convolutions of the image.

The kinoform is an on-axis hologram. Elimination of the carrier grating results in large savings of plotter capacity, since the resolution cell of the plot need be no smaller than the sampling interval of the phase function $\phi(x,y)$.

2.2.3 Phase and Amplitude Holograms

A true phase and amplitude hologram has recently been synthesized in a multilayer color film.⁽³²⁾ Phase and amplitude functions are recorded in separate layers of the same emulsion. If the reconstruction wavelength is in the red region of the spectrum, then the transmitted amplitude will be controlled primarily by the red absorbing layer of the

emulsion, and patterns recorded in the blue and green absorbing layers will cause a phase shift of the transmitted wave. This method allows synthesis of an on-axis hologram with 100% maximum theoretical reconstruction efficiency.

The holograms are synthesized on Kodachrome II in order to take advantage of the highly standardized commercial processing. The color film is very grainy compared to high resolution holographic emulsions, and this may cause noise problems in applications such as spatial filtering. Also, the medium is not suitable for immersion in a liquid gate, because the phase image depends on surface relief effects.

The phase shift due to dye absorption in the medium may be calculated from an integral derived from the Kramers-Kronig relations. Given the spectral density function $D(\omega)$, the phase shift $\phi(\omega_0)$ is shown in Appendix P to be

$$\phi(\omega_0) = \frac{\ln(10)}{\pi} \text{P.V.} \int_0^{\infty} \frac{D(\omega)\omega_0 d\omega}{\omega^2 - \omega_0^2} \quad (2.35)$$

where P.V. means the principal value integral.

Numerical calculations using this integral and experimental data from Kodachrome II film indicate that the expected phase shift at the wavelength of the HeNe laser due to absorption in the dye is much less than one wavelength (see Figure 10). The original data for the dye absorption, obtained from Kodak, showed a peak density of about one. The data for the calculation were multiplied by three to represent a more heavily dyed layer. A program for computing integrals of the above form is shown in Appendix P.

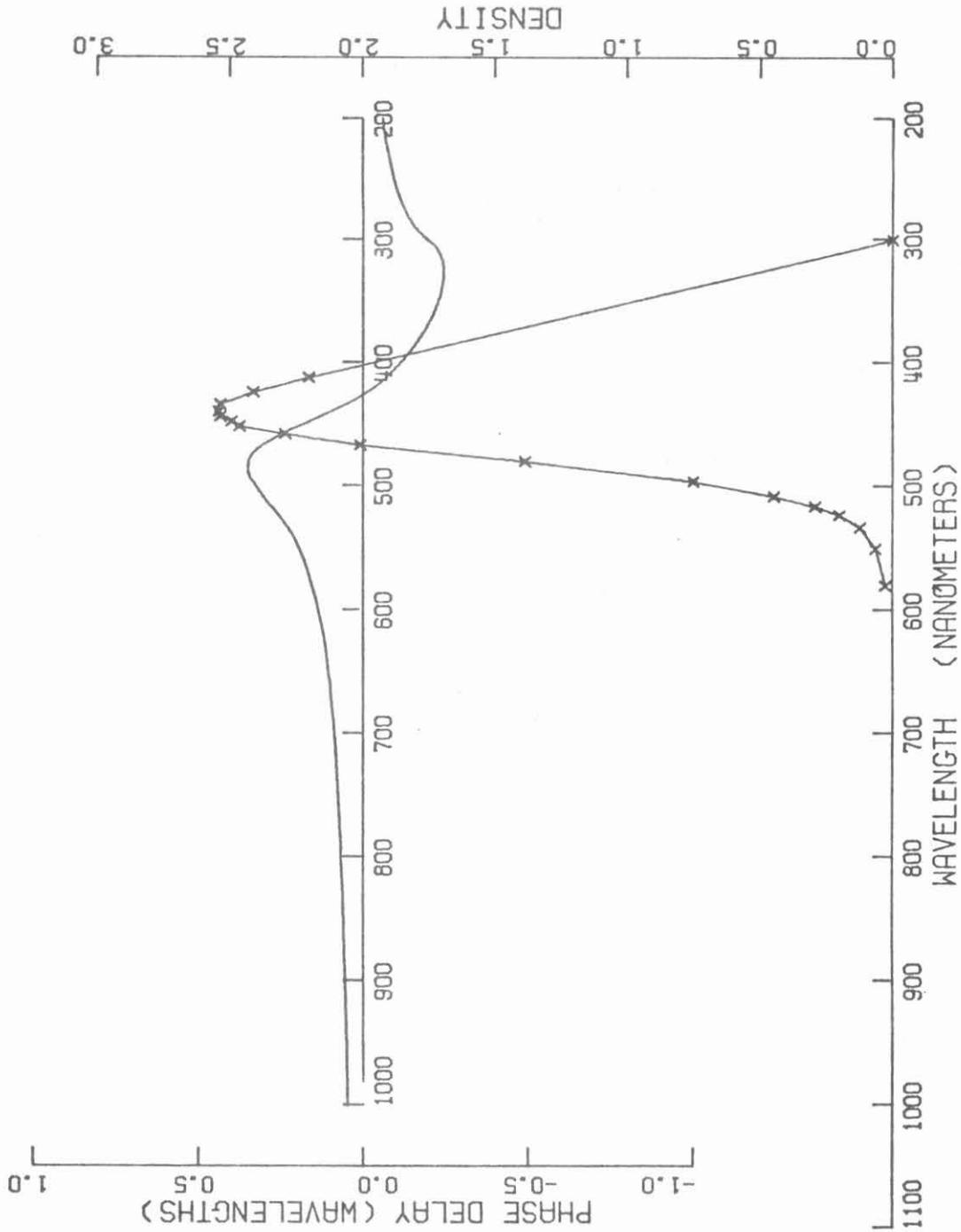


Fig. 2.10 Calculated phase delay due to the blue absorbing dye in Kodachrome II. Dye absorption data have been arbitrarily upscaled to represent a more heavily dyed layer.

For application as a hologram, a phase shift of at least one wavelength is necessary, and it would be desirable if this phase shift did not depend on surface relief. A dye with a sharp edge in the absorption characteristic near the operating wavelength could possibly produce the necessary phase shift.

A rather unique hologram, which may be synthesized in any holographic medium, is that of King et al.⁽³³⁾ In this method, a series of perspective views of a three-dimensional object are computed and drawn as ordinary two-dimensional images. The series is recorded by ordinary holographic techniques in sections on a large holographic plate. The result is a computer-generated holographic stereogram which simulates the calculated three-dimensional image. The hologram shown in the paper was recorded on dichromated gelatin and shows a bright reconstruction of an assembly of random intersecting lines. This method has recently been applied to the generation of displays.⁽³⁴⁾ Application as a spatial filter does not seem possible.

2.2.4 Space-Bandwidth Limitations

Information for the synthesis of a computer-generated hologram must be stored in the form of a finite number of discrete samples. This will impose certain fundamental limitations on the reconstructed wavefront analogous to the bandwidth limitations on a sampled function of time.

Consider a square hologram of side L which is constructed from N by N uniformly spaced samples (see Figure 11). Ideally, the

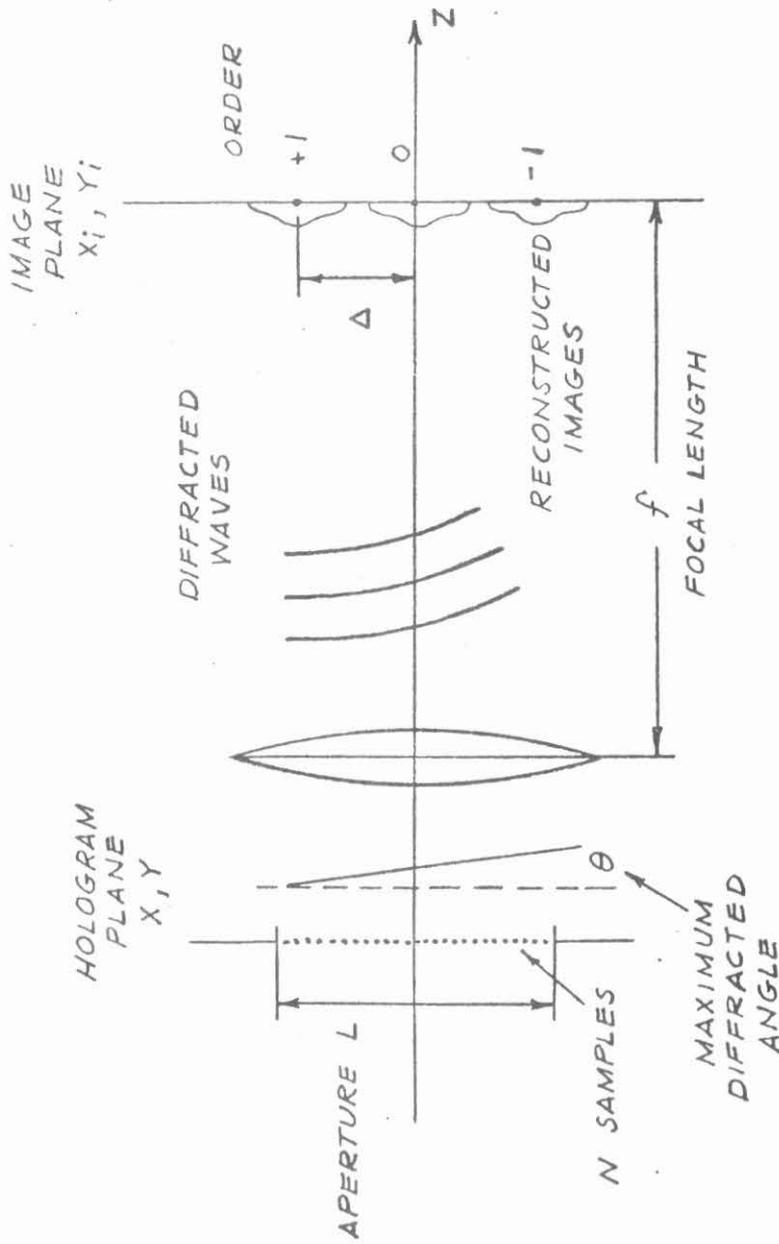


Fig. 2.11 Illustration of space-bandwidth limitations in digital holograms.

hologram will be an array of N by N small areas, each with its own independently controlled phase and amplitude transmittance.

The diffracted waves are imaged in the back focal plane of a lens. Each point in the focal plane corresponds to a unique spatial frequency in the wavefront $\tilde{a}(x,y)$ transmitted by the hologram. In a real hologram the lens is often incorporated into the hologram as a spherical phase factor, but is shown separately here for clarity.

If we apply the sampling theorem to the reconstructed wavefront $\tilde{a}(x,y)$, we find that \tilde{a} must have a spatial frequency band limit $B = 1/2d$, where d is the sample spacing. Considering both positive and negative spatial frequencies, the space-bandwidth product in one dimension of $\tilde{a}(x,y)$ will be

$$2BL = 2L/2d = N$$

The space-bandwidth product in two dimensions is then N^2 , the total number of samples.

The regular array of N by N samples will reconstruct several images in the focal plane, one for each diffracted order of the array. In order that the images not overlap, they must have a maximum size Δ equal to the spacing of the diffracted orders.

$$\Delta = \theta f = (N\lambda/L)f \quad (2.36)$$

The minimum separation of two resolvable spots in the image is given by the Rayleigh criterion as:

$$\delta = (\lambda/L)f \quad (2.37)$$

The total number of resolvable spots in the two-dimensional image is then

$$N_s = (\Delta/\delta)^2 = N^2 \quad (2.38)$$

which is the same as the space-bandwidth product of the wavefront $\tilde{a}(x,y)$.

Another way of stating the space-bandwidth limitation is that the reconstructed wavefront must have a limited angular bandwidth. For the holograms of Chapter 3, $L = 25$ mm, $N = 1000$, and $\lambda = 0.5\mu$. The maximum angular bandwidth is then

$$\theta = N\lambda/L = 0.02 \text{ radians}$$

A useful criterion for comparing various digital holograms is the space-bandwidth efficiency, or ratio of number of resolution cells required of a plotting device to the space-bandwidth product of the reconstructed wavefront. Ransom⁽³⁵⁾ has compared several hologram coding schemes on this basis. (Space-bandwidth efficiency as defined above is the inverse of Ransom's p factor.)

The binary hologram (Figure 8) typically requires 10-50 resolution cells per sample, depending on how much noise is tolerable in the image. Holograms generated with a variable density plotter typically require far less resolution. Lee's hologram, as modified by Burckhardt (Figure 9) requires three resolution cells per sample. The kinoforms and color film holograms require one resolution cell per sample. The modulated grating hologram (Chapter 3) requires one cell per sample.

2.3 Holographic Spatial Filters

One of the major applications of computer generated holograms is spatial filtering. Spatial filtering can be used to perform a number of interesting operations on images, including recognition of predetermined patterns⁽³⁶⁾, coding of images⁽³⁷⁾, and restoration of blurred images.⁽³⁸⁾ The process of linear, spatial-frequency filtering of images is often performed by using a digital computer. This is especially convenient for images which are already in digital form, such as spacecraft photos.⁽³⁹⁾ The computer is also able to perform a large variety of nonlinear operations on an image.⁽⁴⁰⁾ The major advantage of coherent image processing is speed. An optical processor performs instantaneously, Fourier transforming and filtering operations which take nearly a minute on the fastest computer.⁽⁴¹⁾

2.3.1 A Simple One-Lens Processor

A simple one-lens optical processor for performing spatial filtering operations with holographic filters is illustrated in Figure 12.

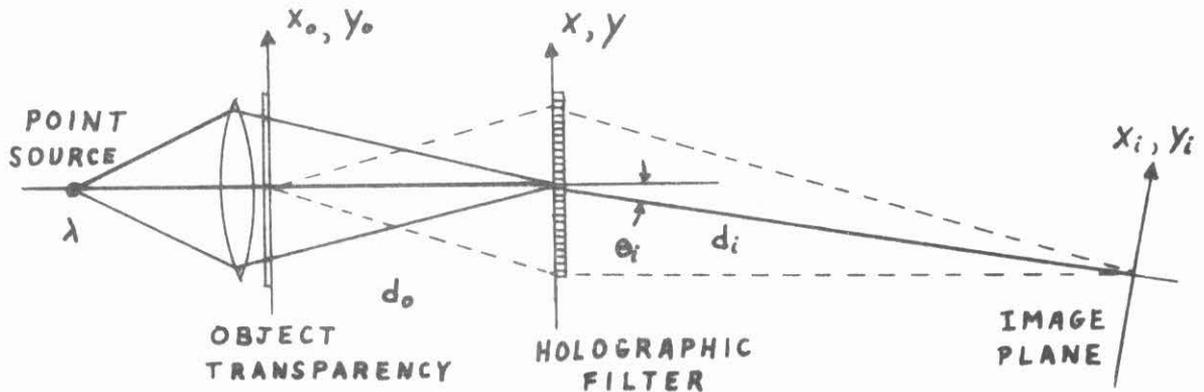


Fig. 2.12 A simple, one-lens, optical processor using a holographic spatial filter. Vertical dimensions are greatly exaggerated for clarity.

Coherent optical processors occur in many configurations. A description of the more basic types can be found in several texts^(42,43) and in the paper by Arsenault⁽⁴⁴⁾. A processor of the type shown in Figure 12 was used by Ragnarsson⁽⁴⁶⁾ in experiments with deblurring filters. An analysis of this processor is given in Appendix N. A brief intuitive description of its operation follows.

The object transparency containing the image to be filtered $t(x_0, y_0)$ is illuminated with a converging spherical wave, focal distance d_0 . The holographic spatial filter placed at this focal plane contains a lens-like phase factor which images points on the transparency onto points in the output plane. Imaging of the central point is shown by the dashed lines.

The object transmittance function can be written as an integral, or summation of its Fourier components

$$t(x_0, y_0) = \iint \tilde{T}(v_x, v_y) e^{i2\pi(x_0 v_x + y_0 v_y)} dv_x dv_y \quad (2.39)$$

v_x, v_y : spatial frequencies

Each Fourier component deflects part of the energy in the converging wave to a new location (x, y) in the filter plane, the angle of deflection being proportional to the spatial frequency.

$$\begin{aligned} x &= \lambda d_0 v_x \\ y &= \lambda d_0 v_y \end{aligned} \quad (2.40)$$

The light incident on the filter is not an exact Fourier transform of $t(x_0, y_0)$, but differs by a spherical phase factor associated

with the distance d_0 . This spherical phase factor is compensated in the hologram transmittance function:

$$\tilde{H}(x,y) e^{-iKx - ik(x^2 + y^2)/2f} \quad (2.41)$$

$$K = k \sin \theta_i$$

$$k = 2\pi/\lambda$$

$$f = [1/d_0 + 1/d_i]^{-1}$$

The function $\tilde{H}(x,y)$ is the complex filter transmittance, and the phase factor is the paraxial approximation to an off-axis converging spherical wave. The focal length f is chosen to image the transparency onto the output plane. The processor is, in effect, an imaging system in which each spectral component of the image may be modified by a complex factor \tilde{H} .

The impulse response of the filter may be obtained at the output plane by placing a pinhole (delta function) at the center of the input plane. The amplitude at the output plane can be written as a convolution of this impulse response with the input function

$$\tilde{a}_i(x_i, y_i) \approx t * \tilde{h}(x_i, y_i) \quad (2.42)$$

The analogy between optical processors and electrical filters is often stressed in optical texts. A significant difference is that the impulse response for an electrical filter must be zero for time < 0 , resulting in a fixed relation between the real and imaginary parts of the filter transfer function $\tilde{H}(\omega)$.⁽⁴⁵⁾ No such restriction exists for spatial filters, and we are free to choose any arbitrary complex filter function.

2.3.2 Deblurring Filters

An application of spatial filtering which has received considerable attention is the restoration of blurred images. (38,46,47)

A simple incoherent blurring process, such as camera motion, mis-focus, or certain types of instrumental defects, may be represented as a linear, space-invariant convolution.

$$I_b(x,y) = I_o * b(x,y) \quad (2.43)$$

$I_o(x,y)$: original image intensity

$b(x,y)$: point-spread function

$I_b(x,y)$: blurred image intensity

The essential assumption here is that the point-spread function does not change over the image area to be filtered. The treatment of space-variant processes is considerably more complicated. (48)

The blurring process may be written in terms of the Fourier transforms of the above intensity distributions:

$$\tilde{\mathcal{I}}_b(v_x, v_y) = \tilde{\mathcal{I}}_o(v_x, v_y) \tilde{B}(v_x, v_y) \quad (2.44)$$

$$\tilde{\mathcal{I}} = \mathcal{F}\{I\}$$

$$\tilde{B} = \mathcal{F}\{b\}$$

Ideally, one may recover the original image spectrum, and hence the original unblurred image, by a process of Fourier transform division

$$I_o(x,y) = \mathcal{F}^{-1}\{\tilde{\mathcal{I}}_b/\tilde{B}\} \quad (2.45)$$

In the system of Figure 12, this process could be accomplished by recording the blurred image on the input medium so that the amplitude

transmittance is proportional to blurred image intensity,

$$t_o(x_o, y_o) \propto I_b(x_o, y_o) \quad \text{linear recording} \quad (2.46)$$

The spectral distribution incident on the filter will then be the blurred image spectrum,

$$\tilde{T}(v_x, v_y) \propto \tilde{\mathcal{J}}_b(v_x, v_y) \quad (2.47)$$

A filter must be synthesized which is inverse to the blurring function:

$$\tilde{H}(v_x, v_y) = \tilde{B}^{-1}(v_x, v_y) \quad (2.48)$$

The amplitude at the output plane will then be proportional to the original unblurred image intensity:

$$\begin{aligned} a_i(x_i, y_i) &\propto I_o(x_o, y_o) & (2.49) \\ x_i &= (-d_i/d_o)x_o \\ y_i &= (-d_i/d_o)y_o \end{aligned}$$

The problem is that the spectral blurring function B may have zeroes. A simple example is that of blurring due to mis-focus. The point-spread function is a small disc of uniform intensity,

$$\begin{aligned} b(r) &= \frac{1}{\pi r_o^2} \text{circ}(r/r_o) & (2.50) \\ \text{circ}(r/r_o) &= \begin{cases} 1 & r \leq r_o \\ 0 & r > r_o \end{cases} \\ r &= (x^2 + y^2)^{1/2} \end{aligned}$$

The spectral blurring function for this circularly symmetric case is found by a Fourier-Bessel transform.⁽⁴⁹⁾

$$B(\rho) = J_1(2\pi\rho)/\pi\rho \quad (2.51)$$

$$\begin{aligned} \rho &= r_0 v \quad \text{normalized spatial frequency} \\ v &= (v_x^2 + v_y^2)^{1/2} \end{aligned}$$

This function and its approximate inverse is shown in Figure 13. Any physically realizable filter will have a limited dynamic range, and the poles in the inverse filter must be truncated in some fashion. This truncation results in holes in the overall transfer function, as shown in Figure 14. It is desirable to have a filter with as large a dynamic range as possible, in order to minimize the loss of information associated with these gaps in the recovered spectrum.

The dynamic range is defined as the ratio of maximum to minimum usable filter transmittance,

$$H_{\min} < H(v_x, v_y) < H_{\max} \quad (2.52)$$

This ratio will depend on the process by which the filter is made. Filters generated by analog methods are typically limited by the linearity range of the recording medium. Computer-generated filters may avoid the nonlinearity problem by pre-distortion of the filter function in the computer. The dynamic range limitation in this case will be due to noise in the medium or in the recording processes, including quantization noise inherent in the digital process. The fundamental limitation will ultimately be noise in the holographic filter medium. This is the subject of Chapter 4.

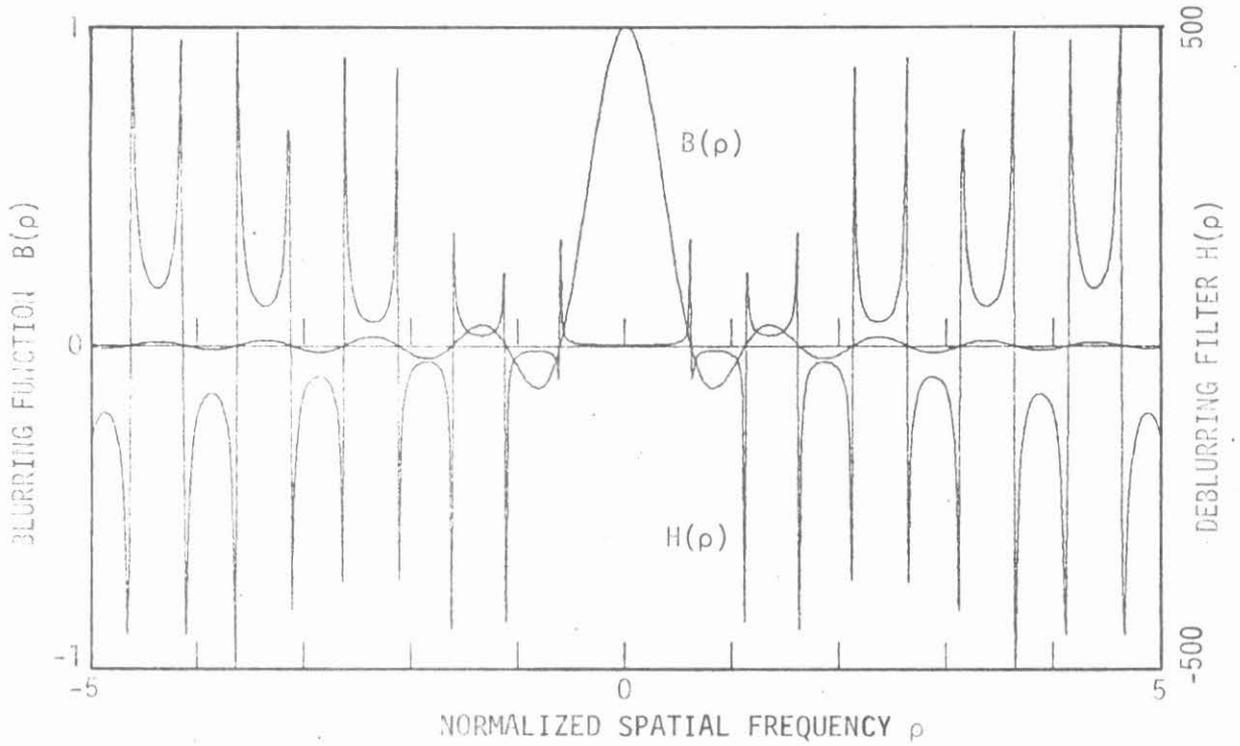


Fig. 2.13 Spectral blurring function $B(\rho)$ and optimum inverse filter $H(\rho)$ for an image blurred by a circular point-spread function.

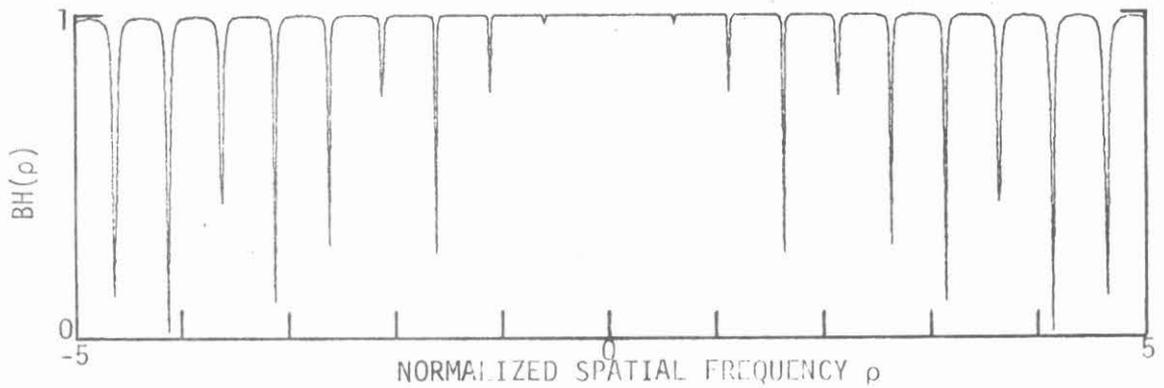


Fig. 2.14 Overall transfer function for the linear blurring-deblurring process of Figure 13.

Noise in the input medium will also fundamentally limit the de-blurring process. For this case an optimum linear estimation filter was derived for Helstrom⁽⁵⁰⁾ by the method of linear mean-square estimation.⁽⁵¹⁾

Assuming the noise is Gaussian and additive, Helstrom showed that the linear filter which results in the least overall mean-square deviation of the recovered image from the original is given by:

$$\tilde{H}(v_x, v_y) = \tilde{B}^{-1}(v_x, v_y) [1 + \Phi_N / (\Phi_S |\tilde{B}|^2)]^{-1} \quad (2.53)$$

$\tilde{B}(v_x, v_y)$: Spectral blurring function

$\Phi_N(\rho)$: Noise power spectral density, isotropic

$\Phi_S(v_x, v_y)$: Signal power spectral density, $|\mathcal{G}_0|^2$

In areas of the filter where the blurred-image power spectral density, $\Phi_S |\tilde{B}|^2$, exceeds Φ_N , the optimum filter is simply \tilde{B}^{-1} . Where the blurred signal falls below the noise, the filter transmittance drops rapidly to zero.

The maximum quantity of information in a blurred, noisy image may be stated, by analogy with the information capacity of a noisy communications channel,⁽⁵²⁾ as

$$Q = N^2 \log_2 [1 + \Phi_S \tilde{B}^2 / \Phi_N] \quad (2.54)$$

where N^2 is the number of independent image components, or equivalently, the space-bandwidth product. The above formula would apply to an optimally-coded image degraded by Gaussian, additive noise. Quite

often in images the noise is multiplicative and non-Gaussian, and the above formula must be applied cautiously. The point to be made here, however, is that the blurring of an image in the presence of noise entails loss of information. In a deblurring process the remaining information is merely modified to a form more suitable for assimilation by an observer. Thus, characters which are unreadable in a blurred image may be made easily legible in the deblurred image without violating the laws of information theory.

2.3.3 Synthesis of Deblurring Filters

Spatial filters for image deblurring may be divided into two classes, depending on the method of synthesis. In analog methods the filter is produced by optical techniques from the recorded point-spread function. In digital methods the filter is computed, and appropriate masks generated to produce a hologram with the desired transmittance.

Analog Methods

The synthesis of spatial filters for image enhancement has been intensively studied for many years. Maréchal and Croce⁽⁵³⁾ first demonstrated the principle of spatial filtering by increasing the contrast of an image with a simple attenuative filter. Tsujiuchi⁽⁵⁴⁾ constructed a filter with positive and negative real transmittance values. The 180° phase shift was obtained by vacuum deposition of a thin layer of MgF_2 in zones where the filter transmittance was negative. The use of a hologram as a spatial filter was demonstrated by VanderLugt,⁽⁵⁵⁾ in his early experiments with character recognition filters. The use of a hologram allows the synthesis of filters with arbitrary complex transmittance functions.

A method of generating a holographic spatial filter for image deblurring was shown by Stroke and Zech.⁽⁵⁶⁾ This filter is generated as follows: The point-spread function $b(x,y)$ is first recorded on film. This film is placed as the input to a processor, and the blurring function \tilde{B} is obtained at the transform plane. The essence of this method is the realization that the filter \tilde{B}^{-1} may be obtained as a product of two easily recorded filters:

$$\tilde{B}^{-1}(v_x, v_y) = (\tilde{B}^*)(1/|\tilde{B}|^2) \quad (2.55)$$

The first filter, \tilde{B}^* , is available as the conjugate wave in a hologram recorded at the transform plane. The second filter is a purely attenuative (real) transmittance obtained by recording the intensity pattern $|\tilde{B}|^2$ on a photographic negative. Exposure and development parameters must be carefully controlled on the second filter, and even then the desired inverse characteristic is obtained only over a limited range. Extension of this range requires a series of carefully controlled masking steps in the generation of the second filter.⁽⁵⁷⁾ Examples of images deblurred by this method are shown in several papers.^(47,58,59)

An alternative method for recording holographic filters for deblurring is the "weak reference" filter of Ragnarsson.⁽⁴⁶⁾ This method produces in a single recording step, filters of high diffraction efficiency and relatively large dynamic range. Normally, in a holographic recording the reference wave is made much stronger than the object wave in order to produce a linear recording of the object. An inversion effect may be obtained by making the reference wave much weaker and exposing the hologram in the saturation region. With a constant, weak reference wave, increasing the amplitude of the object wave increases the total exposure and drives the medium further into the saturation region. With the proper combination of exposure and development, the desired inverse characteristic can be obtained over a relatively large range. Ragnarsson used a special monobath developer and a chloride bleach to obtain holograms with a 250:1 exposure range.

Tichenor⁽⁶⁰⁾ has extended the range of weak reference filters by sandwiching the filter with a "primary filter" which performs the same function as the attenuative filter in Stroke's hologram. The large dynamic range is absorbed in the primary filter, allowing the holographic filter to operate within its optimum range. By using a cascade of such filters, the dynamic range may be extended beyond the range of each emulsion.

An alternative to cascading of filters is to place them side-by-side in a multichannel processor.⁽⁶¹⁾ Each filter operates only on that part of the object spectrum within a limited amplitude range. The different spectral components of the image are recombined at the output plane. Alignment of the system is critical in order to insure proper relative phase of each of the components.

The major problem in synthesis of deblurring filters by analog methods has been the achievement of large linear ranges in media which are inherently nonlinear. Also, since the point-spread function is first recorded on film in the above methods, the class of filters which may be synthesized is limited to those filters with a real impulse response.

Digital Methods

The problem of nonlinearities in the recording of filters by analog methods is avoided in digital generation of filters. Whatever nonlinearity exists in the various media and processes in the generation of a digital hologram may be pre-compensated by distortion of the functions in the computer, provided only that the processes are well

enough controlled to produce a repeatable nonlinear characteristic. The inherent flexibility of the digital computer also allows synthesis of any arbitrary complex filter function.

Digital holographic filters may be synthesized by any of the methods discussed in Section 2.2. The principal limitation in a digitally recorded filter is the relatively small space-bandwidth product compared to analog filters. The problem is not in the recording of the required filter function \tilde{H} , which typically has a small space-bandwidth product, but in generating the high-frequency carrier grating which is necessary to insure adequate separation of the diffracted orders (cf., Fig. 2.11). If the filter were recorded on a minimum-frequency grating, the small image corresponding to the impulse response of the filter would just touch the images from adjacent orders at the output plane. The minimum necessary grating frequency in a spatial filter will be determined by the required separation of orders, which depends on the size of the image to be filtered.

On-axis holograms avoid the separation-of-orders problem, but may be subject to other limitations. The color-film holograms, at least with presently manufactured emulsions, may be too noisy for high quality spatial filters. The possibility of using this type of hologram as a spatial filter is presently being investigated.⁽⁶²⁾ The kinoforms are out of the question because of noise problems.

A method has recently been proposed to produce a digital filter on a high-frequency carrier.⁽⁶³⁾ The filter function is first synthesized on an ordinary binary hologram. The hologram is placed as an

input to a processor, and the first-order wave (filter function \tilde{H}) is extracted using a properly located aperture at the transform plane. The function \tilde{H} may now be recorded on a second hologram with a high grating frequency. The result is a filter with a much larger free spectral range, but the dynamic range can be no larger than that of the initial digital hologram.

The modulated grating hologram (Chapter 3) was designed with the spatial filtering application in mind. In principle, it offers large dynamic range and large free-spectral range. Presently, it suffers from certain technological problems which prevent its application as a precision spatial filter.

2.3.4 Noise Problems in Coherent Optical Systems

Images which have been deblurred in coherent optical systems are generally of very poor quality in comparison to what can be done digitally. Much of the problem is not related to dynamic range limitation in the filter. The major practical problem is that of obtaining a clean, linear recording of the blurred image. Noise and nonlinearities in the recording of the blurred image will severely degrade the restoration, as shown in the computer simulation following.

In a coherent imaging system noise is often much worse than in incoherent imaging systems. Noise arising from dust and scratches is coherent with the image and can cause large fluctuations in image intensity. Multiple reflections of the highly collimated light from lens surfaces can produce false images overlapping the primary image. The

presence of hard apertures may cause "ringing" at the edges of bright areas in the image. Liquid gates are often necessary to eliminate phase noise at emulsion surfaces. A spherical wave passing through a flat liquid gate will suffer aberrations. In general, extreme care is needed in the design and use of a coherent optical processor.

2.4 Computer Deblurring Experiment

In order to illustrate the deblurring process and some of its practical limitations, and to show what might ultimately be done with a coherent optical filter, a computer simulation was carried out.

A spoke target was generated on a 512 by 512 array, shown in Figure 15a. This array was transformed using a two-dimensional complex Fast-Fourier-Transform (Appendix C). The power spectrum of the transformed target is shown in Figure 15b. The full density range in this image corresponds to six decades on a logarithmic intensity scale.

The point-spread function for this demonstration was chosen as a circle function (equation 2.50) with $r_0 = 10$ sample spacings. The discrete PSF can be written as:

$$b_{mn} = \frac{1}{\pi r_0^2} \text{circ}[(m^2 + n^2)^{1/2} / r_0] \quad (2.56)$$

The discrete spectral blurring function is

$$B_{\mu\nu} = J_1(2\pi\rho_{\mu\nu}) / (\pi\rho_{\mu\nu}) \quad (2.57)$$

$$\rho_{\mu\nu} = r_0 s_{\mu\nu} = 10 s_{\mu\nu}$$

$$s_{\mu\nu} = (\mu^2 + \nu^2) / 512$$

$$\mu, \nu = [-256, 255]$$

Application of this blurring function to the spectrum of 15b results in the power-spectral distribution of 15c. The transform of this blurred spectrum gives a blurred image, 15d. The loss of high spatial frequencies in an image spectrum corresponds to loss of fine detail and sharp edges in the image.

Inspection of the spokes in 15d shows areas of reversed contrast. The spectral blurring function is negative for these frequencies. The phase reversals are more apparent in a high-contrast version of the same image, 15e. This image was obtained by using a binary intensity scale, with the threshold set at the middle of the intensity range of 15d. It can be seen from this figure that no simple increase in contrast can deblur an image when the blurring process is severe enough to cause phase reversals in certain spatial frequencies.

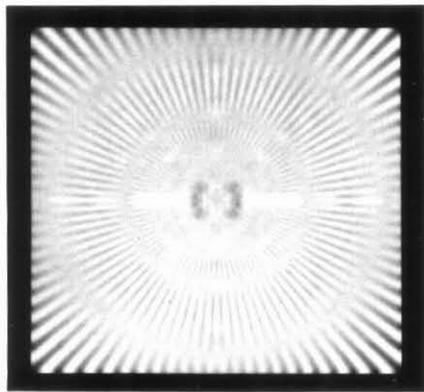
A deblurring filter of the form of an optimum linear filter (equation 2.53) was simulated by assuming a constant signal-to-noise power spectral density ratio. The discrete filter is then

$$H_{\mu\nu} = B_{\mu\nu}^{-1} [1 + CB_{\mu\nu}^{-2}]^{-1} \quad (2.58)$$

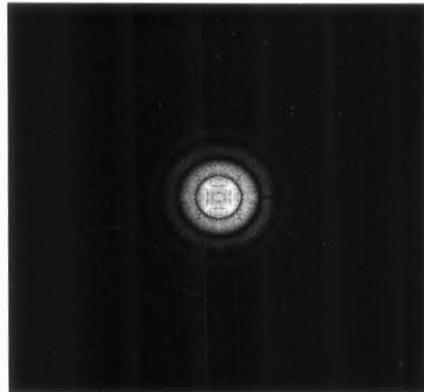
$$C = \Phi_N / \Phi_S = 10^{-6}$$

The plot in Figure 13 was obtained from a scan through the center of this filter. The peak values for a continuous filter with the above value of C would be ± 500 . The irregularity of the peaks in the discrete filter is due to the sampled nature of the function. The overall transfer function of the blurring-deblurring process is shown in Figure 14, which was obtained from a product of the two discrete functions (2.57 and 2.58).

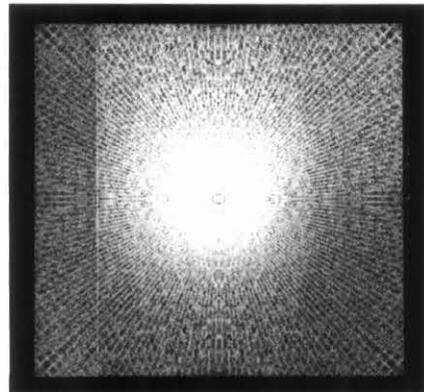
An impulse response of the deblurring filter was obtained by transforming the discrete filter (2.58). A scan through the center of this impulse response is shown in Figure 16. The prominent features



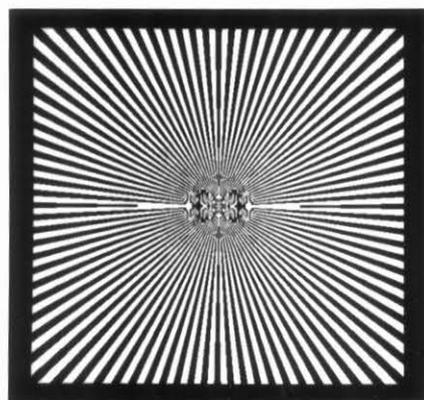
d) Blurred target



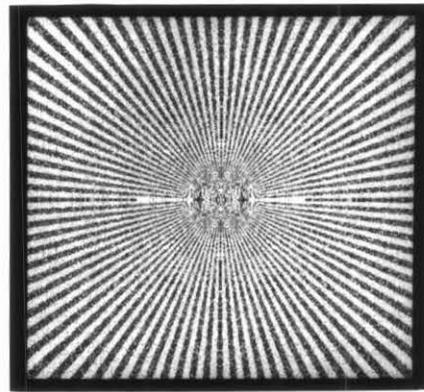
c) Blurred spectrum



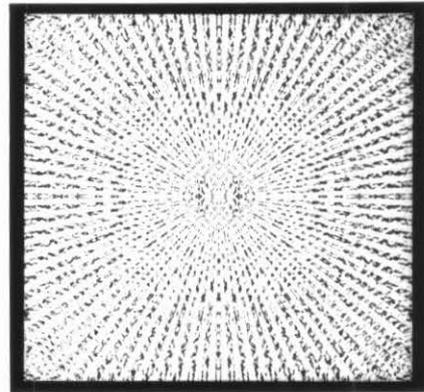
b) Spectrum of target



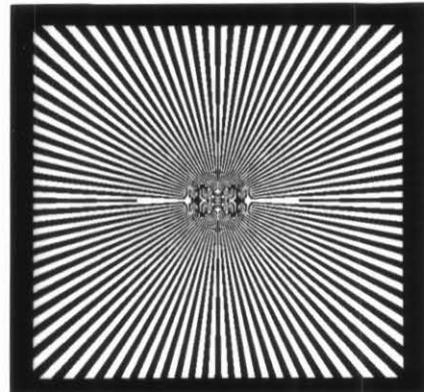
a) Original spoke target



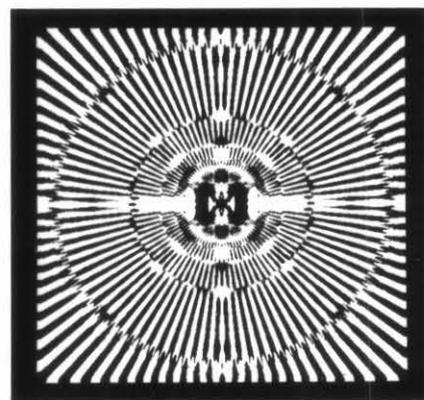
h) Quantization noise



g) Effects of nonlinearity



f) Deblurred target



e) High contrast

Fig. 2.15 Computer deblurring experiment. A spoke target, generated in the computer, was blurred by a circular point-spread function. An inverse estimation filter was simulated, and the effects on the deblurred image, of imperfections in the recording of the blurred image, are shown in g) nonlinear recording, and h) quantization to 128 intensity levels.

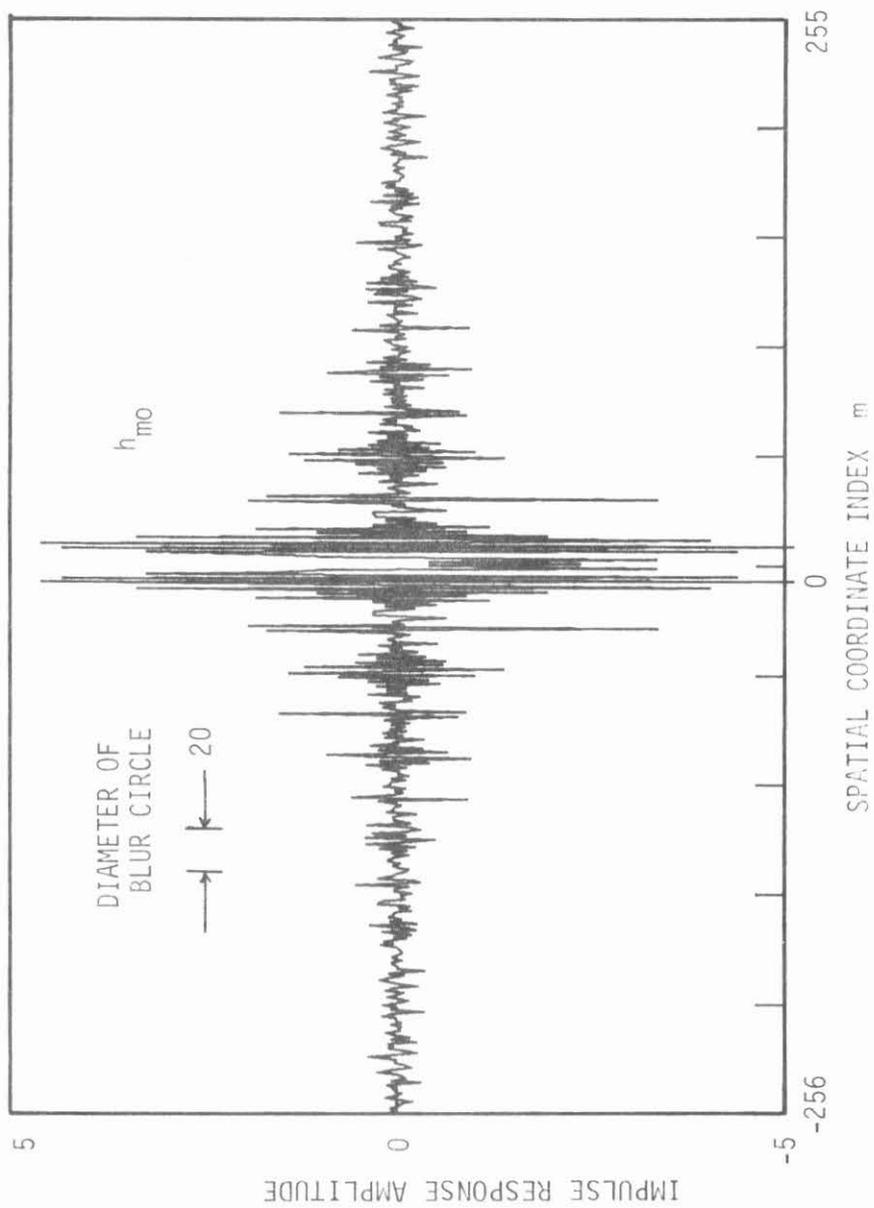


Fig. 2.16 Amplitude of the impulse response h_{mn} of the deblurring filter H_{UV} along a line through the center of the field.

in this pattern are a series of sharp concentric rings of alternating sign. The spacing between groups of these rings is about the diameter of the blur circle. It is easy to imagine how such a function, convolved with the blur circle, would give a restored function nearly zero everywhere except at the center where the circle lines up exactly with the rings.

A real-symmetric point-spread function results in a real-symmetric (but possibly negative) deblurring filter. Likewise, the filter impulse response will be real-symmetric.

Application of the deblurring filter (2.58) to the blurred image, 15d, results in the restored image, 15f. The restoration is visually nearly perfect, although a small amount of noise can be seen in the original photograph.

Figure 15f shows what might ultimately be possible in a coherent processing system. The performance of such a system, however, will be severely degraded by noise and nonlinearities in the blurred image. Blurred images are typically found in one of two forms: a photographic record or a digital record. Photographic films are subject to nonlinearities, and digital image intensity values are typically quantized to levels representable by a few binary bits.

A simple nonlinearity was applied to the values representing the blurred image, 15d. A plot of this nonlinear function, which resembles a film characteristic, is shown in Figure 17. Application of the deblurring filter (2.58) resulted in a restored image with degradations, 15g.

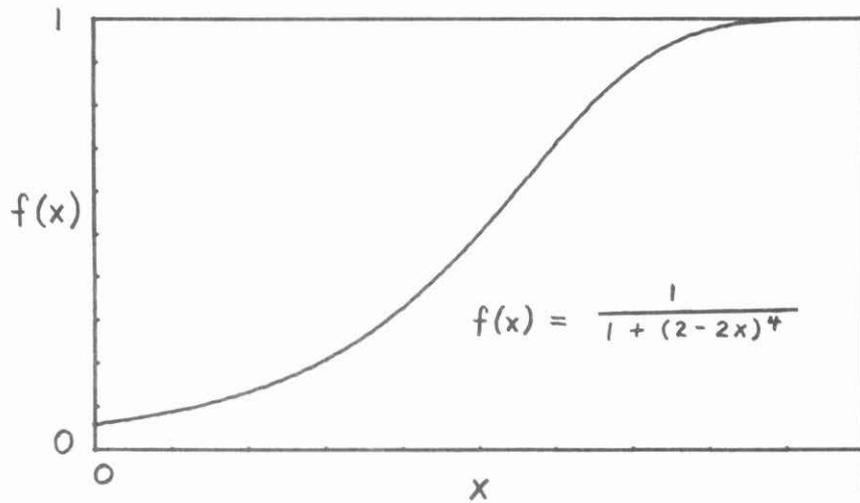


Fig. 2.17 Nonlinear scaling function applied to the blurred image 15d, resulting in degradation of the restored image, 15g.

The effect of quantization noise was simulated by truncating the low-order bits of integers representing the intensity distribution of the blurred image, 15d. Quantizing the blurred image to 128 levels resulted in the degradation of the restored image shown in 15h. Quantization to 256 levels eliminated most of the noise, and quantization to 32 levels totally obliterated the image.

The effect of various degradations on the spectrum of a continuous tone image can be seen in a paper by Anderson and Huang.⁽⁶⁴⁾

References

1. M. Born, E. Wolf, Principles of Optics, 4th ed., (Pergamon, 1970) Sect. 1.4, eq. 54; see also discussion of Wiener's experiments, p.280.
2. G. F. Carrier, M. Krook, C. E. Pearson, Functions of a Complex Variable (McGraw-Hill, 1966), p. 228.
3. Harold Wayland, Complex Variables Applied in Science and Engineering (Van Nostrand, 1970), eq. 8.77, p. 220.
4. J. W. Goodman, "Film-Grain Noise in Wavefront-Reconstruction Imaging" J. Opt. Soc. Am. 57, 493 (1967).
5. J. W. Goodman, G. R. Knight, "Effects of Film Nonlinearities on Wavefront-Reconstruction Images of Diffuse Objects", J. Opt. Soc. Am. 58, 1276 (1968).
6. N. George, J. W. Matthews, "Holographic Diffraction Gratings", Appl. Phys. Lett. 9, 212 (1966).
7. C. V. Raman, N. S. Nath, Proc. Indian Acad. Sci. A2, 406 (1935); A3, 75 (1936).
8. P. Phariseau, "On the Diffraction of Light by Progressive Supersonic Waves", Proc. Indian Acad. Sci. 44, 165 (1956).
9. M. V. Berry, The Diffraction of Light by Ultrasound (Academic Press, 1966).
10. C. B. Burckhardt, "Diffraction of a Plane Wave at a Sinusoidally Stratified Dielectric Grating", J. Opt. Soc. Am. 56, 1502 (1966).
11. C. B. Burckhardt, "Efficiency of a Dielectric Grating", J. Opt. Soc. Am. 57, 601 (1967).
12. H. Kogelnik, "Coupled Wave Theory for Thick Hologram Gratings", Bell System Tech. J. 48, 2909 (1969).
13. F. G. Kaspar, "Diffraction by Thick, Periodically Stratified Gratings with Complex Dielectric Constant", J. Opt. Soc. Am. 63, 37 (1973).

14. G. L. Fillmore, R. F. Tynan, "Sensitometric Characteristics of Hardened Dichromated-Gelatin Films", J. Opt. Soc. Am. 61, 199 (1971).
15. M. V. Berry, The Diffraction of Light by Ultrasound (Academic Press, 1966), Ch. 5, "An Exact Scalar Wave Theory".
16. M. T. Chang, "Holographic Dielectric Gratings: Theory and Practice", Ph.D. Thesis, California Institute of Technology (1969).
17. A. Ralston, H. S. Wilf, eds. Mathematical Methods for Digital Computers, Vol. II (Wiley, 1967), Ch. 4, p. 94 "The Givens-Householder Method for Symmetric Matrices" by James Ortega.
18. The ALLMAT routine on the IBM SHARE Library was used by Kaspar (Ref. 13) for the general solution.
19. W. R. Klein, B. D. Cook, W. G. Mayer, "Light Diffraction by Ultrasonic Gratings", Acustica 15, 67 (1965).
20. L. E. Hargrove, "Optical Effects of Ultrasonic Waves Producing Phase and Amplitude Modulation", J. Acoust. Soc. Am. 34, 1547 (1962).
21. W. R. Klein, B. D. Cook, "Unified Approach to Ultrasonic Light Diffraction", IEEE Trans. Sonics and Ultrasonics SU-14, 123 (1967).
22. A. W. Lohmann, D. P. Paris, "Computer Generated Spatial Filters for Coherent Optical Data Processing", Appl. Opt. 7, 651 (1968).
23. A. J. MacGovern, J. C. Wyant, "Computer-Generated Holograms for Testing Optical Elements", Appl. Opt. 10, 619 (1971).
24. T. S. Huang, "Digital Holography", Proc. IEEE 59, 1335 (1971).
25. B. R. Brown, A. W. Lohmann, "Complex Spatial Filtering with Binary Masks", Appl. Optics 5, 967 (1966).
26. B. R. Brown, A. W. Lohmann, "Computer Generated Binary Holograms", IBM J. Res. Develop. 13, 160 (1969).
27. R. E. Haskell, B. C. Culver, "New Coding Technique for Computer Generated Holograms", Appl. Opt. 11, 2712 (1972).
28. W. H. Lee, "Sampled Fourier Transform Hologram Generated by Computer", Appl. Opt. 9, 639 (1970).

29. C. B. Burckhardt, "A Simplification of Lee's Method of Generating Holograms by Computer", *Appl. Opt.* 9, 1949 (1970).
30. L. B. Lesem, P. M. Hirsch, J. A. Jordan, Jr., "The Kinoform: A New Wavefront Reconstruction Device", *IBM J. Res. Develop.* 13, 150 (1969).
31. D. Kermisch, "Image Reconstruction from Phase Information Only", *J. Opt. Soc. Am.* 60, 15 (1970).
32. D. C. Chu, J. R. Fienup, J. W. Goodman, "Multiemulsion On-Axis Computer Generated Hologram", *Appl. Opt.* 12, 1386 (1973).
33. M. C. King, A. M. Noll, D. H. Berry, "A New Approach to Computer-Generated Holography", *Appl. Opt.* 9, 471 (1970).
34. T. Yatagai, "Three-Dimensional Displays Using Computer-Generated Holograms", *Opt. Comm.* 12, 43 (1974).
35. P. L. Ransom, "Synthesis of Complex Optical Wavefronts", *Appl. Opt.* 11, 2554 (1972).
36. A. VanderLugt, F. B. Rotz, A. Klooster, Jr., "Character Reading by Optical Spatial Filtering", in Optical and Electro-Optical Information Processing, ed. J. T. Tippett et al. (MIT Press, 1965), pp. 125-141.
37. A. W. Lohmann, H. W. Werlich, "Holographic Production of Spatial Filters for Code Translation and Image Restoration", *Phys. Lett.* 25A, 570 (1967).
38. J. L. Harris, Sr., "Image Evaluation and Restoration", *J. Opt. Soc. Am.* 56, 569 (1966).
39. F. C. Billingsley, "Applications of Digital Image Processing", *Appl. Opt.* 9, 289 (1970).
40. H. C. Andrews, Computer Techniques in Image Processing (Academic Press, 1970).
41. B. R. Hunt, H. J. Trussel, "Recent Data on Image Enhancement Programs", *Proc. IEEE* 61, 466 (1973).

42. J. W. Goodman, Introduction to Fourier Optics (McGraw-Hill, 1968), Ch. 7, "Spatial Filtering and Optical Information Processing".
43. W. T. Cathey, Optical Information Processing and Holography (Wiley, 1974). Ch. 7, "Spatial Filtering".
44. H. R. Arsenault, "Conditions for Space Invariance in Optical Data Processors Used with Coherent or Noncoherent Light", Appl. Opt. 11, 2228 (1972).
45. H. W. Bode, Network Analysis and Feedback Amplifier Design (Van Nostrand, 1945), Ch. 14, "Relations between real and imaginary components of network functions".
46. S. I. Ragnarsson, "A New Holographic Method of Generating a High Efficiency, Extended Range Spatial Filter with Application to Restoration of Defocussed Images", Physica Scripta 2, 145 (1970).
47. G. W. Stroke, "Image Sharpening by Holography", in Optical and Acoustical Holography, ed. E. Camatini (Plenum, 1972), p. 411.
48. A. A. Sawchuk, "Space-Variant Image Restoration by Coordinate Transformation", J. Opt. Soc. Am. 64, 138 (1974).
49. J. W. Goodman, Ibid., p. 16, eq. 2-16.
50. C. W. Helstrom, "Image Restoration by the Method of Least Squares", J. Opt. Soc. Am. 57, 297 (1967).
51. A. Papoulis, Probability, Random Variables and Stochastic Processes (McGraw-Hill, 1965), Ch. 11, "Linear Mean-Square Estimation".
52. J. R. Pierce, Symbols, Signals, and Noise: The Nature and Process of Communication (Harper, 1961), p. 175.
53. A. Maréchal, P. Croce, "Un filtre de fréquences spatiales pour l'amélioration du contraste des images optiques", Compte Rendus 237, 607 (1953).
54. Jumpei Tsujiuchi, "Correction of Optical Images by Compensation of Aberrations and by Spatial Frequency Filtering" in Progress in Optics, Vol. II, ed. E. Wolf (North-Holland, 1963), Ch. 4.

55. A. VanderLugt, "Signal Detection by Complex Spatial Filtering", IEEE Trans. Info. Theory 6, 139 (1964).
56. G. W. Stroke, R. G. Zech, "A Posteriori Image-Correcting 'Deconvolution' by Holographic Fourier-Transform Division", Phys. Lett. 25A, 89 (1967).
57. G. W. Stroke, F. Furrer, D. R. Lamberty, "Deblurring of Motion-Blurred Photographs Using Extended-Range Holographic Fourier-Transform Division", Optics Comm. 1, 141 (1969).
58. G. A. Krusos, "Restoration of Radiologic Images by Optical Spatial Filtering", Opt. Engr. 13, 208 (1974).
59. G. W. Stroke, M. Halioua, "Attainment of Diffraction Limited Imaging in High Resolution Electron Microscopy by a posteriori Holographic Image Sharpening, I.", Optik 35, 50 (1972).
60. D. A. Tichenor, "Extended Range Spatial Filters for Image Deblurring", Ph.D. Thesis, Stanford University Center for Systems Research (May 1974).
61. J. W. Goodman, H. B. Strübin, "Increasing the Dynamic Range of Coherent Optical Filters by Means of Modulating Gratings", J. Opt. Soc. Am. 63, 50 (1973).
62. C. R. Mansfield, EG&G, Inc., Los Alamos (personal communication).
63. S. Lowenthal, P. Chavel, "Reduction of the Number of Samples in Computer Holograms for Image Processing", Appl. Opt. 13, 719 (1974).
64. G. B. Anderson, T. S. Huang, "Errors in Frequency-Domain Processing of Images", AFIPS 1969 Spring Joint Computer Conference 34, p.173.

CHAPTER III

THE MODULATED GRATING HOLOGRAM

The modulated grating hologram is a computer-generated, off-axis hologram made by a multiple exposure holographic technique. This hybrid digital-optical method combines the flexibility of the computer with the advantages of optical methods in the generation of large area holograms. Unlike other methods, the modulated grating hologram can be synthesized in virtually any holographic medium. In applications such as spatial filtering, where dynamic range is of great importance, the choice of a low-noise, high-efficiency medium such as dichromated gelatin will allow the synthesis of holographic spatial filters of exceptionally large dynamic range.

3.1 Basic Idea--Multiple Exposure Wavefront Synthesis

The basic idea in the synthesis of modulated grating holograms is illustrated in Figure 1. Several wavefronts of fixed phase but varying amplitude may be recorded by multiple exposures in a holographic medium. The functions $w_n(x,y)$ are assumed to be real and non-negative since they represent wavefronts obtained from computer-generated masks which control only the intensity of the transmitted wave. The phase ϕ_n is set independently for each exposure. If the total exposure does not exceed the linear range of the holographic medium, then the recorded wavefront may be taken as a linear superposition of the several fixed-phase wavefronts. The reconstruction is illustrated in Figure 2. The diffracted wave is now represented by the complex function $\tilde{w}(x,y)$.

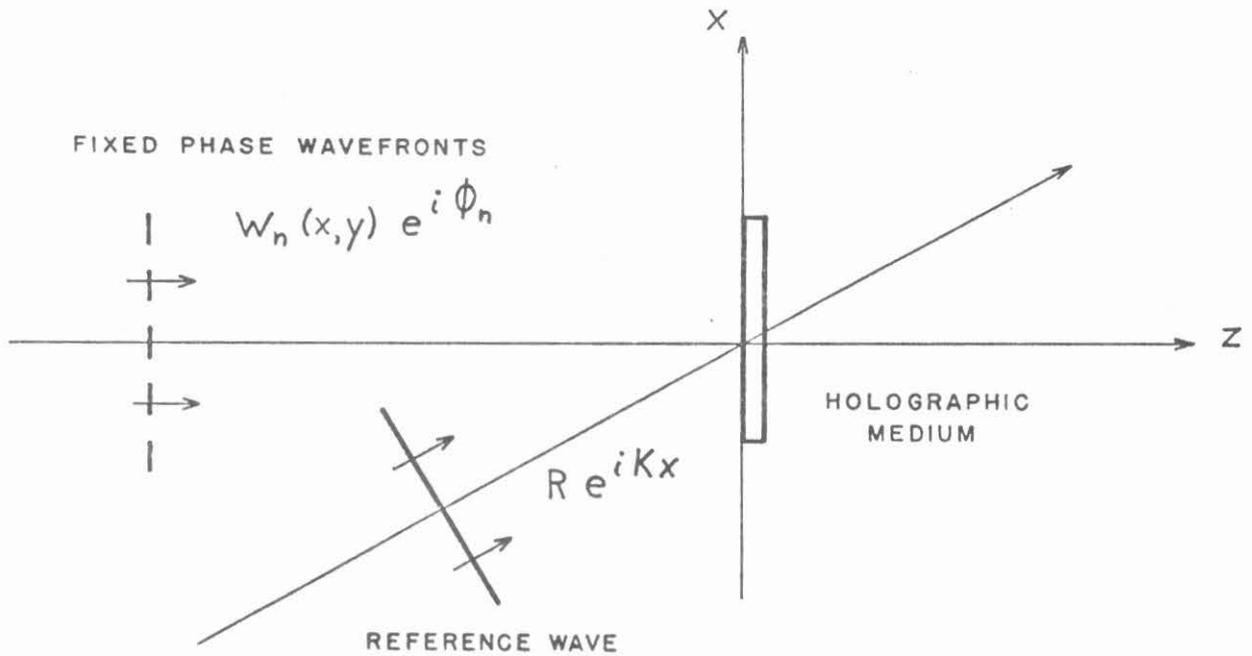


Fig. 3.1. Recording of hologram by multiple exposure of several fixed phase wavefronts.

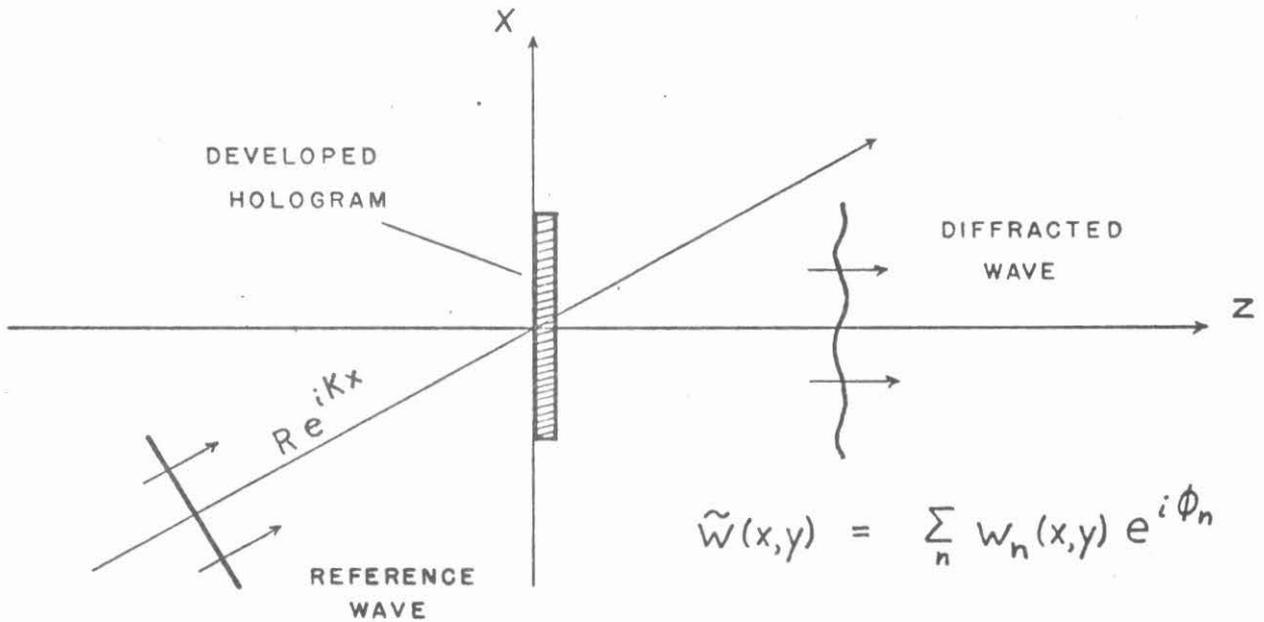
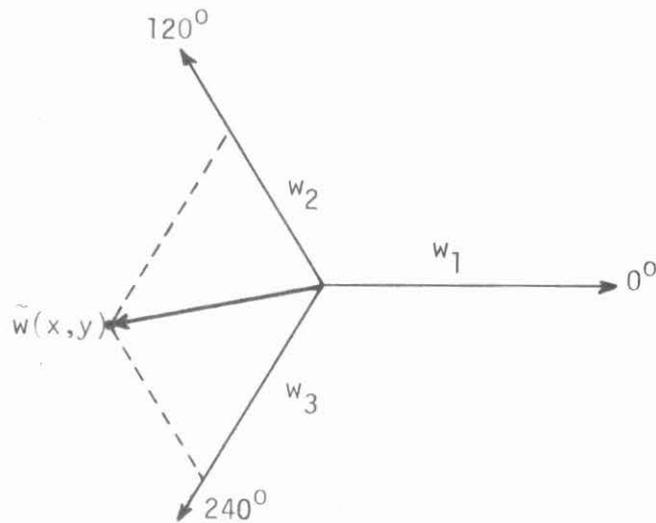


Fig. 3.2. Reconstruction gives wave of desired phase and amplitude.

By appropriate choice of the functions $w_n(x,y)$ and phases ϕ_n one may construct a wave of any desired phase and amplitude. If $\tilde{w}(x,y)$ is a real function (positive and negative values) as in certain classes of spatial filters, then only two functions w_n are needed, with $\phi_n = 0^\circ$ and 180° . If $\tilde{w}(x,y)$ is an arbitrary complex function, then three phases are needed, $\phi_n = 0^\circ, 120^\circ$ and 240° . The decomposition of a complex function into three fixed phase functions is illustrated by the phase diagram, Figure 3. The significance of this simple mathematical relationship to the synthesis of holograms was pointed out by Burckhardt⁽¹⁾. It was, in fact, an earlier paper by Burckhardt and Doherty which was the inspiration for this work⁽²⁾. An algorithm for decomposition of a complex function into its three components is given in Appendix C (subroutine COMPON).



$$\tilde{w}(x,y) = w_1(x,y) + w_2(x,y)e^{i2\pi/3} + w_3(x,y)e^{i4\pi/3}$$

Fig. 3.3. Three fixed phase wavefronts are minimum necessary to generate an arbitrary phase and amplitude wave.

The first problem one must consider in any scheme to generate a hologram by the above method is that of phase aberrations in the wavefronts $w_n(x,y)$. These wavefronts are obtained from computer-generated masks, which may introduce severe phase distortions in the transmitted wave. In an earlier paper⁽³⁾ (Appendix A) several methods were suggested for eliminating phase distortions, including use of a liquid gate and a contact printing method. The latter method is illustrated in Figure 4. The transmission mask for each exposure is placed in direct contact with the holographic medium. If the angle between the two grating-forming waves is small, then each wave suffers the same phase delay due to mask aberrations, and there is no shift in the recorded grating.

The recording process may be thought of as the superposition of two or three component gratings in the same hologram, the amplitude of each grating being independently modulated by its corresponding mask function $M_n(x,y)$. The result is a hologram with a basic grating or

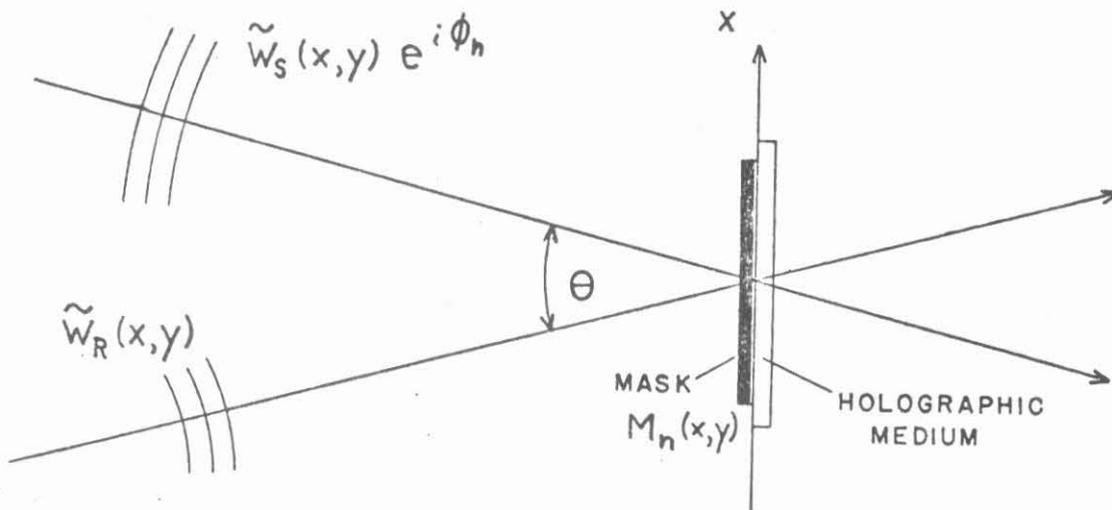


Fig. 3.4. Elimination of phase errors in component waves by contact printing.

"carrier" which is generated optically and a modulation function or "envelope" which is determined by the masks. Several advantages over previous digital holographic methods are apparent:

1. There is complete freedom in the choice of holographic medium. We are not confined to a thin amplitude medium or any special emulsion. Choice of a high-efficiency, low-noise medium such as dichromated gelatin will allow the synthesis of digital holographic spatial filters of exceptionally large dynamic range. A photo-resist medium might be used if one is interested in making a surface relief digital hologram suitable for mass production by an embossing method.⁽⁴⁾
2. The basic grating is generated optically, relieving the computer of the burden of drawing details the size of each fringe. The maximum information capacity of any given plotting device may then be applied to the generation of the modulation functions. This can result in considerable savings in computer time and plotter information capacity in the synthesis of large digital holograms. The maximum theoretical space-bandwidth efficiency is obtained, since one complex transmittance coefficient may be specified for each resolution cell of a continuous tone graphic device. In these experiments, holograms one inch square with a 3μ grating period were synthesized using mask functions with a 1024 by 1024 array of resolution spots. A hologram of the Lee-Burckhardt type with the same grating period and

number of resolution cells would be one millimeter square.

3. By varying the shape of the recording wavefronts, it is possible to incorporate any desired phase function in the hologram without recomputation of the masks. The spherical waves shown in Figure 4 will produce a hologram with a simple lens phase function. If the hologram is to be used in an optical processor, this will allow the elimination of one lens.

A simple analysis may be presented for the recording and reconstruction process in a thin amplitude hologram. We assume an ideal hologram material in which the amplitude transmittance of the developed hologram at each point (x,y) is proportional to the total exposure incident at that point during recording.

$$T_A(x,y) = \alpha \sum_n E_n(x,y) \quad (3.1)$$

Each component exposure is the product of exposure time τ and intensity

$$E_n(x,y) = \tau |M_n \tilde{w}_S e^{i\phi_n} + M_n \tilde{w}_R|^2 \quad (3.2)$$

The mask functions $M_n(x,y)$ are considered real, since any phase factor does not affect the intensity. The total transmittance can be written as the sum of three terms

$$\begin{aligned}
 T_A(x,y) = & \underbrace{\alpha \tau (w_S^2 + w_R^2) \sum_n M_n^2}_{\text{carrier}} \\
 & + \alpha \tau \tilde{w}_S^* \tilde{w}_R \sum_n M_n^2 e^{-i\phi_n} \\
 & + \alpha \tau \tilde{w}_S \tilde{w}_R^* \sum_n M_n^2(x,y) e^{i\phi_n} \quad (3.3) \\
 & \underbrace{\hspace{10em}}_{\text{modulation}}
 \end{aligned}$$

The three terms in the transmittance will be separable in an optical system provided the bandwidth of the modulation does not exceed one-half the carrier frequency. As an example, consider the hologram of Figure 5. The three terms in the transmittance function result in diffraction of three waves. The desired wave propagates along the z-axis and may be separated from the other waves if its angular spectrum does not overlap that of the other waves.

A similar analysis may be carried out for other types of holographic media. In the case of thick holograms the Bragg condition may result in cancellation of all but the first-order wave, and if the grating is very thick there may be a band-limiting effect on the first-order wave.

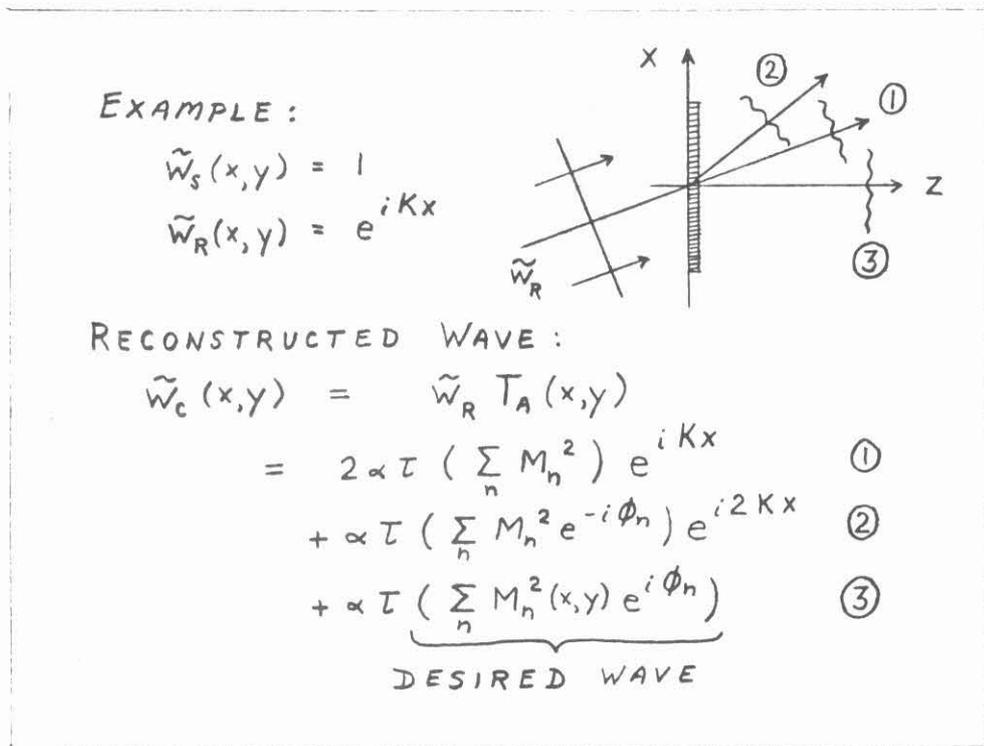


Fig. 3.5. Diffracted waves from a modulated grating hologram in an ideal thin amplitude medium.

3.2 System for Generation of Modulated Grating Holograms

3.2.1 Electronically Stabilized Interferometer

Synthesis of the modulated grating hologram requires that the recording wavefronts be stable in both amplitude and phase over long periods. A schematic of the hologram recording system used in these experiments is shown in Figure 6.

Holograms were recorded on 4" by 5" glass plates with wave curvatures 580 mm and -1740 mm at the hologram plane, giving holograms of 435 mm focal length. The two waves were set to equal intensity by a variable attenuator. The fringes recorded in the holographic medium are modulated in amplitude by a mask held in contact with the hologram. The masks are photographic plates two inches square with density patterns one inch square obtained from a CRT computer plotting device. The 25 μ mask registration tolerance is maintained by an alignment jig, a simple mechanical structure with no adjustments, mounted on the hologram plate holder.

The recording procedure requires two or three exposures, depending on the nature of the hologram transmittance function, each exposure with a separate mask and fringe phase setting. Exposure time is controlled by a shutter and digital timer. Typical exposure times for dichromated gelatin are 20 seconds with 540 $\mu\text{W}/\text{cm}^2$ at the hologram plane. Intensity is stabilized by controlling the laser tube current as shown in Figures 7 and 8. The laser is operated single mode to obtain high fringe contrast at the hologram plane. Mode stability is monitored with a spectrum analyzer, and adjustments are made with the laser etalon prior to each exposure.

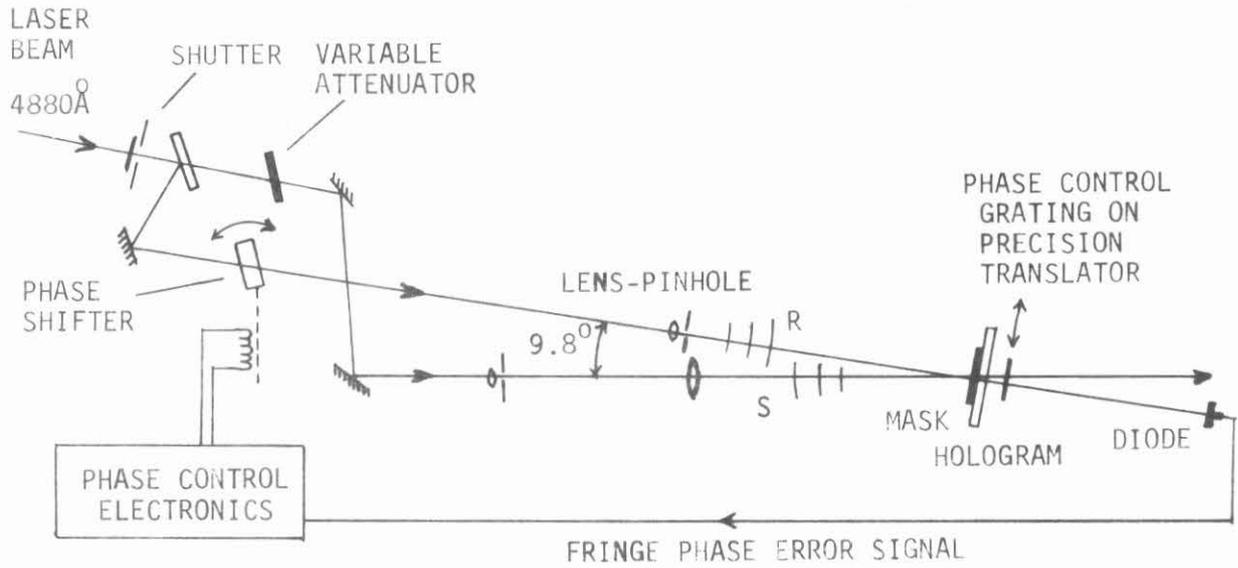


Fig. 3.6 System for Generation of Modulated Grating Holograms

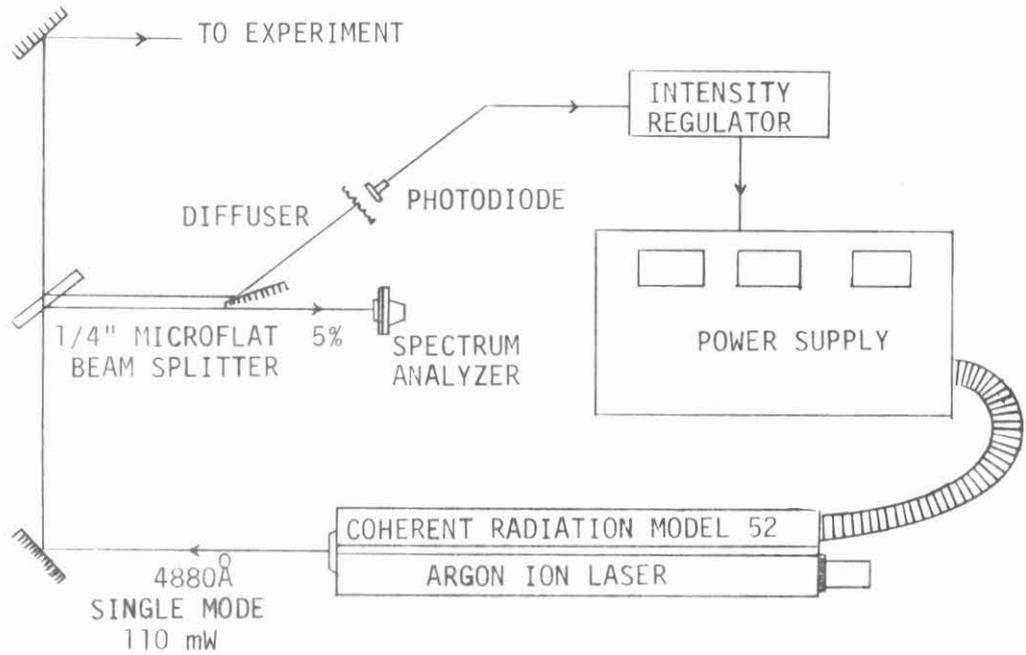


Fig. 3.7 Laser Stabilization System

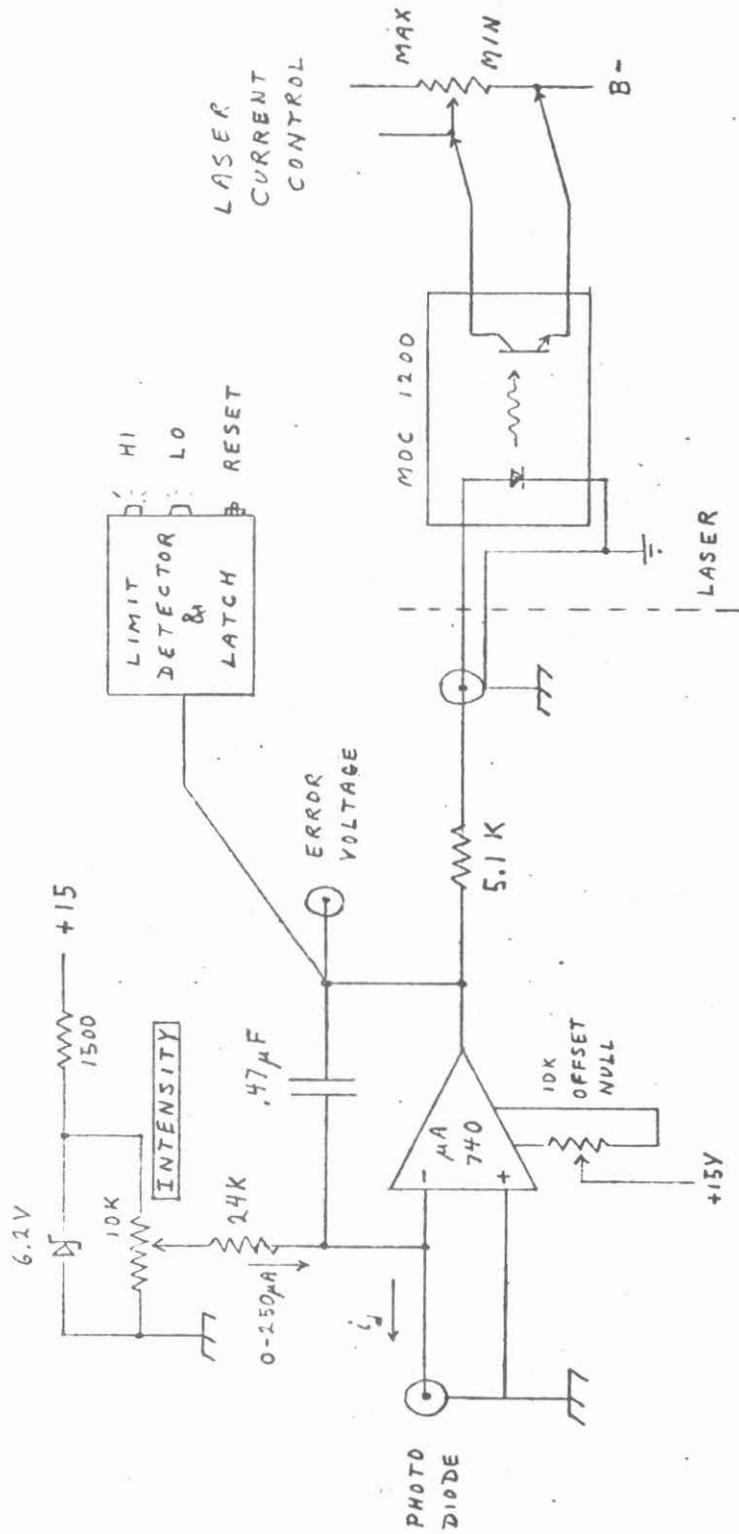


Fig. 3.8 Laser Intensity Regulator Circuit

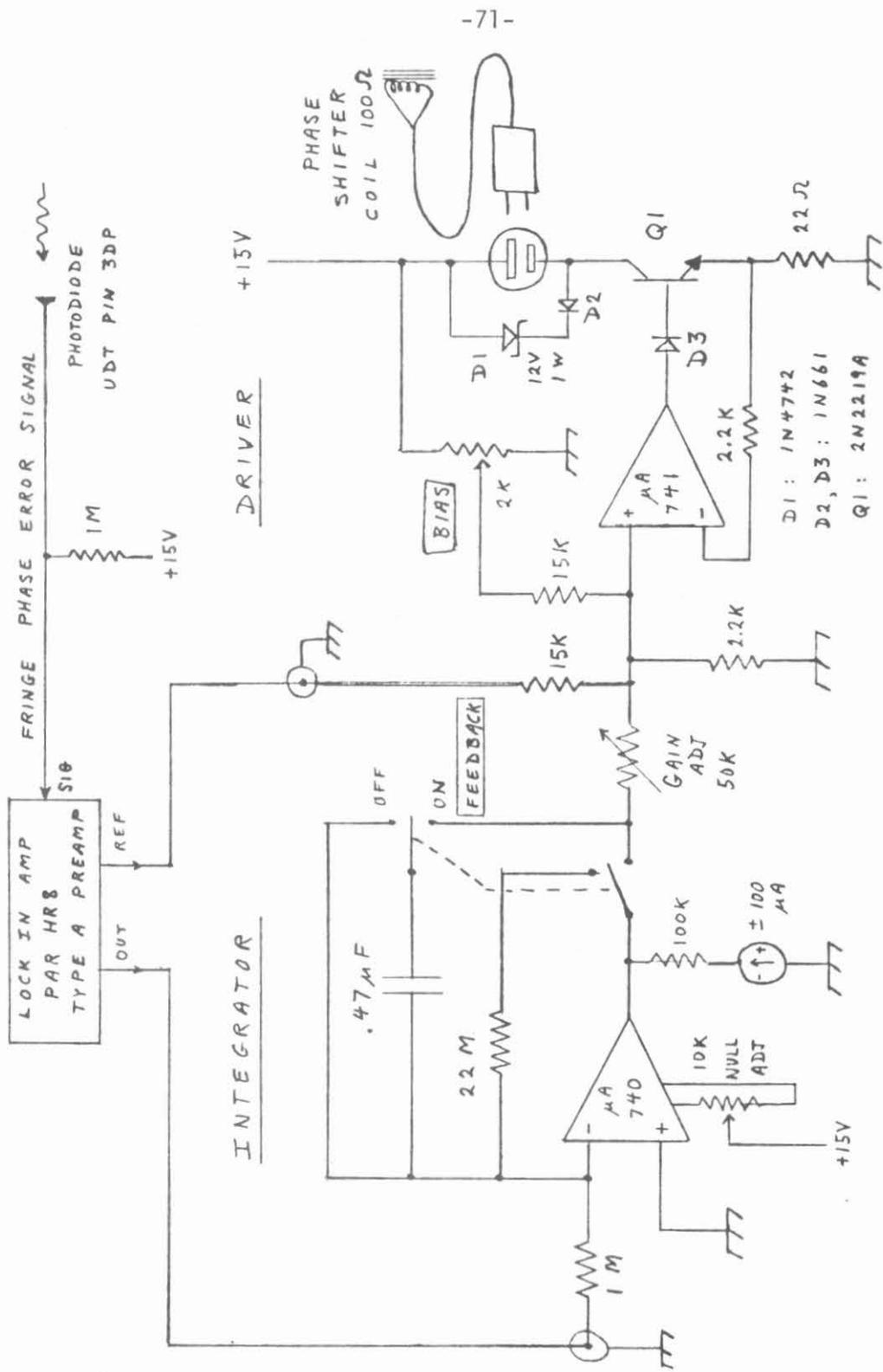


Fig. 3.9 Fringe phase control electronics circuit details

The fringe phase is controlled by rotation of a glass flat in one leg of the interferometer. This phase shifter is part of a servo-loop with the fringe phase controlled at the hologram plane by a small auxiliary grating. Details are discussed below.

All of the optical components are mounted on a cast iron table which is isolated from floor vibrations by a pneumatic suspension system. The interferometer is enclosed in a styrofoam box to eliminate air turbulence.

3.2.2 Mask Generation System

The masks used in these experiments were Kodak precision flat high resolution plates (SP0660) of the type used in the semiconductor industry for integrated circuit mask making. These are two-inch square, .060" thick glass plates with a fine grain emulsion 6μ thick. Characteristics of the emulsion and glass, including flatness specifications, are given in publications by Kodak.^(5,6)

The mask density patterns are first generated on 70 mm film (Kodak S0272) with a computer CRT plotting device.⁽⁷⁾ With this device we generate a 1024 by 1024 array of spots, 25μ spot spacing, and one-inch square total array area. Each spot has 256 density levels corresponding to a density range of 1.2 on the film.

The mask patterns are transferred from the film to the high resolution plates by contact printing. This is necessary to

- put the mask on a suitable substrate for use in the interferometer
- extend the density range 0-3D

- insure proper mask alignment
- eliminate raster in the recorded patterns.

Use of a computer controlled microdensitometer would allow direct writing of the masks and greatly simplify the above procedure.

Generation of the tapes from which the masks were made was done on an IBM370 at the Caltech Computing Center. The programs are listed in Appendix C.

3.2.3 Fringe Phase Control System

Proper recording of the multiple exposure hologram requires a stable, accurate setting of the fringe phase at the hologram plane during exposure. The required fringe stability is obtained by making the phase shifter part of a servo-loop, with the phase being controlled by a small grating just behind the hologram plate (Fig. 6). This grating causes a mixing of the two waves which pass just above the edge of the mask. Large, high contrast fringes are projected onto a photodiode. A high-frequency, small-amplitude oscillation is superimposed on the phase shifter driving signal, and this oscillation is detected in the photodiode current by a lock-in amplifier. When the projected fringe intensity is minimum, the error signal will be null. Deviation of the fringe phase to either side of this null point will result in an error signal of the proper polarity to return the system to equilibrium. Locking onto the fringe minimum in this manner insures that the phase setting is not affected by variations in the fringe intensity arising from fluctuations in the laser or other system instabilities. Details of the phase control electronics are given in Figure 9.

Other systems for fringe stabilization have been presented.^(8,9) In these systems the fringes at the hologram plane are magnified by a microscope objective and projected onto a photomultiplier tube. The very small amount of light power available from one fringe at the hologram plane results in signal-to-noise problems and requires a sensitive photomultiplier. The essential difference in our system is the use of a grating instead of a microscope objective. The light intensity available at the photodetector is the same as at the hologram plane, and a simple silicon photodiode may be used to sense the fringe phase. The disadvantage of the grating method is that each change of the interferometer configuration requires exposure of a new grating.

The control gratings used in this experiment were half-inch squares of dichromated gelatin plate. Exposure was adjusted to give projected fringes of maximum contrast.

Fringe spacing at the hologram plane is 3μ , so a displacement of the control grating over 1.5μ will shift the phase of the recorded fringes by 180° . These small displacements are produced by a differential spring mechanism which is mounted directly on the hologram plate holder so as to minimize any error resulting from thermal gradients or mechanical stresses.

3.2.4 Phase Accuracy of the Interferometer

In order to test the phase errors introduced by the interferometer and masks, several gratings were made with double exposures of equal amplitude and opposite phase. The amplitude of the resultant grating will then be a measure of the phase error, as shown in Figure 10. Variations in the amplitude of the component gratings will also give a non-zero resultant. These variations are due mostly to instabilities in the laser beam mode structure. Intensity at the center of the beam showed fluctuations of about 10% in spite of regulation of the total beam power to better than 0.1%. A 10% error in the amplitude of

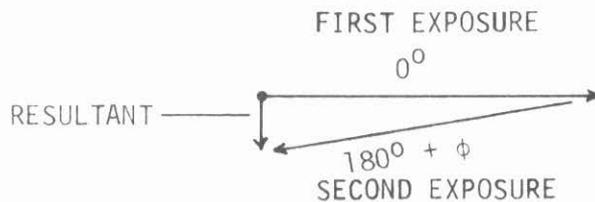


Fig. 3.10 Double exposure test for phase accuracy of fringes.

one component will be equivalent to about 6° phase error.

Figure 11 shows the results of measurements from several double exposure gratings made with various phase settings. The phase shift is assumed proportional to the micrometer setting on the differential spring translator. The dots are for gratings made with no mask. Grating amplitude was computed as the square root of the measured diffraction efficiency. The plus or minus sign ambiguity was resolved in favor of the best fit to a straight line through all the data points.

Most of the points seem to fall within 10^0 of the best fitting straight line. Direct measurements on the interferometer using a second grating and lock-in also indicated about 10^0 error in the resetability of the phase. The error seems to be the result of hysteresis in the action of the springs in the translator. A more careful design could perhaps improve this resolution.

A double exposure was made with a mask substrate in position. The mask was inverted for the second exposure so that the effect of any wedge in the mask substrates would be doubled. Figure 12a is a photograph of the diffracted light from the resulting hologram. A piece of black tape was placed across the lower third of the mask area. The bright areas at the top and bottom of the hologram show the gratings formed by the first and second exposures alone. The light from the

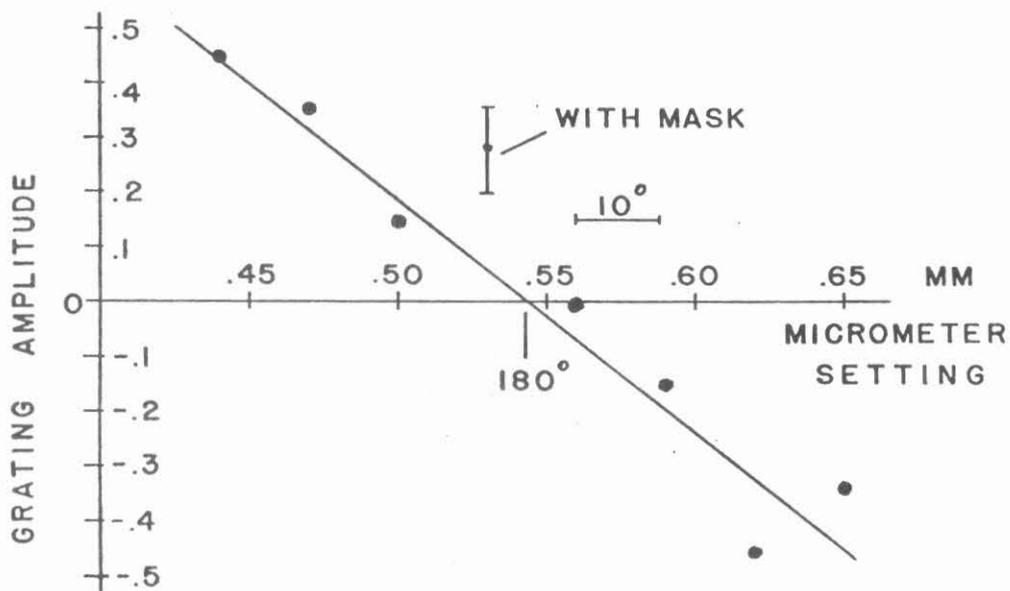
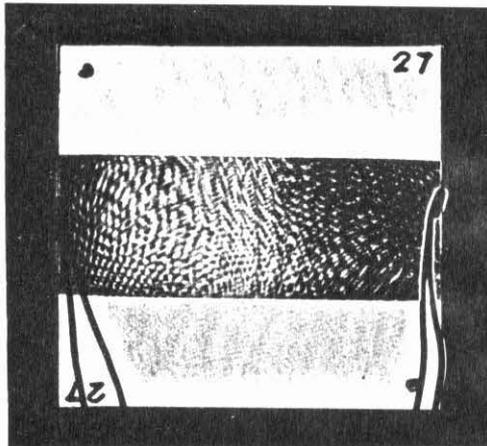


Fig. 3.11 Grating amplitude for several double exposure holograms showing interferometer and mask phase errors.

central area is the result of incomplete cancellation of the two exposures and is an indication of phase errors in the mask. Inspection of the hologram in Figure 12 shows some points in the central area to be about as bright as the top and bottom reference areas, indicating a phase error of about 60° . The accuracy of the interferometer on this hologram was good, as shown by the nearly complete cancellation of the grating outside the mask area.

a)



b)



Fig. 3.12 a) Double exposure with inverted mask substrate showing result of mask phase aberrations. b) Interferogram of the same mask.

Measurements of the diffraction efficiency were made at several points in the central area of the hologram in Figure 12. The area of the laser beam used to probe the hologram was much larger than the fine structure which is evident in this picture, so the measurement must be considered a local average of the diffraction efficiency. The mean and standard deviation for this set of measurements is plotted in Figure 11.

All the holograms in this figure were exposed and processed under the same conditions (see Dichromated Gelatin Process II in Appendix G for details).

3.2.5 Mask Phase Errors

It is apparent that the largest source of phase error in the recorded fringes is the mask substrate. Results similar to Figure 12 were obtained using masks both with and without emulsions.

An analysis of the fringe phase errors due to the mask substrate is illustrated in Figure 13. The two rays A and B, which form the fringes at any given point in the hologram, enter the mask substrate at slightly different points. The fringe phase error will depend on the flatness and polish of surface 1. The phase error due to the surface figure or flatness is given by

$$\phi^0 = N \theta T \cdot 360^0 \quad (3.4)$$

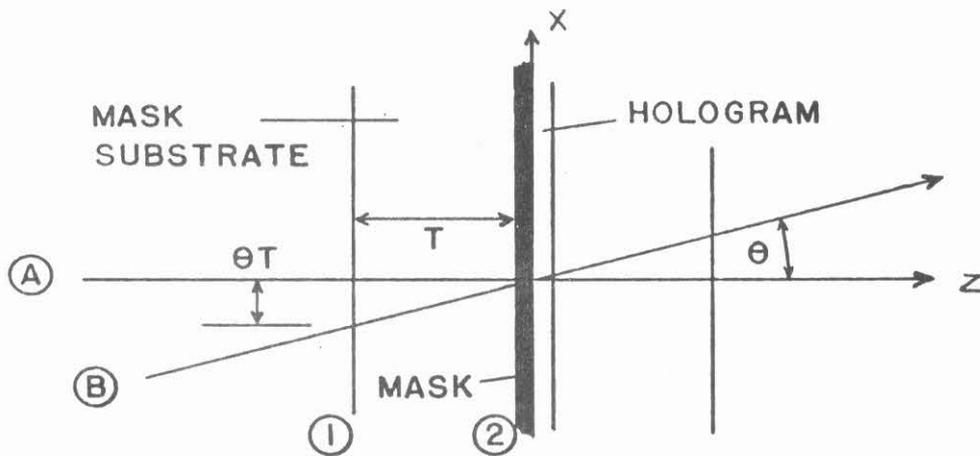


Fig. 3.13 Fringe Phase Error Due to Mask Substrate

where n is the number of fringes per unit distance in the x direction due to surface 1, θ is the beam angle in radians, and T is the mask substrate thickness. Figure 12b shows a transmission interferogram of the two-inch square mask. If we assume that the fringes are due entirely to surface 1, and assume a maximum of five fringes per inch, with $\theta = 10^\circ$ and $T = .060"$, the above formula will give 19° as the maximum phase error. Interferograms of several other possible mask substrates are shown in Appendix D.

The phase error due to imperfections in the polish on surface 1 may be calculated by

$$\phi^\circ = (n-1) \frac{h}{\lambda} \cdot 360^\circ \quad (3.5)$$

where h is the surface height difference for two points separated by a distance θT , n is the refractive index of the mask substrate, and λ is the wavelength in air. A table of surface smoothness for various polished glass specimens is given in the book by Holland.⁽¹⁰⁾ For a fire-polished microscope slide $h_{\text{rms}} = 56\text{\AA}$. If we set $h = 100\text{\AA}$, $n = 1.52$ and $\lambda = 4880\text{\AA}$, equation 5 predicts $\phi = 3.8^\circ$. An h of 1560\AA would be required to produce a phase error of 60° . The phase errors in the hologram of Figure 12 seem to be the result of ripples in the back surface of the mask substrate which are too small to be resolved in the interferogram.

Another possible source of fringe phase error may be failure of the mask to rest flat against the hologram surface due to irregularities in the gelatin. The phase error in this case is

$$\phi^0 = (n-1)(\theta T \delta / \lambda) \cdot 360^0 \quad (3.6)$$

where δ is the angular misalignment in radians. Considerable care is taken in the manufacture of high resolution plates for semiconductor mask making to insure that the gelatin layer is uniform and flat,⁽¹¹⁾ so this should not be too much of a problem. Assuming a worst case deviation of 5μ across a 50 mm plate, we get $\delta = 10^{-4}$ and $\phi = 10.2^0$.

The severity of the phase aberrations in these masks makes the synthesis of a precision spatial filter impossible. Further experimentation is needed with various substrates to find one that has acceptable surface quality and is available at reasonable cost.

3.2.6 Suggested Improvements in the System

Synthesis of high quality spatial filters with the modulated grating hologram will require further development of the apparatus and procedures described above. Suggested here are several improvements which could easily be made in a redesign of the system and some possible solutions to the mask phase error problem.

The speed and accuracy of the fringe stabilizer could be improved. The loop response time for fringe locking is about a third of a second, limited by mechanical resonances in the glass flat phase shifter. This makes phase control difficult with the more sensitive holographic media. A lightweight mirror on a piezoelectric device would be much faster, allowing shorter hologram exposure times.

The interferometer phase accuracy which is currently limited by mechanical problems in the grating translator could be improved by mounting the grating on a sensitive capacitance micrometer.⁽¹²⁾ Such devices have been used to measure displacements as little as 10^{-11} mm.

Location of the control grating behind the hologram plane is inconvenient for exposure of holographic media which scatter or distort the incident waves. Placement of a small mirror or beam splitter in front of the hologram with the control grating off to the side would alleviate this problem.

Exposure intensity stability could be improved by use of a laser with a more stable mode structure or by placing the photodiode so as to regulate intensity at the center of the hologram plane.

The most serious problem seems to be phase errors introduced by the mask substrates. Further testing of substrates is necessary to find one with acceptable phase error available at a reasonable price. The trade-off is between thickness and flatness. Thin substrates are more difficult to grind flat. In the system described above, a substrate of thickness .060" would require surface flatness 1/4 fringe per inch (in transmission) to reduce the phase error to 1° . A very thin substrate, say .005", would require a flatness of only 3 fringes per inch for the same phase accuracy. A fire polish would probably provide adequate surface smoothness.

Another possible solution to the problem might be the use of a liquid gate, either at the hologram plane, or with the mask separate from but imaged onto the hologram plane. Precise index matching could be obtained by mixtures of high index and low index fluids.⁽¹³⁾ Use of liquid gate requires that the mask substrate be thin and homogeneous. A .060" glass plate with index variations $\delta n = 10^{-5}$ would show a phase error of 11° .

Perhaps a set of high quality reusable substrates is the answer. An antireflection coating would eliminate any problems with multiple reflections. The use of a mask medium such as photoresist⁽¹⁴⁾ or a bleachable dye in gelatin suspension would eliminate the film grain noise in the masks. Dichromated gelatin holograms made with such masks might be the ultimate in noise-free digital holograms.

Direct writing of these masks on a computer-controlled micro-densitometer would be very desirable. This would greatly simplify the

mask making procedure and minimize problems with mask alignment and density variations.

3.3 Computer-Generated Holograms in Dichromated Gelatin

One of the major requirements of a computer-generated hologram for use as a spatial filter is large dynamic range. This range will be limited fundamentally by the maximum diffraction efficiency and the minimum noise level of the medium. The high diffraction efficiency and very low light scattering of dichromated gelatin make this a very attractive medium for spatial filters.

Synthesis of holograms in dichromated gelatin is, however, more difficult experimentally than synthesis in the more conventional holographic media. The low sensitivity of the medium requires long exposures with a high power laser. Repeatability of results depends on the utmost care in handling and uniformity of processing. The developed holograms are unstable in humidities above 80%.

3.3.1 Noise and Diffraction Efficiency in a Modulated Grating Hologram

Several holograms were synthesized in dichromated gelatin in order to test the noise and efficiency properties relating to dynamic range. Processing details are given in Appendix G. The holograms were made with a single exposure using the interferometer of Section 3.2. The mask had 17 areas of uniform density.

Measurements of the diffraction efficiency and noise were made as shown in Figure 14. The diffraction efficiency is defined as the ratio of power diffracted into the first order beam to power in the incident beam. Noise was measured with the detector just off the first order spot ($\theta = 7.8$ milliradians in Figure 14). Details of the

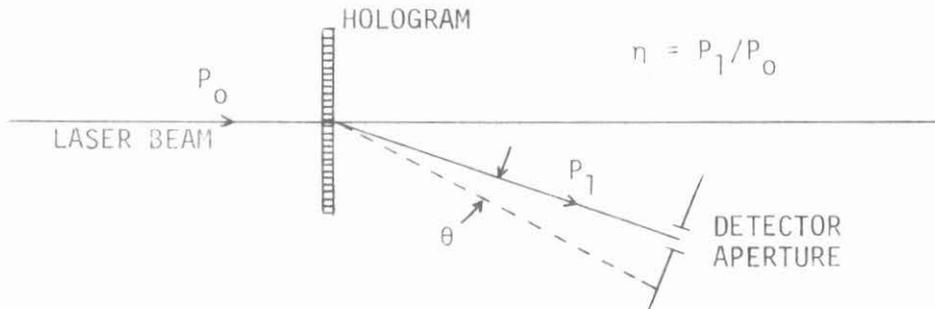


Fig. 3.14 Measurement of diffraction efficiency and noise in a holographic medium.

experimental system are given in Appendix F. Experimental data for three identically processed holograms are shown in Figure 15. The dots indicate diffraction efficiency as a function of exposure and show that with careful processing dichromated gelatin can give uniform and repeatable results.

Relative amplitude of the diffracted wave (square root of the diffraction efficiency) is plotted in Figure 15 to illustrate the linear response of the medium. The theory for an ideal holographic medium predicts a linear relationship between diffracted wave amplitude and grating modulation at low exposures. (15)

The diffracted wave amplitude will reach a saturation level due to depletion of the main beam and diffraction into other orders, even though the medium is perfectly linear. The expected curve for a linear medium will be of the form

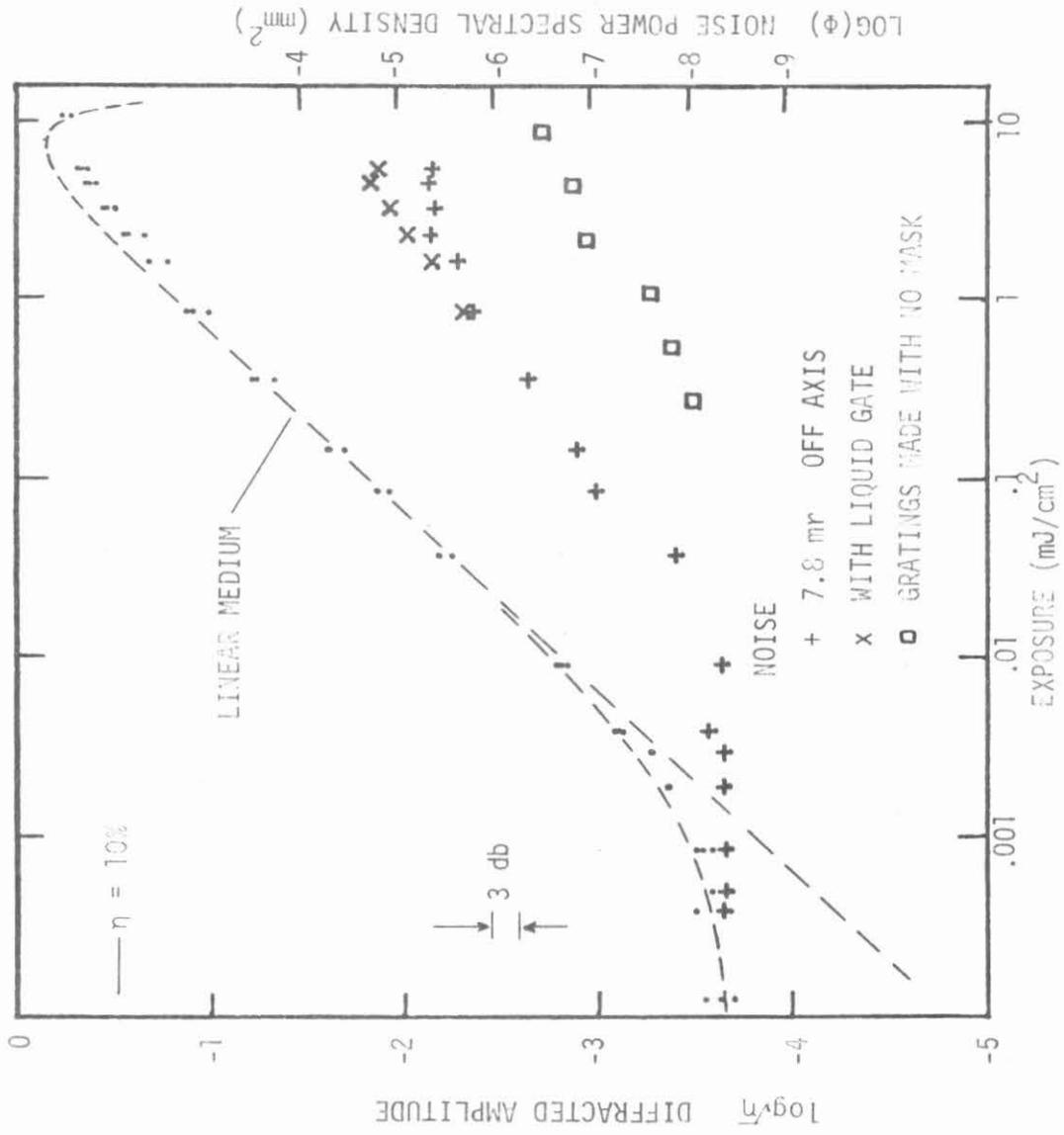


Fig. 3.15 Diffraction efficiency (η) and noise for a series of gratings in dichromated gelatin.

$$\sqrt{\eta} = \sqrt{\eta_{\max}} \sin(\alpha E) \quad (3.7)$$

The exact functional form will vary slightly, depending on the thickness of the grating, but the sine function may be taken as a good approximation, at least over the first peak in the diffraction efficiency. A plot of this function (with $\eta_{\max} = 50\%$ and α adjusted to give a best fit over the linear region) is shown as a dashed line in Figure 15. On the logarithmic scale of this figure, a linear medium would show a fairly sharp cutoff at maximum diffraction efficiency. The more gradual falloff of the experimental data is a result of saturation of the medium for very strong exposures. The curve deviates from a straight line at low efficiencies due to a minimum noise level in the system. A curve of the form

$$\sqrt{\eta} = \sqrt{\eta_{\max}} \sin(\alpha E) + C \quad (3.8)$$

with $C = 2 \cdot 10^{-4}$ is plotted through the data points at low amplitude. The minimum noise level C is a function of the detector solid angle. A smaller detector aperture would measure the same diffraction efficiency but smaller noise level. The signal-to-noise ratio, therefore, cannot be taken as an intrinsic property of the medium, but is still useful in comparing one medium to another.

A more fundamental characterization of the medium is the noise power spectral density, Φ (lines/mm)⁻², which is defined as the fraction of incident power scattered into an aperture of unit spatial frequency. It is related to the fraction of incident power scattered

per unit solid angle by

$$\Phi(v_x, v_y) dv_x dv_y = \Phi_\Omega(\theta_x, \theta_y) d\Omega/\lambda^2 \quad (3.9)$$

where the spatial frequencies v_x, v_y are given by

$$v_{x,y} = \sin \theta_{x,y}/\lambda \quad (3.10)$$

Φ is used rather than Φ_Ω because Φ will be independent of wavelength λ in film noise models which assume a two-dimensional transmittance function such as the random checkerboard or overlapping circular grain models.⁽¹⁶⁾ For a real emulsion, Φ may vary with wavelength due to the wavelength dependent nature of the scattering from film grains in a thick emulsion.

The scale on the right side of the graph applies to the noise measurements

$$\log \Phi = \log[\eta/\delta v_x \delta v_y] \quad (3.11)$$

where η is in this case the noise equivalent of diffraction efficiency.

Noise measurements for one of the holograms in Figure 15 are shown on the graph by + symbols. Comparison of these levels to those for gratings made without a mask shows that the mask is the dominant noise source. The mask noise probably originates from film grain in the 70 mm film (Kodak S0272) on which the density patterns were written. This film has much larger grains than a typical

nolographic emulsion, and the noise from this source will be transferred to the high resolution plates and recorded in the hologram. Direct writing of the masks in a low-noise medium would alleviate this problem.

A liquid gate was formed by placing one of the high resolution plate substrates over the hologram emulsion with a drop of xylene for index matching. This had no effect on the noise levels at low diffraction efficiency, and at high efficiency actually increased the noise scattering. This result is to be expected if the noise is recorded in the hologram and is not the result of surface deformations in the gelatin. The increased scattering at high efficiencies is probably due to scattering of the diffracted wave by the glass-air interface of the liquid gate. The surface relief pattern on these glass plates which is evident in Figure 12a would be expected to add a large amount of low-frequency phase noise to the transmitted wavefront.

The low noise gratings used as a comparison in this experiment were made by varying the exposure time. The noise levels for these gratings, indicated by the squares in Figure 15, are more indicative of the dynamic range which might ultimately be obtained with these holograms. Further measurements and comparison of the dynamic range in various holographic media are presented in Chapter 4.

3.3.2 A Simple Three-Phase Hologram

The multiple exposure wavefront synthesis idea has been illustrated for the two-phase case in the double exposure hologram of Figure 12. To demonstrate the synthesis of a modulated grating hologram in the case of a general complex transmittance function requiring three fixed-phase exposures, a hologram of a simple non-symmetric object was generated.

The object consists of three bright spots on a dark background. The modulating function for this object is given by

$$\tilde{h}(x,y) = e^{i2\pi\nu_a x} + e^{i2\pi\nu_a(x+y)} + e^{i2\pi\nu_b x} \quad (3.12)$$

where x and y are spatial coordinates in the hologram plane (see Figure 16) and the spatial frequencies ν_a and ν_b are related to displacements of the bright spots from the center of the image plane

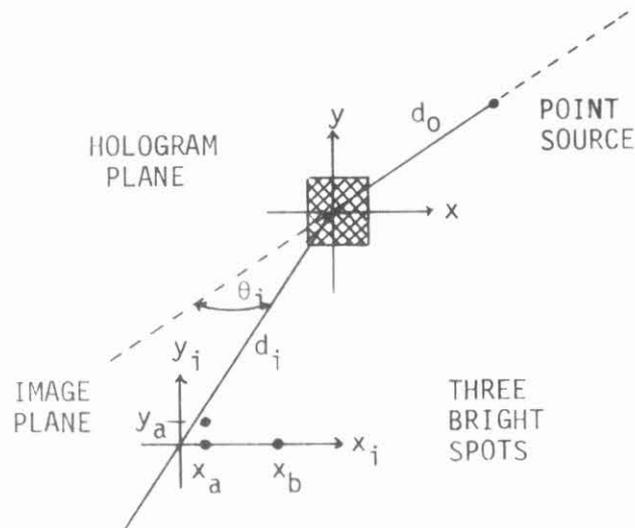


Fig. 3.16. Hologram reconstruction geometry for the three spots hologram.

by

$$v_{a,b} = x_{a,b}/\lambda d_i \quad y_a = x_a \quad (3.13)$$

The total transmittance function for the first order wave of an ideal hologram is given by

$$\tilde{T}_H(x,y) = \underbrace{\tilde{H}(x,y)}_{\text{modulation}} \underbrace{e^{iKx - ik(x^2+y^2)/2f}}_{\text{carrier}} \quad (3.14)$$

The carrier grating has a linear phase term corresponding to the off-set angle θ_i ,

$$K = k \sin \theta_i, \quad k = 2\pi/\lambda \quad (3.15)$$

and a quadratic term which is the paraxial approximation to a spherical wave of radius f . In order to focus the point source onto the image plane, the hologram focal length f must be such that

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (3.16)$$

In this case, the amplitude at the image plane is related to the modulation function by a Fourier transform,⁽¹⁷⁾

$$\tilde{n}(x_i, y_i) = \mathcal{F}\{\tilde{H}(x,y)\} \cdot e^{ik(x_i^2 + y_i^2)/2d_i} \quad (3.17)$$

Since we are interested only in the intensity at the image plane, the phase term may be ignored and we get

$$h(x_i, y_i) = \delta(x_i - x_a, y_i) + \delta(x_i - x_a, y_i - y_a) + \delta(x_i - x_b, y_i) \quad (3.18)$$

In a real optical system, the δ functions will be replaced by functions of finite width which are the transform of the hologram pupil function.

A modulated grating hologram was generated in dichromated gelatin using the function from equation 12. This complex function was computed on a 1024 by 1024 array of sample points and broken into three real functions, using the programs in Appendix C.

$$\tilde{H}(x, y) = H_1(x, y) + H_2(x, y)e^{i2\pi/3} + H_3(x, y)e^{i4\pi/3} \quad (3.19)$$

The real functions were scaled from 0 to 255 and written out as three separate frames on the CRT plotter. No attempt was made to compensate for nonlinearities in the CRT and photographic emulsions.

Contact prints of the three masks are shown in Figure 17a. These three masks were made into a hologram using the apparatus and procedures of Section 3.2. A photograph of the resulting hologram is shown in Figure 17b. The hologram is illuminated from behind with a tensor lamp, and areas where the grating is strong appear bright, due to the diffracted light. The modulated area in the center of this hologram is one inch square. The visual appearance of this hologram is even more striking. The focusing and dispersive properties of the grating cause the observer to see the pattern in a brilliant spectral color which varies from red to blue as the head is moved from side to

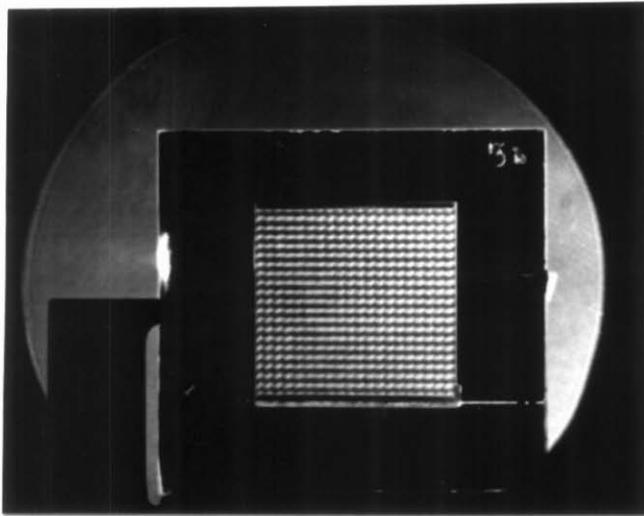
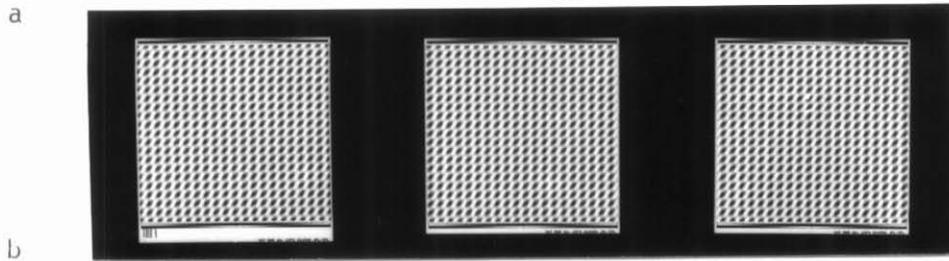
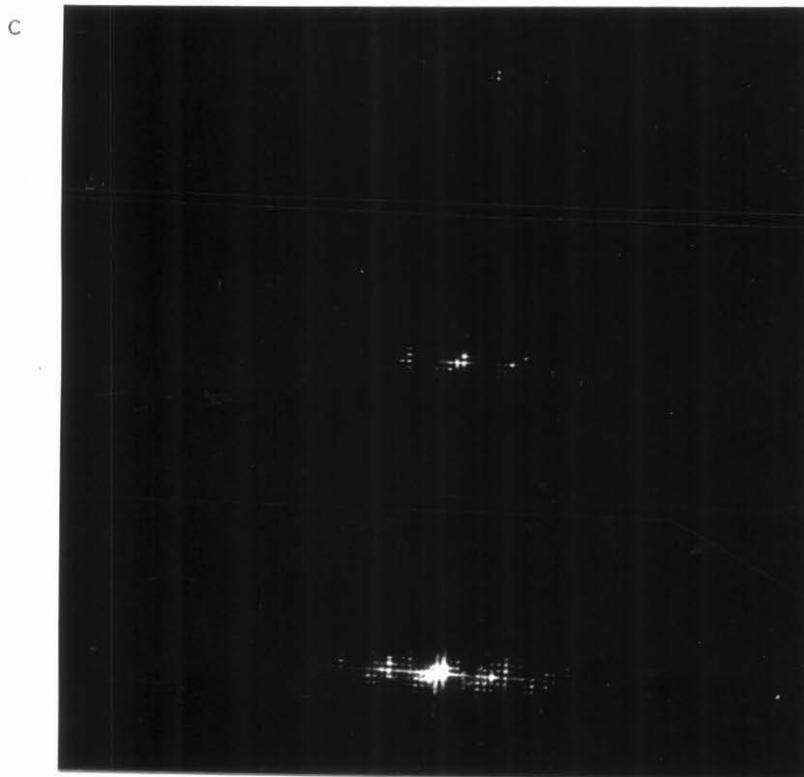


Fig. 3.17. Computer-generated masks (a) used to synthesize a modulated grating hologram (b) which projects an image of three bright spots (c). Increasing image exposure brings out intermodulation spots and other image defects.



Relative Exposure

0 db

+17 db

+27 db

side in the focal region.

The hologram was illuminated with a point source as shown in Figure 16, with an aperture placed so that only the modulated area of the hologram was illuminated. The projected image was recorded directly on high-contrast film placed at the image plane. A series of exposures was made with increasing exposure time to bring out the defects in the image. Prints of three of the negatives from this series are shown in Figure 17c. In addition to the three predicted spots, the image shows a spot at the center of the image plane and some spots which result from nonlinear mixing of the three primary spots. The central spot corresponds to zero spatial frequency in the modulating function \tilde{H} , and is the result of an imbalance in the three exposures. This imbalance is evident also from the incomplete cancellation of the grating in areas outside the mask (see Figure 17b). Errors in the phase of the recorded components can result in terms in the hologram transmittance which are the complex conjugate of the desired transmittance. These conjugate image terms appear as a faint inverted image reflected through the origin of the image plane.

The nonlinearity spots, which appear in an evenly spaced array around each of the bright spots in the image, result from nonlinearities in the computer plotting device, the mask emulsions, and the holographic medium. These nonlinearities cause the appearance of terms at spatial frequencies which are the sums and differences of multiples of the original image spatial frequencies.

Two faint spots are visible 45 mm either side of the center spot on the most heavily exposed photograph in Figure 17c. This distance in the image corresponds to 40 lines/mm at the hologram, which is the raster frequency of the array of sample spots. If each spot were a δ function, the image would be repeated in an infinite array at the x-y plane. Since the spots have a finite width, however, we would expect these higher order images to be attenuated. Inspection of the masks under a microscope shows the spots to be wide enough to merge with their neighbors, and no raster is visible at all. The faint spots in these photographs are all that remain of the higher order images.

An analysis of the nonlinearities could be carried out as follows:

The complex transmittance for the first order wave of a thick phase hologram is related to the index modulation by Kogelnick's coupled wave theory⁽¹⁸⁾ (see Appendix K),

$$\begin{aligned} \tilde{T}_H(x,y) &= i \sin\left[\frac{\pi T}{\lambda_a \cos \theta_B} n_1(x,y)\right] e^{i(\phi_S + \phi(x,y))} \\ \phi_S(x,y) &= Kx - k(x^2+y^2)/2f \end{aligned} \quad (3.20)$$

T , λ_a , and θ_B are the hologram thickness, wavelength in air, and Bragg angle of the grating, which is assumed to have fringes perpendicular to the hologram surface, and ϕ_S is the spherical phase factor for the unmodulated grating. The index modulation parameter n_1 is determined from a Fourier series expansion of the refractive index of the medium.

$$n(x,y) = \sum_{\ell=-\infty}^{\infty} \tilde{n}_{\ell}(x,y) e^{i\ell kx} \quad (3.21)$$

$$\tilde{n}_1(x,y) = n_1(x,y) e^{i\phi(x,y)} \quad (3.22)$$

The index of the medium is computed from the exposure by a Taylor series,

$$n(E) = \alpha_0 + \alpha_1 E + \alpha_2 E^2 + \dots \quad (3.23)$$

The α coefficients may be experimentally determined by methods such as those used by Fillmore and Tynan.⁽¹⁹⁾

The exposure is related to the mask transmittance function by

$$E(x,y) = \sum_n M_n^2(x,y) E_0 [1 + \cos(\phi_S + \phi_n)] \quad (3.24)$$

assuming fringes of maximum contrast. The mask transmittance functions are related to the components of the modulation function \tilde{H} by some function f which represents the nonlinearities in the CRT plotting device and the photographic mask-making processes,

$$\begin{aligned} M_n &= f(H_n) \\ &= f_0 + f_1 H_n + f_2 H_n^2 + \dots \end{aligned} \quad (3.25)$$

A simple modulation function with a few discrete spatial frequencies, such as equation 3.12, will result in a hologram transmittance function (3.20) with a multitude of spatial frequencies which are sums and differences of multiples of the original frequencies. This accounts for the location of the extraneous spots in the images of Figure 17c.

Precompensation of the nonlinearities for the entire process is not difficult. It is not necessary to determine the Taylor coefficients or any details of the process above. As long as some stable relationship exists between the diffraction amplitude of the final hologram and the number 0-255 generated by the computer, a simple routine may be used to precompensate the functions in the computer.

Appendix H presents an analysis of the nonlinearities inherent in the diffraction process and in the response of the hologram medium to a given exposure. The Taylor series method is applied to the analysis of thin amplitude, thin phase, and thick phase holograms. For each of these cases a formula is derived for the magnitude of the lowest order intermodulation products.

3.4 Summary

A new method for synthesis of digital holograms has been presented. The generation of an arbitrary complex wavefront can be accomplished by a linear superposition of at most three fixed-phase wavefronts in a multiple exposure hologram. The fixed-phase wavefronts are obtained from computer-generated transmission masks. The masks may be placed in contact with the hologram to reduce the effect of phase aberrations.

The advantages of this method for generating holograms include: 1) Arbitrary choice of recording medium, including high-efficiency thick phase media. 2) Efficient use of plotter capacity in generating large holograms with high grating frequency. 3) Variation of the recording wavefronts allowing the incorporation of a lens or other phase function in the hologram without recomputation of the masks.

An experimental system was shown for generation of modulated grating holograms. The basic requirement of the recording interferometer is a stable amplitude and a stable, controllable phase of the wavefronts at the hologram plane. Phase control may be accomplished with an electronic feedback loop, including an auxiliary grating to sense the phase at the hologram and a mechanism to shift the phase in one leg of the interferometer.

Tests with double exposure holograms showed that the major source of error in the recorded wavefronts is still phase aberrations in the masks. Synthesis of precision spatial filters will require correction of this problem. Suggestions were made for future improvements in the masks and recording system.

Holograms were synthesized in dichromated gelatin to demonstrate the generation of a digital hologram in a thick phase medium and to explore the properties of this particular medium relevant to the generation of spatial filters. DCG holograms are much more difficult to process than conventional photographic media, but the data of Figure 15 show that it is possible to get repeatable diffraction efficiency over a very large exposure range. This uniformity is essential if nonlinearities are to be precompensated in the computer.

Noise measurements were made on the holograms, and the dominant noise source was found to be the masks. Most likely this noise originates from the coarse grains in the original mask recording films. With an improved mask process, holograms of very large dynamic range could be synthesized in dichromated gelatin.

A simple three-phase hologram was synthesized to demonstrate the generation of a hologram with a general complex transmittance function. The reconstructed image shows the effects of nonlinearities and other defects in the recording process. An analysis of the nonlinearities in a modulated grating hologram is outlined, and further discussion is deferred to Appendix H.

The modulated grating hologram was designed with the spatial filtering application in mind, but it may find application in other areas as well. The generation of large holograms by this method may be useful in the synthesis of wavefronts for testing aspheric optical elements⁽²⁰⁾ or for generation of aspheric hologram lenses.⁽²¹⁾ In such situations the desired wavefront may have no focal region where a

small hologram can cover the entire wavefront. A large digital hologram may be necessary for synthesis of these wavefronts.

References

1. C. B. Burckhardt, "A Simplification of Lee's Method of Generating Holograms by Computer", *Appl. Opt.* 9, 1949 (1970).
2. C. B. Burckhardt, E. T. Doherty, "Formation of Carrier Frequency Holograms with an On-Axis Reference Beam", *Appl. Opt.* 7, 1191 (1968).
3. D. MacQuigg, "Complex Wavefront Synthesis by Multiple Exposure Holography", *Optics Comm.* 8, 76 (1973).
4. R. A. Bartolini, "Characteristics of Relief Phase Holograms Recorded in Photoresists", *Appl. Opt.* 13, 129 (1974).
5. Kodak Publication No. P-47 (1972), "Kodak High Resolution Plate".
6. Kodak Publication No. Q-35 (1971), "Physical Characteristics of Glass for Kodak Photographic Plates".
7. Video Film Converter at the JPL Image Processing Laboratory.
8. D. B. Neumann, H. W. Rose, "Improvement of Recorded Holographic Fringes by Feedback Control", *Appl. Opt.* 6, 1097 (1967).
9. H. W. Rose, H. D. Pruett, "Stabilization of Holographic Fringes by FM Feedback", *Appl. Opt.* 7, 87 (1968).
10. L. Holland, The Properties of Glass Surfaces, Chapman & Hall, London, 1964, Table 2.3, p. 119.
11. Kodak Tech Bits (1972), No. 2, p. 8., "Kodak High Resolution Plates Feature Improved Emulsion Thickness Uniformity."
12. R. V. Jones, J. C. S. Richards, "The Design and Some Applications of Sensitive Capacitance Micrometers", *J. Phys. E, Scientific Instr.* 6, 589 (1973).
13. R. W. Green, "Immersion Media for Liquid Gates in Coherent Optical Processing", Tech. Rep. ECOM-00013-92, Willow Run Laboratories, University of Michigan, 1967.
14. J. P. Kirk, G. L. Fillmore, "Model of Negative Working Photoresists as Continuous Tone Photographic Materials", *Appl. Opt.* 11, 2347 (1972).

15. R. J. Collier, C. B. Burckhardt, L. H. Lin, "Ideal Wavefront Reconstruction and Ideal Recording Material" in Optical Holography, Academic Press, 1971, Section 10.6, p. 273.
16. E. L. O'Neill, Introduction to Statistical Optics, Addison-Wesley, 1963, Chapter 7.
17. J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill, 1968, Ch. 5. Equation 3.17 is similar to Goodman's equation 5-22. Notation is changed and a few irrelevant constants are dropped.
18. H. Kogelnick, "Coupled Wave Theory for Thick Hologram Gratings", Bell System Tech. J. 48, 2909 (1969).
19. G. L. Fillmore, R. F. Tynan, "Sensitometric Characteristics of Hardened Dichromated-Gelatin Films", J. Opt. Soc. Am. 61, 199 (1971).
20. A. J. MacGovern, J. C. Wyant, "Computer-Generated Holograms for Testing Optical Elements", Appl. Opt. 10, 619 (1971).
21. D. H. Close, A. Au, A. Graube, "Holographic Lens for Pilot's Head-Up Display", Hughes Research Labs, Tech. Report to Naval Air Development Center (August 1974).

CHAPTER IV

NOISE LIMITS ON THE DYNAMIC RANGE OF HOLOGRAPHIC SPATIAL FILTERS

The dynamic range of a holographic spatial filter will be limited fundamentally by the maximum diffraction efficiency and the minimum noise level of the holographic medium. This chapter presents a brief discussion of the noise problem in spatial filters and some experimental measurements comparing several low-noise media.

In general, the dynamic range of a spatial filter will be affected by other factors as well as intrinsic noise and diffraction efficiency. Nonlinearities in the hologram medium may result in distortion products which make the usable diffraction efficiency much less than the maximum. In a coherent optical processor, noise and nonlinearities in the input medium as well as other noise sources in the system may be the limiting factor. These other problems will not be considered in detail here. Nonlinearities in a computer-generated hologram may in principle be eliminated by pre-distortion of the functions in the computer. Various techniques exist for extending the linear range of holograms made by optical methods.⁽¹⁻⁴⁾ The only fundamental limitations on the dynamic range of a spatial filter are the maximum diffraction efficiency and the minimum noise level intrinsic to the medium.

The concept of noise power spectral density in photographic media was first introduced by Jones⁽⁵⁾ as a method of characterizing photographic granularity. Since then, there have been many studies

of noise in photographic media for use in coherent optical systems⁽⁶⁻¹²⁾ and many studies which apply specifically to holographic media⁽¹³⁻²¹⁾.

The results of these studies are difficult to apply to the evaluation of holographic media for use as spatial filters. The problem is that the noise is treated as originating only from scattering of the zero-order or specular wave in the hologram. This is a valid approximation for weakly diffracting holograms, but in a strong hologram the dominant noise source is often scattering from the first-order wave due to surface irregularities in the emulsion or other noise sources of low spatial frequency. The study which comes closest to treating the problem is that of Lee and Greer,⁽¹⁷⁾ who measured signal-to-noise ratios in media for holographic memories. The low-frequency effects are not apparent in these data, probably the result of low diffraction efficiency and the presence of intermodulation noise. No attempt was made to eliminate the intermodulation noise, as this was one of the main effects being studied.

In the experiments reported here, noise measurements were made in several low-noise holographic media using single point images, so as to eliminate intermodulation products. The scattered noise light is shown to be strongly dependent on diffraction efficiency.

4.1 Dynamic Range and Signal-to-Noise Ratios

The dynamic range of a holographic spatial filter for image deblurring will be defined here as the ratio H_{\max}/H_{\min} , where H_{\max} and H_{\min} are maximum and minimum useful values of the filter transmittance function. The concept of dynamic range is normally applied to nonlinear media, in which the transmittance may vary over only a finite range without introducing unacceptable distortion in the recorded function. In a filter for which the nonlinearity has been corrected, a dynamic range may still be specified, based on the minimum acceptable signal-to-noise ratio at the output plane of the processor in which the filter is to be used.

In the processor of Figure 1, the signal is shown as a single bright point, although in general it may be a diffuse image. The noise occurs as scattered light within the image area. The hologram is assumed to have a transmittance (for the first-order wave) of:

$$\tilde{H}(x,y) e^{-iKx - ik(x^2+y^2)/2f} \quad (4.1)$$

The function \tilde{H} is the complex filter transmittance, and the phase factor represents an off-axis converging spherical wave. An analysis of the operation of this processor is presented in Appendix N.

H_{\max} for the filter will be limited by the maximum diffraction efficiency of the medium. H_{\min} for a typical deblurring filter (see Chapter II) will be the transmittance at the center of the filter. Lowering H_{\min} to achieve greater dynamic range will result in less

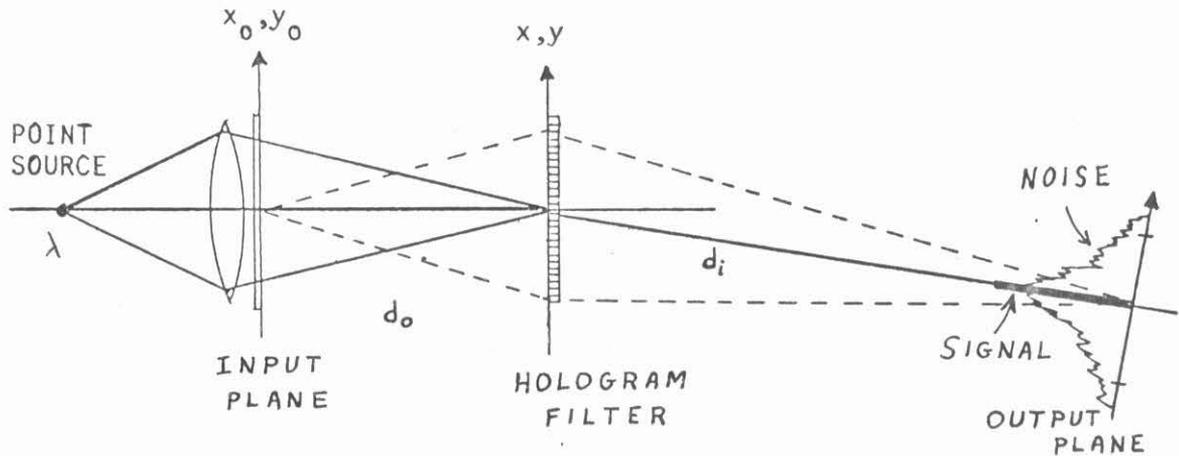


Fig. 4.1 Coherent Optical Processor showing degradation of the image (signal) by noise at the output plane.

image intensity at the output plane, and lower signal-to-noise ratio. The exact relationship between dynamic range and signal-to-noise ratio will in general depend on the particular characteristics of the filter function and image distribution. Tichenor⁽²¹⁾ has derived a relationship for a filter in which all of the scattered light is assumed to originate from the bright central spot in the filter plane. This ignores the high diffraction efficiency effects mentioned earlier.

A signal-to-noise ratio may be defined as the ratio of maximum intensity of the impulse response to average noise intensity, $I_S / \langle I_N \rangle$. A more careful definition for a coherent optical system must take account of the random fluctuations due to interference of the signal with the noise,⁽¹⁴⁾

$$I_S / \sigma = I_S / [\langle I_N \rangle^2 + 2I_S \langle I_N \rangle]^{1/2} \quad (4.2)$$

σ is the root-mean-square fluctuation in intensity due to noise in the coherent image.

Signal-to-noise ratios are useful for comparing one medium to another under similar circumstances, but do not represent a fundamental property of the holographic medium. For a point image the signal-to-noise ratio will be proportional to the hologram area. For a diffuse image the signal-to-noise ratio will depend on the image area.⁽¹⁶⁾

A more fundamental characterization of the medium is $\Phi(p,q)$, the noise power spectral density, or fraction of incident wave power scattered per unit spatial frequency, at frequency (p,q) . Complete specification of this function will allow calculation of signal-to-noise ratios in any particular situation.

4.2 Noise in Holographic Media

The sources of noise in a thick holographic medium are illustrated in Figure 2. In addition to the diffracted waves, which arise from periodic modulations in the absorption or refractive index of the medium, there will be a portion of the incident energy which is scattered randomly. This random scattering results from film grains or other inhomogeneities in the medium and from irregularities at the surface of the emulsion.

In a simple model, the total effective amplitude transmittance of the hologram can be written as:

$$\tilde{T}(x,y) = \underbrace{\tilde{T}_S(x,y)}_{\text{signal}} \cdot \underbrace{\tilde{T}_N(x,y)}_{\text{noise}} \quad (4.3)$$

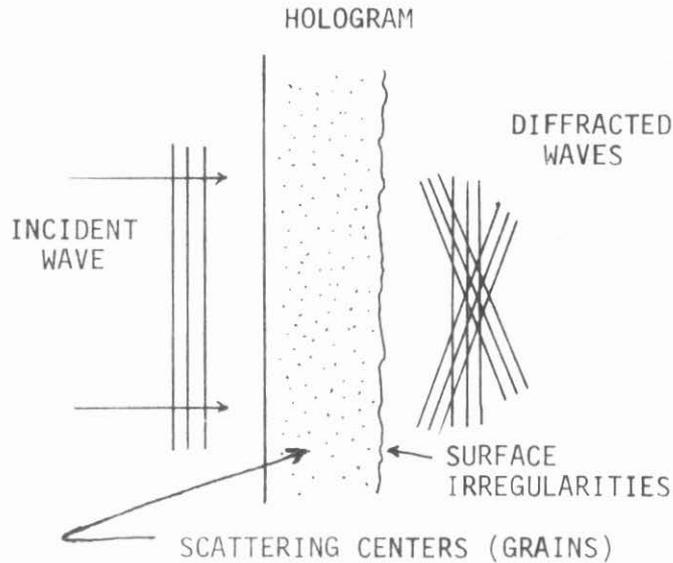


Fig. 4.2 Sources of noise in a holographic medium.

This model ignores any dependence of the noise on the signal transmittance T_S , and also ignores the finite thickness of the medium.

T_S represents the deterministic part of the hologram transmittance, and includes all diffracted orders:

$$\tilde{T}_S(x,y) = \sum_{\ell} \tilde{T}_{\ell}(x,y) e^{i\phi_{\ell}(x,y)} \quad (4.4)$$

ϕ_{ℓ} is the spherical phase factor for the ℓ^{th} order wave.

T_N includes all random deviations from T_S and is assumed to have a mean value 1:

$$\begin{aligned} \tilde{T}_N(x,y) &= 1 + \tilde{t}_N(x,y) \\ \langle \tilde{t}_N \rangle &= 0 \end{aligned} \quad (4.5)$$

Consider now the hologram imaging situation of Figure 3. A point source (pinhole) illuminates a hologram which projects an image

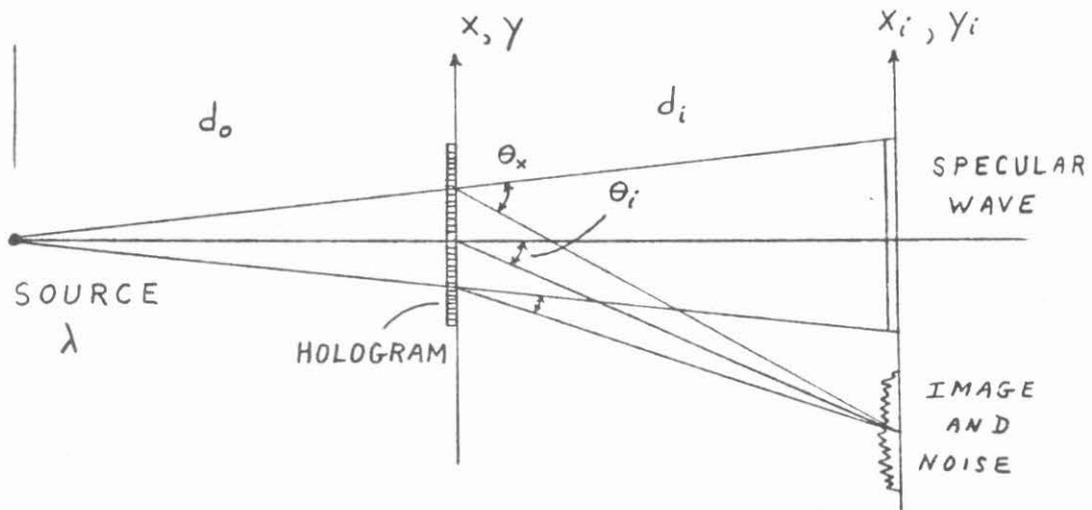


Fig. 4.3 Geometry for analysis of noise in a holographic medium.

onto the plane (x_i, y_i) . The angle θ_i must be large enough that the image lies outside the specular wave. In the case of the spatial filter of Figure 1, the image would be the impulse response of the filter.

Only three terms in the total transmittance contribute significantly to the image and noise within the image area.

$$\tilde{T}(x,y) \approx \underbrace{\tilde{T}_1 e^{i\phi_1}}_{\text{signal}} + \underbrace{\tilde{T}_0 e^{i\phi_0} \tilde{t}_N}_{N_0} + \underbrace{\tilde{T}_1 e^{i\phi_1} \tilde{t}_N}_{N_1} \quad (4.6)$$

noise terms

The first of these is the signal wave, and in the case of the spatial filter it may be assumed to have the same form as equation 1,

$$\tilde{T}_1(x,y) = \tilde{H}(x,y)$$

The second term represents scattering of the zero-order wave, or specular term, into the image aperture. This will produce a distribution of noise light over the image area which is uniform and uncorrelated with the signal, as shown in Figure 4. The third term represents scattering of the first order wave by low-frequency noise in the medium. The noise power spectral density is much higher at low spatial

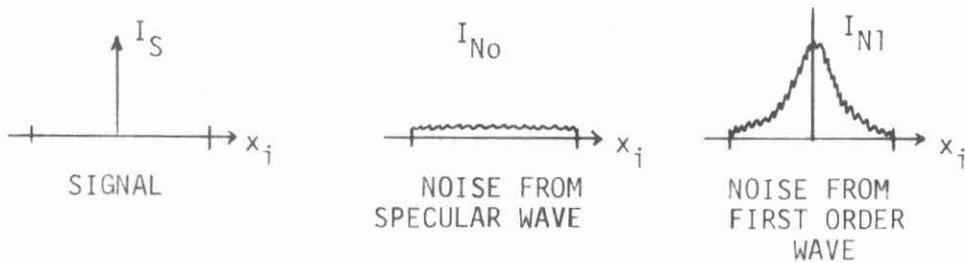


Fig. 4.4 Signal and noise intensity distributions at the image plane for a single point signal.

frequencies, and the noise from this term is therefore sharply peaked around a single point image. For a diffuse image, the noise will be more evenly distributed, but will still increase with diffraction efficiency. Noise scattering from other diffracted orders will generally be of the same form as I_{No} and can be included in this term.

The average noise intensity at the image plane in Figure 3 due to scattering from the zero-order wave is easily evaluated. The average noise power passing through a small aperture $\Delta x_i \Delta y_i$ at the image

plane, due to the scattering from a small area $dx dy$ at the hologram may be written as:

$$d \langle I_{NO}(x_i, y_i) \rangle \Delta x_i \Delta y_i = I_0(x, y) dx dy \Phi(p, q) \Delta p \Delta q \quad (4.7)$$

$$\Delta p \Delta q = \Delta x_i \Delta y_i / (\lambda d_i)^2$$

$I_0(x, y)$ is the intensity of the zero-order wave transmitted by the hologram. $\Phi(p, q)$ is the noise power spectral density for the medium at spatial frequencies (p, q) . Using paraxial approximations for the angles in Figure 3, these spatial frequencies are:

$$p(x, x_i) \approx (\sin \theta_x) / \lambda \approx (x_i / d_i - x / d_i - x / d_0) / \lambda \quad (4.8)$$

$$q(y, y_i) \approx (\sin \theta_y) / \lambda \approx (y_i / d_i - y / d_i - y / d_0) / \lambda \quad (4.9)$$

Integration of equation 7 gives

$$\langle I_{NO}(x_i, y_i) \rangle = \frac{1}{\lambda^2 d_i^2} \iint_{A_h} I_0(x, y) \Phi(p, q) dx dy \quad (4.10)$$

with A_h being the aperture at the hologram plane. Kozma⁽¹⁶⁾ has deduced this same relationship (with I_0 constant) on the basis of a scalar diffraction theory presented by Goodman.⁽¹⁴⁾

In the derivation of equation 10, we have assumed that each element $dx dy$ contributes incoherently to the total noise at the output aperture $\Delta x_i \Delta y_i$. This assumption breaks down if the hologram area $dx dy$ is not much larger than the correlation area $\ell_x \ell_y$ of the noise processes at hologram plane, or if the sample aperture $\Delta x_i \Delta y_i$ is not much larger than the speckle correlation area at the image plane.

Since the average speckle size is inversely related to the hologram

area, this latter condition sets a minimum area for the probe beam in a measurement system (such as Figure 5).

The expected noise intensity due to scattering from the first-order, or signal wave, is more difficult to evaluate since this noise term is correlated with the signal. In this case the noise amplitude at the image plane can be written as a convolution of the signal amplitude with the random part of the filter impulse response,

$$\tilde{a}_{N1}(x_1, y_1) = \frac{1}{\lambda d_i} \iint \tilde{a}_s(\eta) \tilde{h}_N(\xi - \eta) d\eta \quad (4.11)$$

$$\xi = (x_1, y_1) = (x_i + (K/k)d_i, y_i)$$

(coordinates centered on first-order wave)

The noise-free signal is given by:

$$\tilde{a}_s(x_1, y_1) = \frac{a_0}{\lambda d_i} \mathcal{F}\{\tilde{T}_1(x, y) P(x, y)\} (v_x, v_y) \quad (4.12)$$

a_0 : amplitude incident on hologram

$P(x, y)$: hologram pupil function

$(v_x, v_y) = (x_1/\lambda d_i, y_1/\lambda d_i)$ spatial frequencies

and the noise impulse response by:

$$\tilde{h}_N(x_1, y_1) = \frac{1}{\lambda d_i} \mathcal{F}\{\tilde{t}_N(x, y) P(x, y)\} (v_x, v_y) \quad (4.13)$$

The expected value for the noise intensity is:

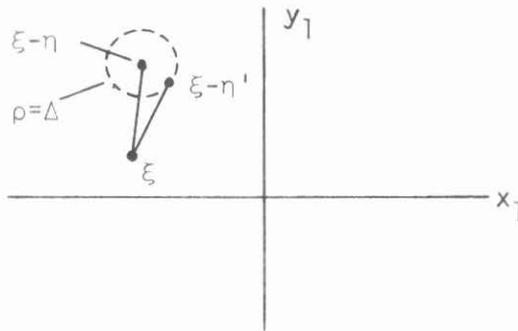
$$\langle I_{N1}(x_1, y_1) \rangle = \langle |\tilde{a}_{N1}|^2 \rangle$$

$$\langle I_{N1}(\xi) \rangle = \frac{1}{\lambda^2 d_i^2} \iiint \tilde{a}_s(\xi - \eta) \tilde{a}_s^*(\xi - \eta') \langle \tilde{h}_N(\eta) \tilde{h}_N^*(\eta') \rangle d\eta d\eta' \quad (4.14)$$

Further reduction of this integral requires some specific assumptions

about the noise statistics and the signal distribution:

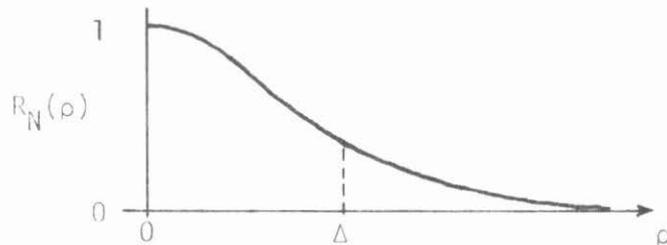
- i) $\langle \tilde{h}_N(\eta) \tilde{h}_N^*(\eta') \rangle$ decorrelates over a distance $\rho = |\eta - \eta'|$ much smaller than the overall variation of $\langle |h_N|^2 \rangle$, and within this distance the autocorrelation function is isotropic. Physically, this means that the average speckle is much smaller than the total noise distribution and has no preferred orientation.



- ii) The autocorrelation function is separable as a product,

$$\langle \tilde{h}(\eta) \tilde{h}^*(\eta') \rangle = \langle |\tilde{h}(\eta)|^2 \rangle R_N(\rho) \quad (4.15)$$

Not all distributions which satisfy i) will meet ii), but we will proceed with this assumption.



- iii) The signal distribution $\tilde{a}_S(\xi)$ does not vary significantly over the correlation distance Δ .

$$\tilde{a}_S(\xi - \eta) \tilde{a}_S^*(\xi - \eta') = I_S(\xi - \eta) \quad (4.16)$$

This is perhaps the most restrictive assumption of the three, as it limits our consideration to images well below the resolution limit, with no details the size of the speckle.

With the above assumptions the expected noise intensity can be written as:

$$\langle I_{N1}(\xi) \rangle = \frac{1}{\lambda^2 d_i^2} \iint I_S(\xi - \eta) \langle |\tilde{h}_N(\eta)|^2 \rangle \int_0^\infty R_N(\rho) 2\pi\rho \, d\rho \quad (4.17)$$

The integral over ρ is roughly the average speckle area,

$$\int_0^\infty R_N(\rho) 2\pi\rho \, d\rho = \alpha\Delta^2 = \lambda^2 d_i^2 / A_h \quad (4.18)$$

A_h : hologram area

The noise impulse is related to the noise power spectral density by

$$\langle |\tilde{h}_N(\eta)|^2 \rangle = (A_h / \lambda^2 d_i^2) \Phi(\eta / \lambda d_i) \quad (4.19)$$

The expected noise intensity in terms of measurable quantities is then

$$\langle I_{N1}(\xi) \rangle \approx \frac{1}{\lambda^2 d_i^2} \iint I_S(\xi - \eta) \Phi(\eta / \lambda d_i) \, d\eta \quad (4.20)$$

By means of this formula and the noise power spectral density data for a point image, one can calculate the expected noise intensity in a more general image distribution, due to scattering of the image wave. The formula is strictly true only when the assumptions i) - iii) are satisfied, but it should provide a useful approximation in any case. A more

accurate result would require evaluation of the integral (4.14) with a specific image distribution, or a more exact specification of the noise statistics for the particular case being considered.

4.3 Measurements of Diffraction Efficiency and Noise in Several Media

Measurements were made in several low-noise media to determine their relative performance as holographic spatial filters. The method of these measurements is illustrated in Figure 5. Experimental details are given in Appendix F. A laser beam with a Gaussian profile was passed through the test hologram. The hologram which was recorded with spherical waves, focuses the first order beam onto the output plane. A photodiode with a small aperture samples the signal and noise intensities at this plane. The aperture is large enough to include all of the power in the signal spot. Recording and measurement wavelengths were both 4880\AA .

The recorded signal was a single bright point, in order to eliminate any intermodulation noise which could occur with a more complicated

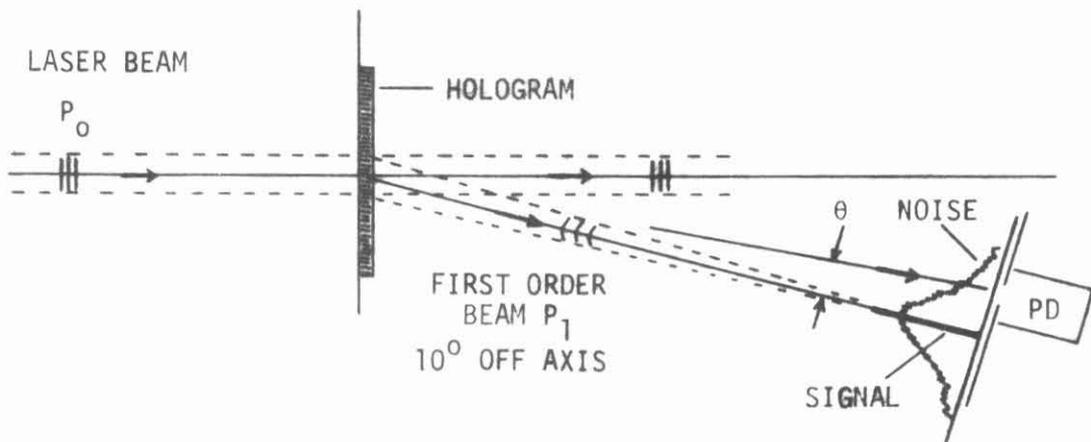


Fig. 4.5 Measurement of noise power spectra. θ is proportional to spatial frequency. PD is a photodiode.

signal. The measured noise pattern should be similar to that which might occur in a real spatial filter in which the impulse response is very localized.

Noise power spectra were plotted for several gratings in each medium at different levels of diffraction efficiency. The scattering angle θ measured from the central spot is proportional to spatial frequency at the hologram plane.

Signal-to-noise intensity ratios were measured in all cases with the detector 7.8 milliradians off the signal spot, in order to facilitate comparison of the media. This angle was chosen to be as close to the signal spot as possible without interference from light in the sidelobes of the signal. The Gaussian beam profile is essential to reduce these sidelobes. The shape of the power spectra does not vary much from one medium to another, and the measurements at 7.8 mr can be taken as roughly proportional to the total noise power.

The plots in Figures 6 through 12 show the results of these measurements. In each graph the measurement system noise levels are plotted as dashed lines. Any point more than 3 db above these levels can be considered entirely hologram noise. The levels labeled Φ_{M_0} were determined from measurements of the noise scattered from a glass plate 10^0 off axis. These levels represent the minimum scattering of the zero-order wave, and were determined from measurements on a clean glass plate obtained from one of the holograms with the emulsion removed.

The levels Φ_{M_1} are 58 db below the signal spot, and represent the noise associated with the first-order wave of a noise-free hologram. This noise is the result of scattering of the laser beam by mirrors, lenses, hologram substrate, and other elements of the measurement system. Lacking a perfect hologram with which to make this measurement, a worst case calculation was made from the measurements on a glass plate. Auxiliary lenses were placed in front of the hologram to provide beam convergence, and the noise was measured from a glass plate 7.8 milliradians off the main spot. This noise level was 58 db below the main spot. The auxiliary lenses introduce several db additional noise into the measurement. This figure, therefore, represents a worst case for the measurement system with a noise-free hologram. Details of these measurements are given in Appendix F.

The dichromated gelatin media (Figures 6 and 7) show a linear relationship between diffracted wave amplitude and exposure over a very wide exposure range. The silver bromide media show no such behavior. The simplicity of the photochemical mechanism in dichromated gelatin as compared to the complexity of reactions in a photographic grain undoubtedly accounts for this behavior.

Several of the spectra show an anomalous bump at 10-12 mr. Heavily exposed photographs of these noise patterns showed these bumps to occur in two rings around the signal spot. The origin of this noise was found to be the inside edges of the apertures used to define the beams in the recording system. Larger apertures, placed further from the pinhole spatial filters would alleviate this problem.

The HNO_3 bleach gratings showed very large scattering at low frequencies. Inspection of the emulsion showed a very rough surface, probably the result of dissolution of some of the gelatin by the nitric acid. This is probably an erratic result, due to subtleties in the method of processing. This process, when it works, is supposed to produce very low noise holograms.⁽²²⁾ Both the HNO_3 processed plates and Hariharan's bleach plates showed considerable darkening during the experiment, due to the print-out effect of unstabilized AgBr grains.

Details of the processing of each of these media are given in Appendix G.

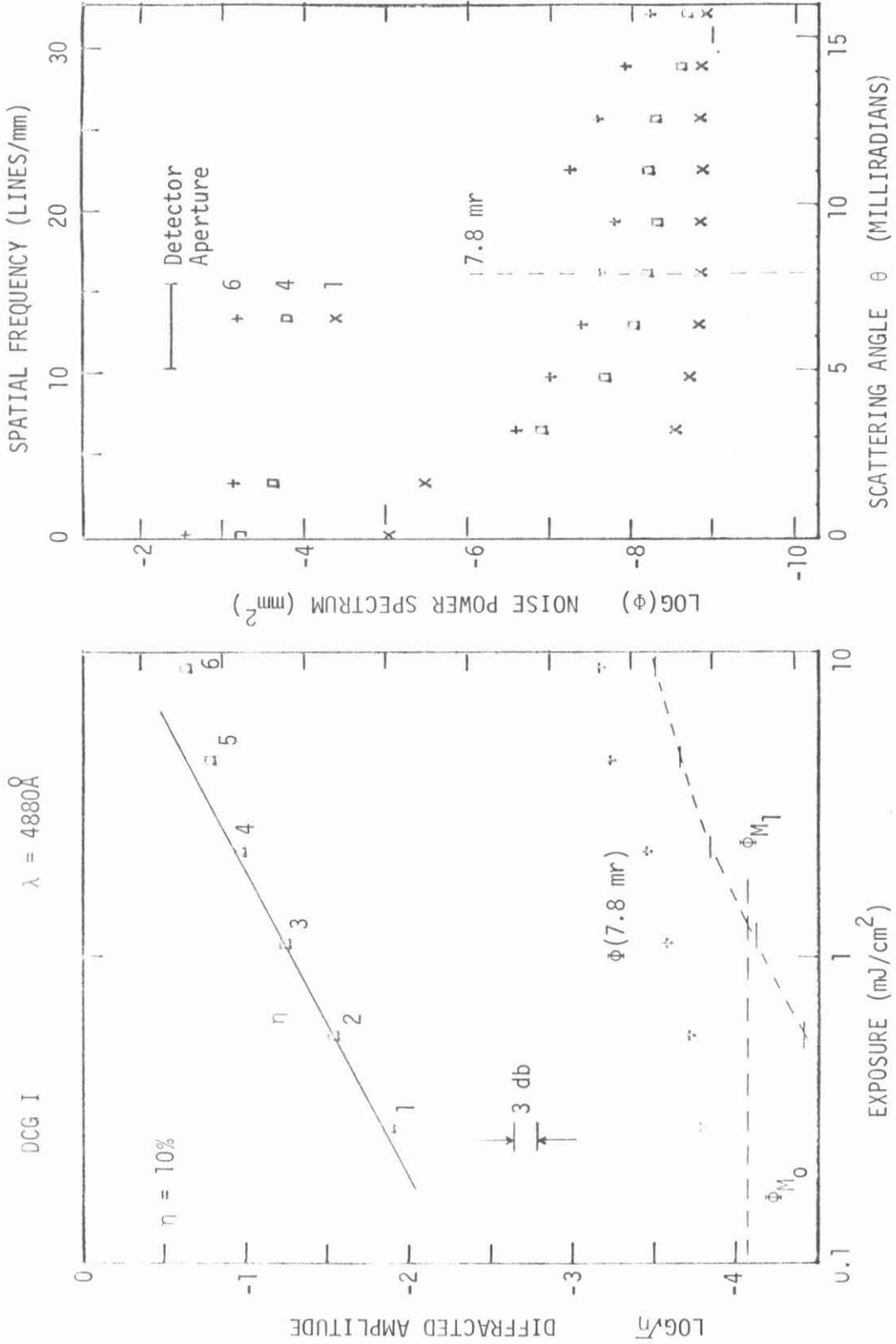


Fig. 4.6 Diffraction efficiency and noise for gratings in dichromated gelatin (Process I).

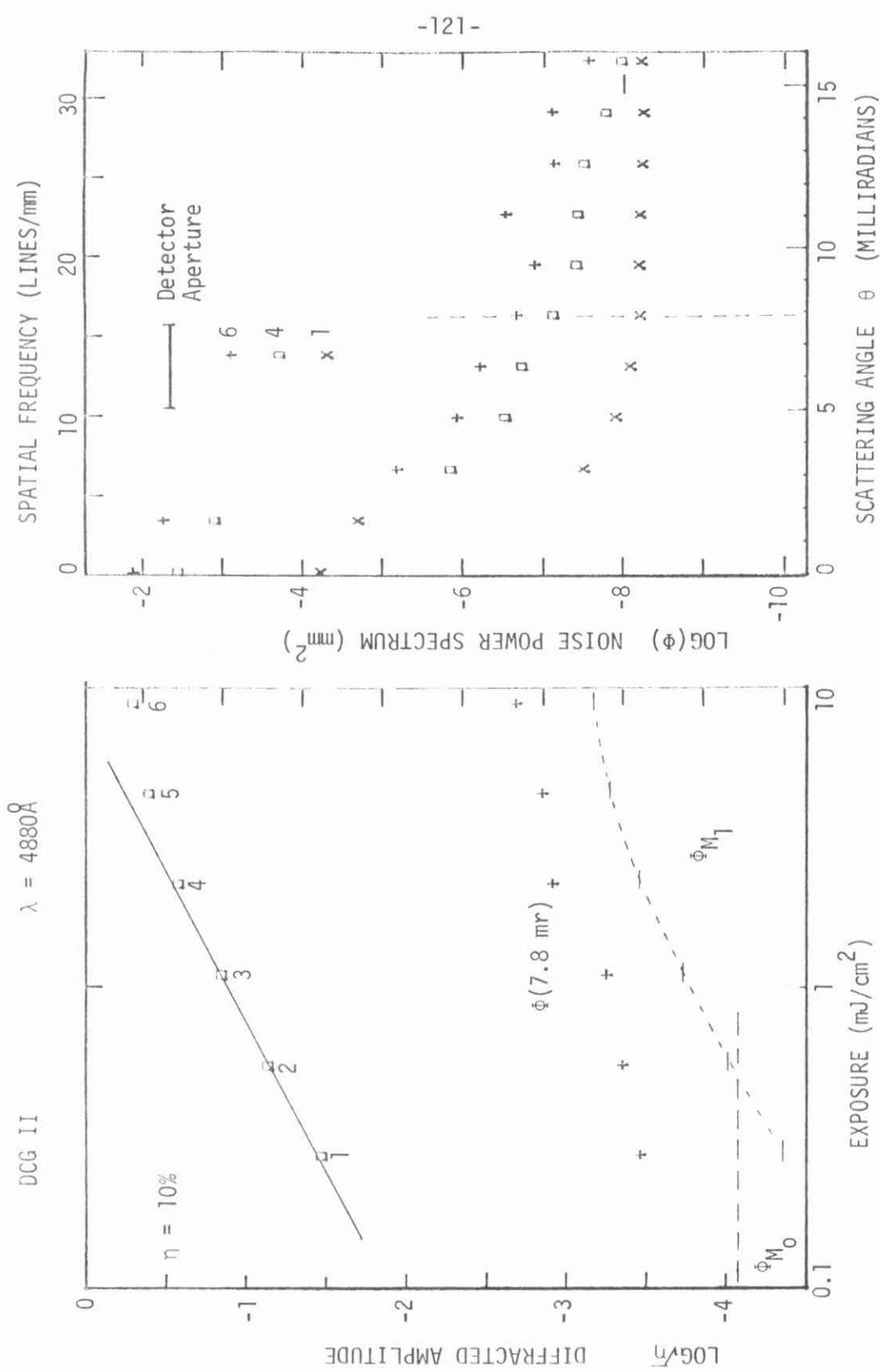


Fig. 4.7 Diffraction efficiency and noise in dichromated gelatin (Process II).

649 F $\lambda = 4880\text{\AA}$

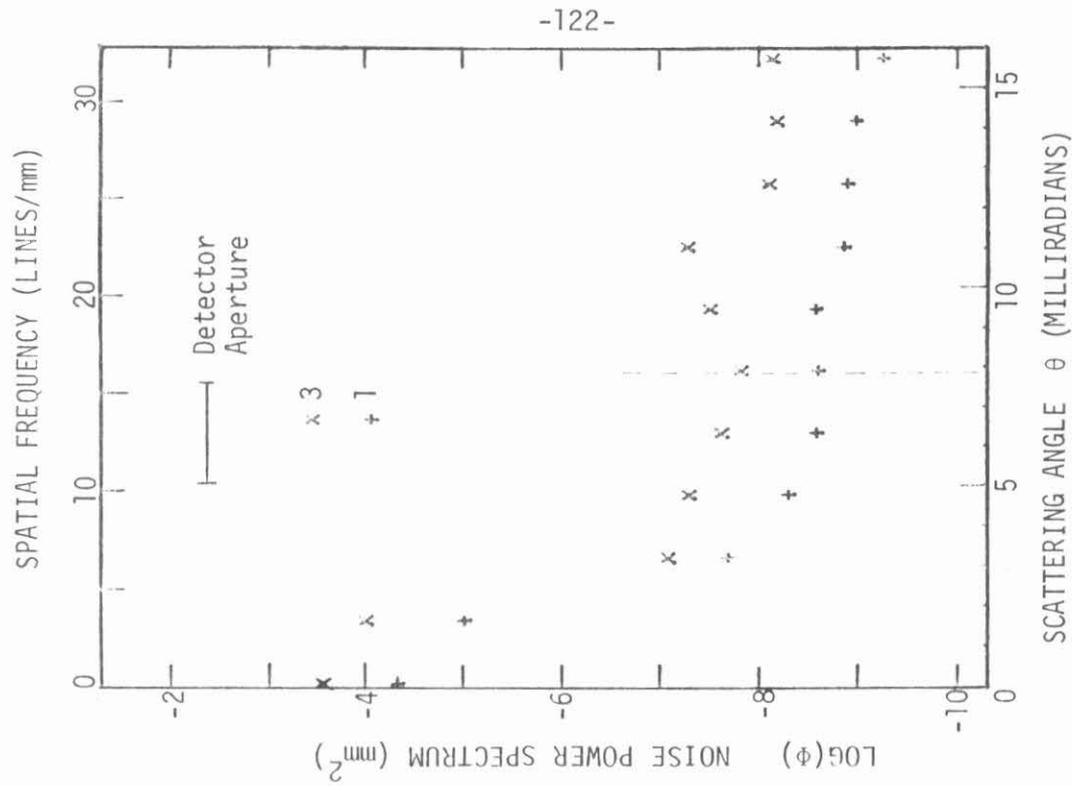
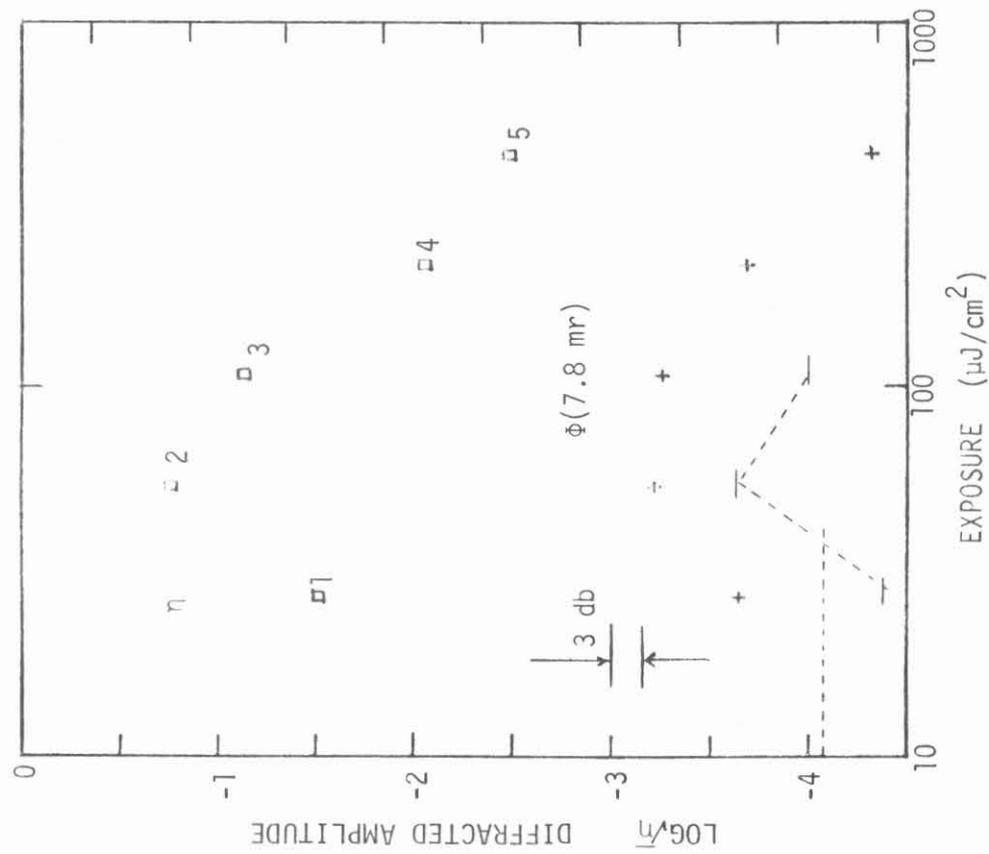


Fig. 4.8 Diffraction efficiency and noise in 649F plates.

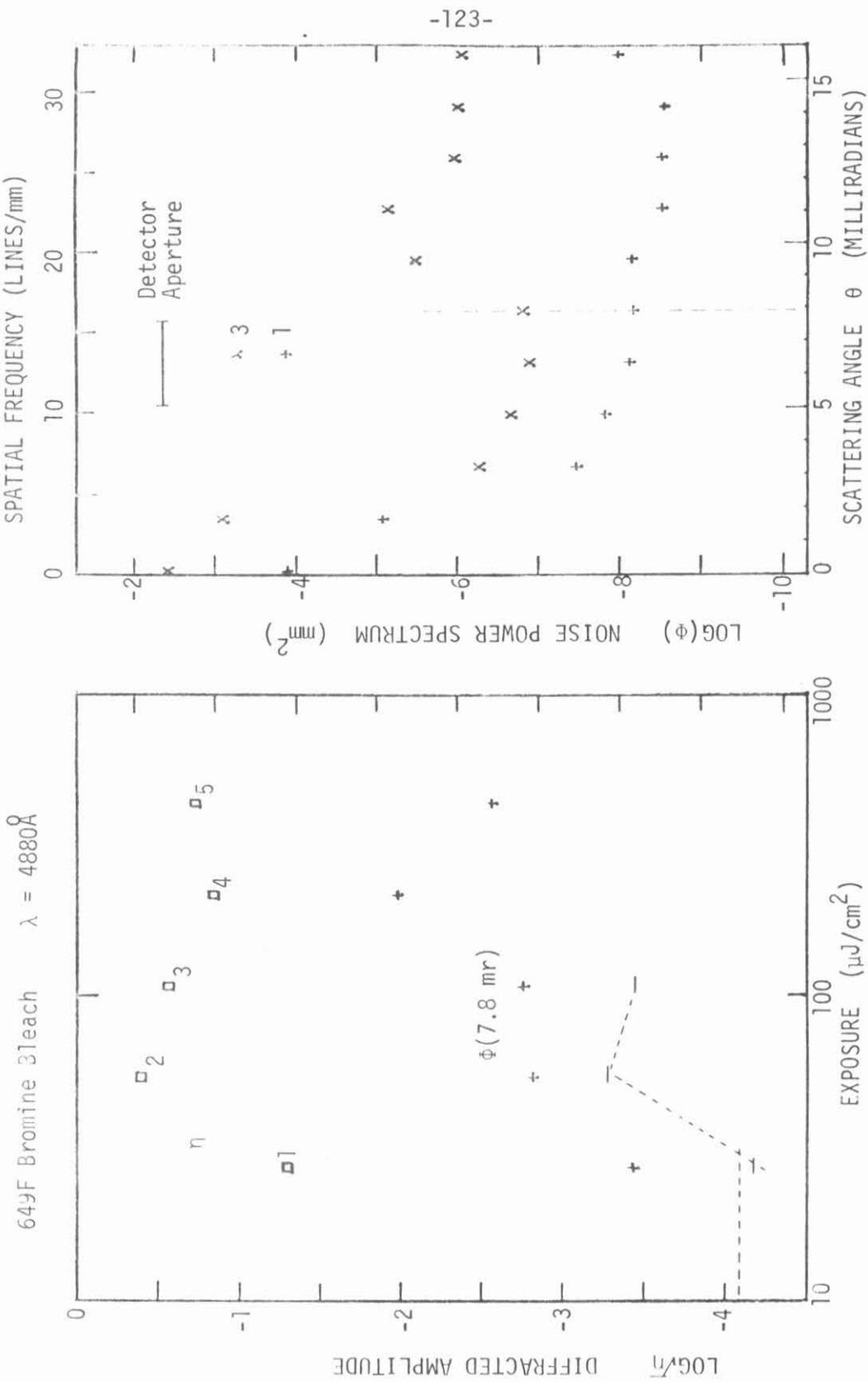


Fig. 4.9 Diffraction efficiency and noise in 649F plates with bromine gas bleach.

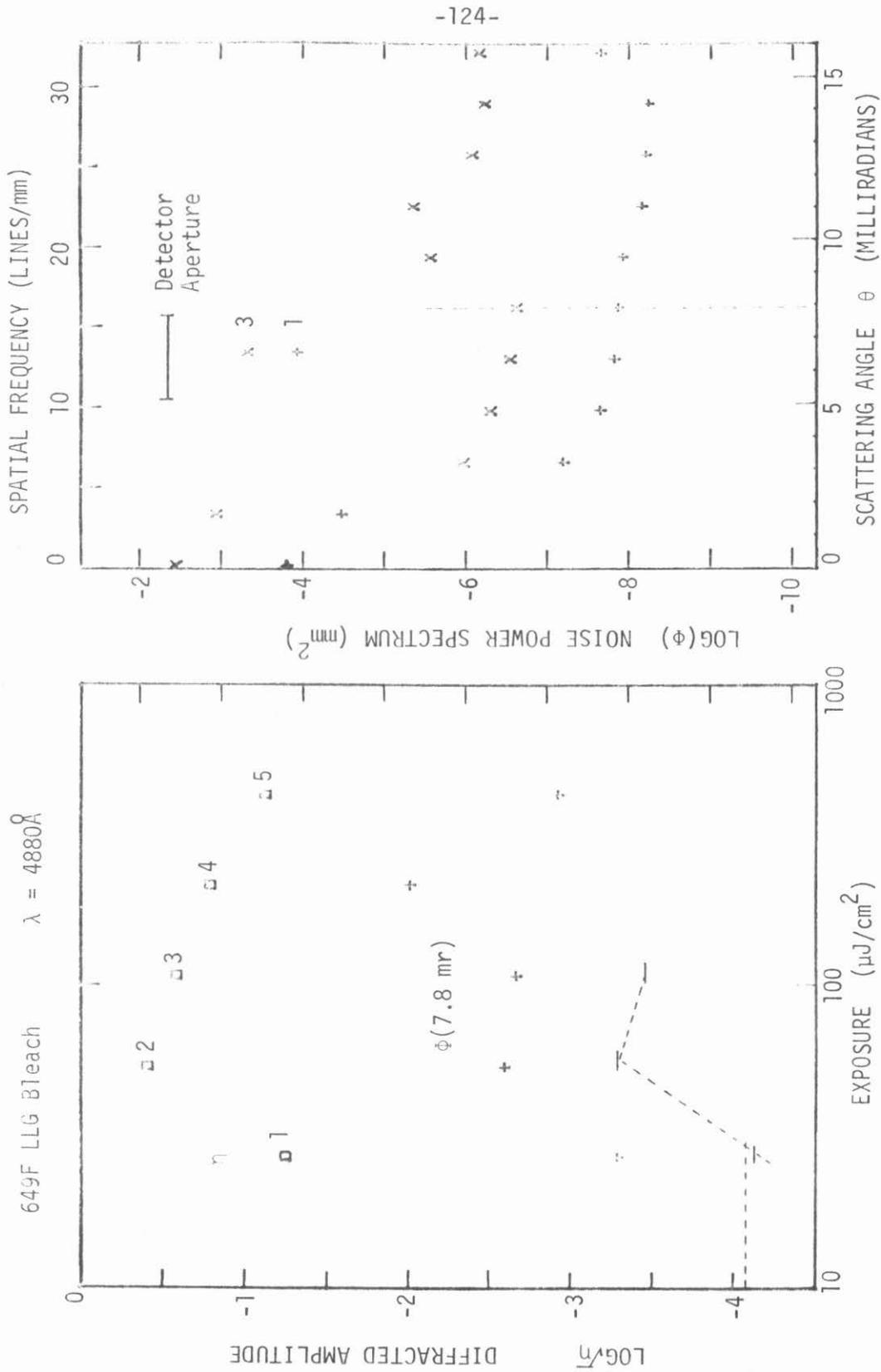


Fig. 4.10 Diffraction efficiency and noise in 649F plates, with liquid bleach (Lehmann, Lauer, Goodman).

649F Hariharan's Bleach $\lambda = 4880\text{\AA}$

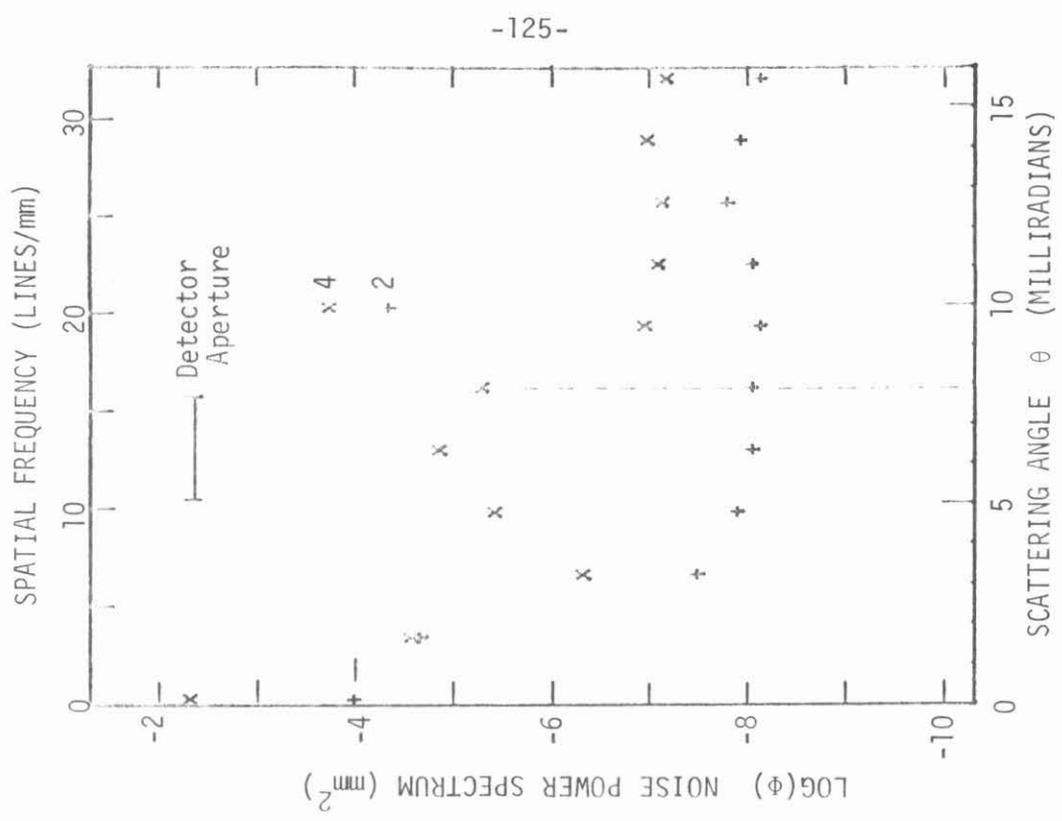
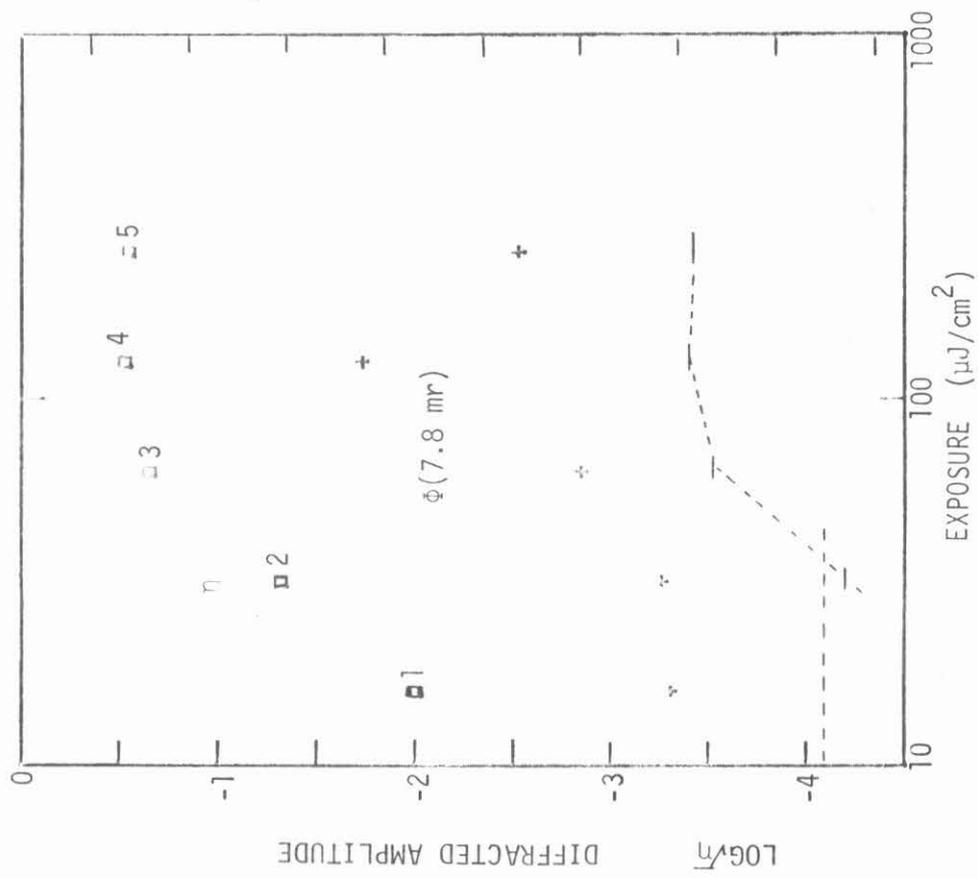


Fig. 4.11 Diffraction efficiency and noise for gratings in 649F plates, Hariharan's bleach.

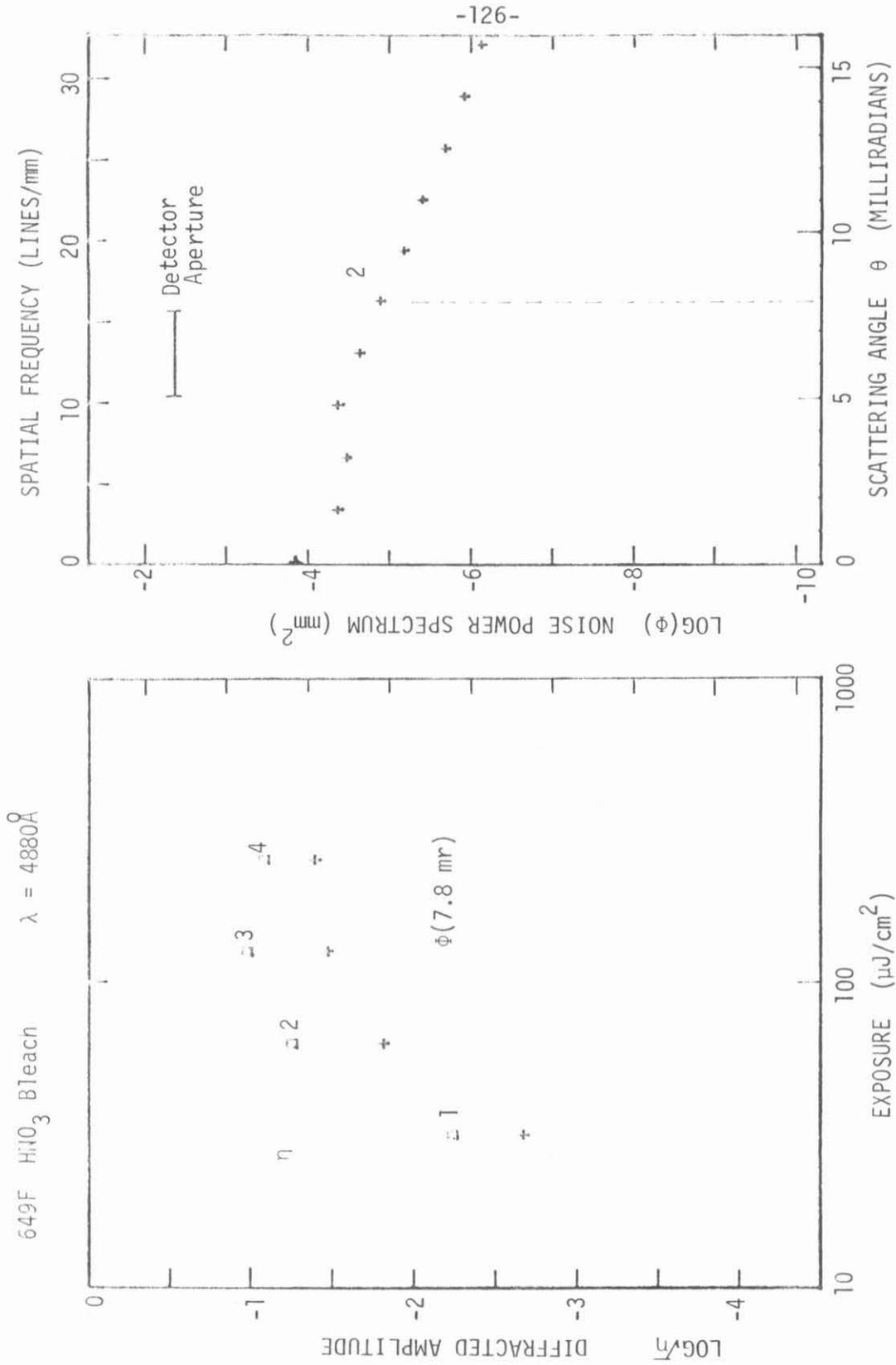


Fig. 4.12 Diffraction efficiency and noise in 649F plates with nitric acid bleach. Excessively high noise levels may be a fluke in the process.

4.4 Summary

The dynamic range of a holographic spatial filter is limited fundamentally by noise originating in the holographic medium. The definition of dynamic range, which is normally applied to nonlinear media, can also be applied to a linear medium in which the only limitation is signal-to-noise ratio at the output plane of a coherent optical processor. Signal-to-noise ratios do not represent a fundamental property of the medium, but are useful in comparing one medium to another under similar circumstances.

A simple model for the noise sources in a hologram was presented. The model leads to two terms for the average noise intensity at the image plane. One of the terms is dependent on diffraction efficiency, and this term will be the dominant noise source in holograms of high diffraction efficiency.

Signal-to-average-noise intensity ratios were measured for each of the media in Figures 6 through 11. These ratios are plotted in Figure 13 versus the diffraction efficiency of the gratings.

The dichromated gelatin media showed about the same signal-to-noise ratios as the silver bromide plates. This was surprising considering the visual appearance of these media. Apparently the relatively noisy appearance of the bleached media is due to wide angle scattering by film grains, and the small-angle, or low-frequency scattering, which is the dominant noise source in these holograms, is about the same for the various media. The coincidence of the noise

levels for the cleanest dichromate process and the 649F unbleached plate suggests a common source of noise. The gelatin for the DCG plates was obtained by fixing unexposed 649F plates. The removal of photographic grains may leave voids or small inhomogeneities in the gelatin which could cause the same scattering as the original grains. Measurements of noise in unexposed, fixed emulsions of various thicknesses support this hypothesis.⁽²⁰⁾ Further experimentation with freshly coated gelatin layers could perhaps give media with much higher signal-to-noise ratios.

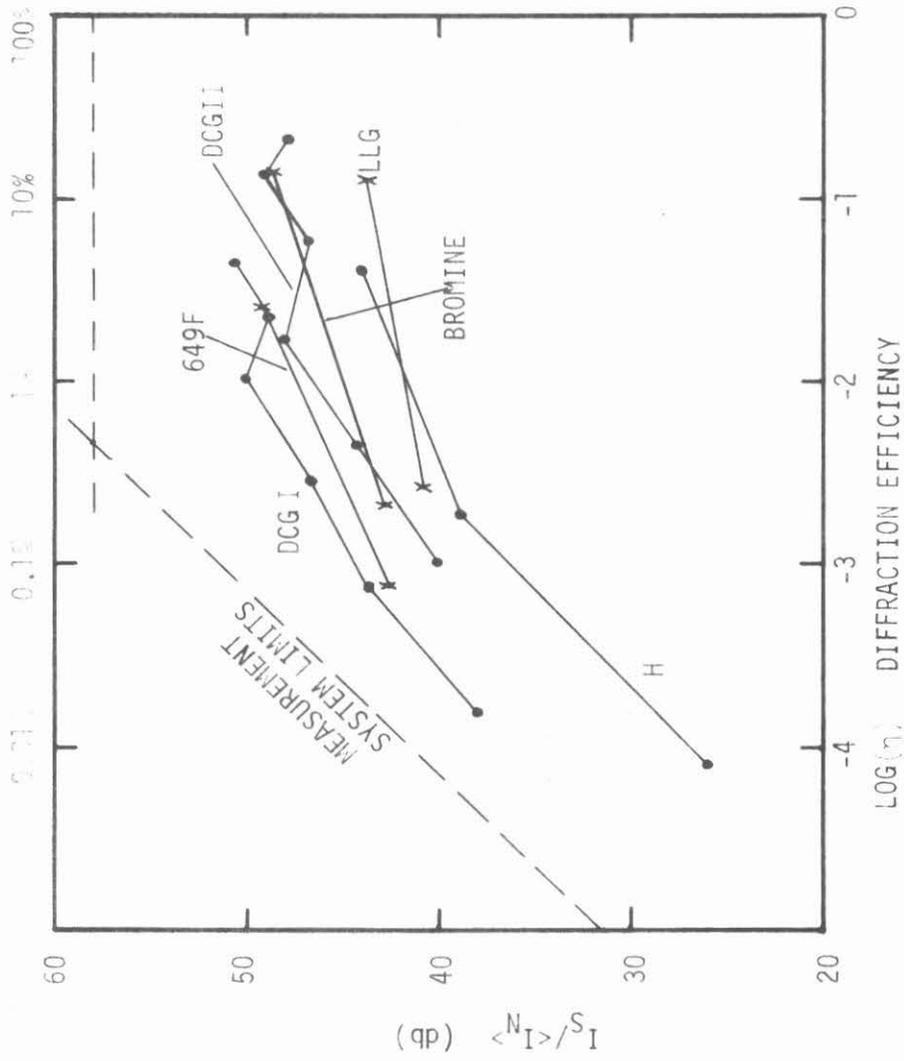


Fig. 4.13 Signal-to-Noise ratios for several low-noise holographic media

References

1. G. W. Stroke, F. Furrer, D. R. Lamberty, "Deblurring of Motion-Blurred Photographs Using Extended-Range Holographic Fourier-Transform Division", *Optics Comm.* 1, 141 (1969).
2. S. I. Ragnarsson, "A New Holographic Method of Generating a High Efficiency, Extended Range Spatial Filter with Application to Restoration of Defocussed Images", *Physica Scripta* 2, 145 (1970).
3. J. W. Goodman, H. B. Strübin, "Increasing the Dynamic Range of Coherent Optical Filters by Means of Modulating Gratings", *J. Opt. Soc. Amer.* 63, 50 (1973).
4. D. A. Tichenor, "Extended Range Spatial Filters for Image Deblurring", Thesis: Stanford University Center for Systems Research (May 1974).
5. R. C. Jones, "New Method of Describing and Measuring the Granularity of Photographic Materials", *J. Opt. Soc. Amer.* 45, 799 (1955).
6. A. L. Ingalls, "The Effect of Film Thickness Variations on Coherent Light", *Phot. Sci. Engr.* 4, 135 (1960).
7. E. N. Leith, "Photographic Film as an Element of a Coherent Optical System", *Phot. Sci. Engr.* 6, 75 (1962).
8. H. Stark, W. R. Bennett, M. Arm, "Design Considerations in Power Spectra Measurements by Diffraction of Coherent Light", *Appl. Opt.* 8, 2165 (1969).
9. C. B. Brandt, "Coherent Optical Power Spectra of Photographic Materials", *Appl. Opt.* 9, 1424 (1970).
10. D. H. R. Vilkomerson, "Measurements of the Noise Spectral Power Density of Photosensitive Materials at High Spatial Frequencies", *Appl. Opt.* 9, 2080 (1970).
11. H. Stark, "Some Film-Noise Measurements by Diffraction of Coherent Light", *Appl. Opt.* 10, 333 (1971).

12. C. E. Thomas, "Film Characteristics Pertinent to Coherent Optical Data Processing Systems", *Appl. Opt.* 11, 1756 (1972).
13. C. W. Helstrom, "Image Luminance and Ray Tracing in Holography", *J. Opt. Soc. Amer.* 56, 433 (1966).
14. J. W. Goodman, "Film-Grain Noise in Wavefront-Reconstruction Imaging", *J. Opt. Soc. Amer.* 57, 493 (1967).
15. C. B. Burckhardt, "Storage Capacity of an Optically Formed Spatial Filter for Character Recognition", *Appl. Opt.* 6, 1359 (1967).
16. A. Kozma, "Effects of Film-Grain Noise in Holography", *J. Opt. Soc. Amer.* 58, 436 (1968).
17. W. H. Lee, M. O. Greer, "Noise Characteristics of Photographic Emulsions Used for Holography", *J. Opt. Soc. Amer.* 61, 402 (1971).
18. W. H. Lee, "Effect of Film-Grain Noise on the Performance of Holographic Memory", *J. Opt. Soc. Amer.* 62, 797 (1972).
19. L. S. Watkins, "Noise Effects of Granularity in Thin Phase Holograms", *J. Opt. Soc. Amer.* 62, 1233 (1972).
20. H. M. Smith, "Light Scattering in Photographic Materials for Holography", *Appl. Opt.* 11, 26 (1972).
21. D. A. Tichenor, "Fourier-Plane Scattering in Image Restoration" Ch. 5 in Extended Range Spatial Filters for Image Deblurring, Thesis, Stanford University Center for Systems Research (1974).
22. N. George, personal communication.

COMPLEX WAVEFRONT SYNTHESIS BY MULTIPLE EXPOSURE HOLOGRAPHY†

David MacQUIGG

California Institute of Technology, Pasadena, California 91109, USA

Received 26 March 1973

An arbitrary complex wavefront can be synthesized by a multiple exposure of three photographic transparencies on a single holographic plate with 120° phase shift in the reference wave between exposures. This method has several advantages over ordinary computer generated holograms. Applications include spatial filtering, hologram optical elements, and image subtraction.

1. Basic wavefront synthesis idea

An arbitrary complex wavefront, represented by the complex function $\tilde{w}(x, y)$, can be synthesized as the sum of three real-valued, non-negative functions of fixed phase:

$$\tilde{w}(x, y) = w_1(x, y) + w_2(x, y)e^{2\pi i/3} + w_3(x, y)e^{4\pi i/3},$$

$w_n(x, y)$ real, non negative.

As shown in fig. 1, a minimum of three fixed phase vectors are required to cover the complex plane. A holographic diffraction grating which will generate the desired complex wavefront from a plane wave can be produced from three real wavefronts as shown in fig. 2. Each of the three component wavefronts is record in succession on a holographic plate with 120° phase shift in the reference wave between exposures. The result is a hologram with an effective (off-axis) transmittance $\tilde{w}(x, y)$. The wavefronts $w_n(x, y)$ are obtained from photographic transparencies, which may be generated by computer or by other means.

Let the amplitudes of the reference and subject beams at the hologram plane be $R \exp[i(kx - \phi_n)]$ and $w_n(x, y)$, $n = 1, 2, 3$; $\phi_n = 0, 2\pi/3, 4\pi/3$. The total film exposure is then:

$$\begin{aligned} E &= \sum_n E_n = \sum_n \tau_e |R \exp[i(kx - \phi_n)] + w_n(x, y)|^2 \\ &= \tau_e \left[3R^2 + \sum_n w_n^2 + R \exp(ikx) \sum_n w_n \exp(-i\phi_n) \right. \\ &\quad \left. + R \exp(-ikx) \sum_n w_n \exp(i\phi_n) \right], \end{aligned}$$

where τ_e is the exposure time for each step. Assuming linear recording of the hologram, the last term in the above expression gives the desired complex transmittance on a carrier $\exp(-ikx)$. The other terms occur at widely separate spatial frequencies.

The addition of wavefronts by multiple exposure holography has been demonstrated in the context of image subtraction [1, 2] and holographic interferometry [3, 4]. The wavefront synthesis idea is also related to the grid coding techniques of Pennington [5] and the methods of Burckhardt and Doherty [6] except that the "grid" here is formed by the interference of the two waves.

The multiple exposure technique has several advantages over purely digital techniques [7, 8] in the synthesis of wavefronts. The generation of the basic grating by optical means relieves the computer of the burden of drawing details the size of each fringe and allows the production of much larger holograms. If we think of the hologram as a high frequency carrier (grating) modulated by a two-dimensional complex-valued envelope, then it is clear that all the information

† Research supported in part by the Air Force Office of Scientific Research.

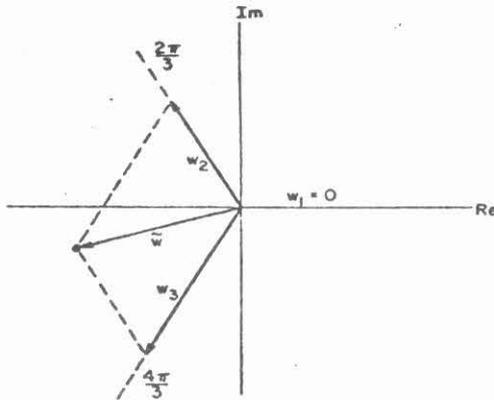


Fig. 1. Illustrating the generation of an arbitrary complex function as the sum of three fixed phase functions with non-negative coefficients.

is contained in the envelope and none in the carrier. The optimum use of the information capacity of any computer plotting device will result if it is used only to generate the envelope function. This can result in enormous savings in the generation of certain image enhancement spatial filters where the envelope is a slowly varying function.

Another advantage of this method is that the hologram be generated directly in a low noise high efficiency thick medium such as dichromated gelatin [9, 10] whereas the digital techniques require a thin hologram. The maximum theoretical diffraction efficiency for a thick phase grating is 100% as compared to 6.25% for

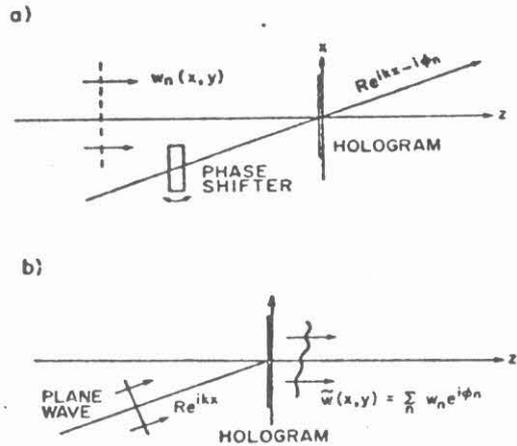


Fig. 2. (a) Recording of component wavefronts by multiple exposure. (b) Reconstruction gives desired wavefront.

a thin sinusoidal amplitude grating and 10% for a binary amplitude grating [11].

2. Technique for generating the desired grating

The major problem encountered in the making of these plates is phase aberrations in the masks used to generate the component wavefronts $w_n(x, y)$. These wavefronts need not be perfectly plane. In fact, the optical phase may show a large systematic error due to other components in the system. However, the

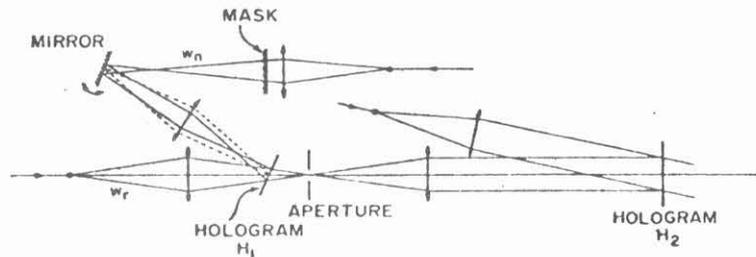


Fig. 3. Wavefront modulator plate (H_1) used to eliminate phase errors in mask function w_n .

phase at the hologram plate must be consistent within a few degrees from mask to mask over the entire (several cm) aperture.

The brute force solution of the phase problem involves precise index matching in a liquid gate with a mixture of high index and low index fluids. In addition the mask substrate must be thin and homogeneous. A 0.040" glass plate requires that index inhomogeneities Δn be less than 10^{-5} to hold the phase error under 7° .

Another approach is to place the masks in direct contact with the hologram plate. The phase aberrations of the substrate will then affect both beams equally and result in no net distortion of the grating on the hologram plate [12].

A third approach is the use of a second hologram as a "wavefront modulator plate". Referring to fig. 3, each of the aberrated wavefronts w_1, w_2, w_3 is separately recorded with reference beam w_r on hologram H_1 with a small change in incident angle. The optics is arranged so that rotation of the mirror causes wavefront w_n to pivot about the center of hologram H_1 . The transmittance of H_1 contains terms $w_1^* w_r, w_2^* w_r,$ and $w_3^* w_r$. When re-illuminated with the original aberrated wavefronts, the reconstructed terms become $|w_1|^2 w_r, |w_2|^2 w_r$ and $|w_3|^2 w_r$ which can be recorded by multiple exposure on hologram H_2 . The phase aberrations in the reconstructed wavefronts are due to w_r and the substrate of H_1 and are consistent from mask to mask. Interfering terms $w_1 w_2^* w_r$, etc. fall off axis and are easily separated out by a small aperture. The advantage of this method over contact printing is the ability to filter out noise and raster in the computer generated masks by placing a small aperture in the focal plane of beam w_r .

An interesting feature of both the contact printing and modulator plate methods is the squaring of the mask functions. This can be very useful in the making of spatial filters as it allows a doubling of the dynamic range of the filter. If the available range of transmittance of the computer generated mask is 10^3 then the final filter can have a range of 10^6 .

3. Applications of multiple exposure wavefront synthesis

The multiple exposure method of wavefront synthesis should prove most applicable in cases where the

desired wavefront varies slowly in phase and amplitude with respect to some easily generated reference wavefront. Such applications include spatial filters for image enhancement, pattern recognition and image coding. Other applications such as generation of artificial Fresnel holograms, in which the wavefront shows very large fluctuations over short distances, may be impractical because of the necessity of keeping each mask aligned within micron tolerances.

Synthesis of hologram lenses and optical elements may be possible in cases where an approximate wavefront can be generated optically. For example, using a plane wave and a spherical wave it would be possible to synthesize a hologram which converts an incident plane wave to a parabolic wave.

Image subtractions may be accomplished by using two exposures with 180° shift in the reference wave. This has been done by Gabor [1] and by Collins [2], but they avoided the phase aberration problem by using the same object for both exposures, blocking part of the object with an opaque mask for the second exposure. Bromley [13] used photographic film in a liquid gate for his image subtractions. To the best of my knowledge, the contact printing method has not been used in image subtraction experiments.

Applications of the above method to the generation of spatial filters for image enhancement are now being studied and results will be reported later.

References

- [1] D. Gabor, G.W. Stroke, R. Restrick, A. Funkhouser and D. Brumm, *Phys. Letters* 18 (1965) 116.
- [2] L.F. Collins, *Appl. Opt.* 7 (1968) 203.
- [3] L.O. Heflinger, R.F. Wuerker and R.E. Brooks, *J. Appl. Phys.* 37 (1966) 642.
- [4] G.W. Stroke and A.E. Labeyrie, *Appl. Phys. Letters* 8 (1966) 42.
- [5] K.S. Pennington, P.M. Will and G.L. Shelton, *Opt. Commun.* 2 (1970) 113.
- [6] C.B. Burckhardt and E.T. Doherty, *Appl. Opt.* 7 (1968) 1191.
- [7] T.S. Huang, *Proc. IEEE* 59 (1971) 1335.
- [8] P.L. Ransom, *Appl. Opt.* 11 (1972) 2554.
- [9] T.A. Shankoff, *Appl. Opt.* 7 (1968) 2101.
- [10] M. Chang, *Appl. Opt.* 10 (1971) 2550.
- [11] R.J. Collier, C.B. Burckhardt and L.H. Lin, *Optical holography* (Academic Press, New York, 1971) p. 224.
- [12] J.W. Goodman, W.H. Huntley Jr., D.W. Jackson and M. Lehmann, *Appl. Phys. Letters* 8 (1966) 311.
- [13] K. Bromley, M.A. Monahan, J.F. Bryant and B.J. Thompson, *Appl. Opt.* 10 (1971) 174.

APPENDIX B

Computer Programs for Analysis of Thick
Holographic Diffraction Gratings

The following are listings of the subroutine WAVES used in generating the plots of Chapter 2, and a typical program using the WAVES subroutine. This subroutine was written in collaboration with Alan Mickelson. The method of computation is presented in Chapter 2. This is a general purpose subroutine to compute the amplitudes of the diffracted waves from an unslanted phase grating of arbitrary thickness and arbitrary (but symmetric) grating fringe profile.

Usage:

CALL WAVES (WVLG, THETA, D, T, EPS, NL, VL)

WVLG: Wavelength in air

THETA: Incident wave internal angle (degrees)

D,T: Grating period and thickness

EPS: Array containing Fourier coefficients of the grating
index profile

NL: Number of orders retained in the calculation

VL: Complex array of wave amplitudes

EPS and VL must be dimensioned at least NL. Wave amplitudes are returned with zero-order coefficient in VL(L0). $L0 = (NL + 1)/2$.

Timing: Approx. 120 msec on IBM 370/158

Storage: 8936 bytes

Subroutines Called^{*}

HSEHLD: Finds eigenvectors and eigenvalues of a real-symmetric matrix,
by the Givens-Householder method.

EQSOV: Simultaneous equation solver uses Gaussian elimination with
iterative improvements.

* Documentation available at the Caltech Computing Center.

```

SUBROUTINE WAVES( WVLG,THETA,D,T,EPS,NL,VL )
C
C GIVEN WAVELENGTH AND INTERNAL ANGLE OF THE INCIDENT WAVE, GRATING
C PERIOD D AND THICKNESS T, AND THE FOURIER COEFFICIENTS OF THE GRATING
C PROFILE EPS, SUBROUTINE RETURNS COMPLEX AMPLITUDES VL FOR THE NL DIF-
C FRACTED WAVES CENTERED ABOUT THE ZERO ORDER.
C
DIMENSION EPS(21),A(21,21),U(21,21),GAM(21,5),B(21),C(21)
DIMENSION UBLCK(50,21)
EQUIVALENCE(UBLCK(1,1),A(1,1))
REAL KO,KO2,KG,KST,KOZEPS(21),NO,TUPI/6.2831853/
COMPLEX VL(21),E(21)
LO = (NL+1)/2           LO : Index of zero order
KO=TUPI/WVLG
KO2=KO*KO
KG=TUPI/D
NO = SQRT( EPS(1) )     n0 = √ε0
KST = KO*NO*SIN( THETA*0.0174533 )
DO 100 N=1,NL
100 KOZEPS(N)=KO2*EPS(N)
DO 200 I=1,NL
  L = I - LO
  ALFL=I*KG-KST
200 A(I,I)=KOZEPS(1)-ALFL*ALFL } Compute matrix A
  DO 300 J=2,NL
    JM1=J-1
    DO 300 I=1,JM1
      N=J-I+1
300 A(I,J)=KOZEPS(N)
  CALL HSEHLD(NL,21,A,U,GAM) } Find eigenvectors U and
  DO 500 J=1,NL } eigenvalues GAM
  DO 500 I=1,NL
500 UBLCK(I,J)=U(I,J)
  B(LO) = 1. } Given boundary conditions Bj
  MAX = 5 } Solve for coefficients Cj
  CALL FQSOV(NL,UBLCK,B,MAX,1.E-3,C,IT,0)
  IF(IT.GE.MAX) PRINT 60, MAX,WVLG,THETA,D,T,(EPS(N),N=1,6),NL
60 FORMAT (' NO CONVERGENCE ',I3,1P10E10.3,I5)
C RECN HERE IF ONLY T HAS CHANGED
  ENTRY NEWT(T)
700 DO 710 J=1,NL } Compute exponential factors
  GAMMAJ=GAM(J,1)
  IF(GAMMAJ.LT.0) GO TO 705
  PHI=SQRT(GAMMAJ)*T
  E(J)=CMPLX(COS(PHI),SIN(PHI))
  GO TO 710
705 E(J) = EXP( -AMINI(SQRT(-GAMMAJ)*T,50.) ) } AMINI to avoid
710 CONTINUE } underflow
P00 DO 850 I=1,NL } Compute diffracted
  VL(I)=(C.,0.) } orders
  DO 850 J=1,NL
850 VL(I)=VL(I)+C(J)*U(I,J)*E(J)
  RETURN
  END

```

```
// SET ICC=10
// EXEC FCRTG
//SYSPLTDN DD SYSOUT=N
//FCRT DD *
C PROGRAM TO PLOT FIRST ORDER DIFFRACTION EFFICIENCY FOR PHASE
C TRANSMISSION GRATINGS.
  DATA PI /3.14159265/
  DIMENSION EPS(12),P(12),X(60),Y(60,10),QI(10)
  COMPLEX VL(12)
C INITIAL DATA VALUES
  DATA QI/0.1,1.,2.,3.,4.,5.,6.,8.,10.,20./
  REAL NO/1.54/
  WVLG = 0.5
  T = 12.0
  NL = 12
C INITIAL CALCULATED VALUES
  LO = (NL+1)/2
  EPS(1) = NO*NO
  LO : ZERO ORDER
  EPS(1) = E0
C
  DO 500 IQ = 1,10
  Q = QI(IQ)
  D = SQRT( 2.*PI*WVLG*T/(NO*Q) )
  THETAB = ARSIN(WVLG/(2.*NO*Q))
  C1 = WVLG*NO*COS(THETAB)/(2.*PI*T)
  THETA = (180./PI)*THETAB
  WRITE (6,27) NO,WVLG,THETA,D,T,C1,NL,Q
  27 FORMAT (/1X,5F10.3,E15.3,I10,F10.3)
  Given Q
  Calculate D
  BRAGG ANGLE
C BEGIN LOOP TO FILL PLOTTING ARRAYS X,Y.
  DO 400 I = 1,50
  PHI = .04*PI*I
  EPS(2) = C1*PHI
  CALL TIMFIN
  CALL WAVES (WVLG,THETA,C,T,EPS,NL,VL)
  CALL TIMOUT
  EPS(2) = EH
C CALCULATE POWER IN DIFFRACTED ORDERS
  SUM = 0.
  DO 405 L = 1,NL
  VR = REAL( VL(L) )
  VI = AIMAG( VL(L) )
  P(L) = VR*VR + VI*VI
  405 SUM = SUM + P(L)
  WRITE (6,29) PHI,F,SUM
  29 FORMAT (1X,14F9.4)
C STORE DATA FOR PLOTTING
  Y(I,IQ) = P(LO+I)
  400 X(I) = PHI
  CALL TIMSUM
  500 CONTINUE
  Timing Summary
C
C SAVE DATA FOR LATER PLOTTING
  PUNCH 50, X,Y
  50 FORMAT (20A4)
  WRITE (6,31) (X(I),(Y(I,IQ),IQ=1,10),I=1,50)
  31 FORMAT (//(1X,11F11.3))
C BEGIN PLOTS
  CALL SCALE( 3.,2.,5.,5.,0.,2.*PI,0.,1. )
  DO 700 IQ = 1,10
  700 CALL PLECTLN( X,Y(1,IQ),50 )
  CALL AXIS( 0,0.,0.,2.*PI,10 )
  CALL AXIS( 90,0.,C.,1.,10 )
  STOP
  END
```

APPENDIX C

Computer Programs for Image Processing and Generation of Holograms

<u>C.1 Image Processing</u>	<u>page</u>
FFT2	144
FILTER JINC JNCINV	145
Filtering Program	146

<u>C.2 Hologram Generation</u>	
COMPON	147
SCALEX SCALEQ	148
PHASE W2048	149

This appendix provides listings of the programs used in processing the images in Figure 2.15 and generating the holograms of Section 3.3. Several programs of general usefulness are listed also.

FFT2 is a 512 by 512 complex Fast Fourier Transform. FILTER generates a two-dimensional filtering function from a given radial function. JINC and JNCINV generate the circle blurring function and its inverse. COMPON separates a complex function into three fixed-phase components. SCALEX and SCALEQ apply nonlinear transfer functions to compensate film nonlinearities in the generation of digital holograms. PHASE and W2048 are useful for calculating phase factors for diffusers and computation of Fresnel integrals using the Fast Fourier Transform.

FFT2

This routine is essentially a large matrix transpose with a one-dimensional Fourier transform (Subroutine FORT) on the input records

(horizontal transforms) and final output records (vertical transforms). IN and IOUT are FORTRAN file numbers for the input and output files. File 99 is used as an intermediate. The Fourier transform operation performed by this program can be represented by:

$$\tilde{A}_{\nu\mu} = (-)^{(\mu+\nu)} \sum_{m,n=1}^N \tilde{X}_{mn} (-)^{(m+n)} e^{i(2\pi/N)(m\mu+n\nu)} \quad (C.1)$$

\tilde{X}_{mn} : input complex array (N by N)

$\tilde{A}_{\nu\mu}$: output transformed array (N by N)

The factors $(-)^{(m+n)}$ and $(-)^{(\mu+\nu)}$ are applied to shift the origin of the transform to the center of the array [i.e., zero spatial-frequency at point (257,257)]. Notice that the transformed array is transposed. In this application the orientation of the matrix is unimportant. If a re-transpose is necessary, this may be accomplished by running the program with a dummy FORT routine.

A description of the Fast Fourier Transform method^(1,2) and the specific algorithm used in the FORT routine⁽³⁾ can be found elsewhere.

Timing: (on IBM 370/158)

97 sec	for transforms
<u>19 sec</u>	for data shuffling (Dummy FORT)
116 sec	Total

Storage: 142K bytes

Files Required: Two or three random-access disk files, 512 records, 4096 bytes per record. (The output file may be the same as the input).

SCALEX, SCALEQ

These routines were written to apply a nonlinear scaling function to a computed image array for the purpose of compensating film and other process nonlinearities. The standard linear interpolation routines (YINTERP) would be far too slow for these large arrays. A compromise must be made between speed and accuracy. SCALEQ determines scaled values by indexing in a precomputed table. This routine is extremely fast, but limited in accuracy. The input array must be normalized 0 to 1000 prior to calling this routine. SCALEX performs a binary search in a precomputed scale table. For a limited number of output functional values (e.g., 0-255) this routine provides an exact scaling, but it is, of course, slower than SCALEQ.

	<u>SCALEX</u>	<u>SCALEQ</u>
Timing: * (IBM 370)	65 msec	14 msec
Storage:	2K	3K bytes

PHASE

This routine simulates a "fly's-eye" phase diffuser. In generation of digital holograms phase coding is often necessary to disperse the spectral energy and lower the dynamic range of the recorded wavefront.^(4,5) An image represented by the real array a_{mn} will have typically a large fraction of its spectral energy concentrated near zero frequency. PHASE performs the following operation:

* FORTRAN H compiled. Timing for one line of 1024 values.

$$\begin{aligned}\tilde{a}_{mn} &= a_{mn} e^{i\phi_{mn}} \\ \phi_{mn} &= \phi_{\max} \cos(2\pi m/32) \cos(2\pi n/32)\end{aligned}\quad (C.2)$$

The degree of phase dispersion can be adjusted with ϕ_{\max} .

W2048

In computing the Fresnel integral, a factor $\tilde{w} = e^{i\phi}$ must be evaluated at each point in the field. If 2048 uniformly spaced values for ϕ (over the interval $0-2\pi$) is sufficient resolution, then this routine will provide an extremely fast evaluation of \tilde{w} .

$$\tilde{w}_n = e^{i2\pi n/2048}\quad (C.3)$$

The routine uses a pre-computed cosine table for speed and a bit shifting and masking technique to reduce the total number of stored values in the table to 513. A similar routine could be written for any set of values $\tilde{w}_n = e^{i2\pi n/N}$, where N is a power of two.

Timing: 130 μ sec to return one complex value

Storage: 3K bytes

Documentation for routines FORT, BESJ1, WRTLN, YNTERP, IAND, and SHIFTR is available at the Caltech Computing Center.

References

1. J. W. Cooley, J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series", *Mathematics of Computation* 19, 297-301 (1965).
2. W. T. Cochran, et al., "What is the Fast Fourier Transform?", *Proc. IEEE* 55, 1664-74 (1967).

3. IBM Scientific Subroutine Package--Programmer's Manual H20-0205-2 (1967), p. 148-149, subroutine HARM.
4. W. J. Dallas, "Deterministic Diffusers for Holography", Appl. Opt. 12, 1179 (1973).
5. D. C. Chu, "Spectrum Shaping for Computer Generated Holograms", PhD Thesis, Stanford University (1974).

```

C PROGRAM TO PERFORM TWO DIMENSIONAL FAST FOURIER TRANSFORM ON
C A COMPLEX ARRAY, 512 BY 512
  SUBROUTINE FFT2 (IN, IOUT )
    DIMENSION S(128)
    COMPLEX BLOCK(512,32), BUFFER(32,32), BUFF1(512), BUFF2(512)
    EQUIVALENCE (BUFF1(1), BUFFER(1,1)), (BUFF2(1), BUFFER(1,17))
C INITIALIZE SINE TABLE AND IO FILES
  CALL FORT (BLOCK, S, S, C, IERR)
  DEFINE FILE 99(512,1024,U,NXT99)
C
C FILE IS BROKEN INTO 16 BLOCKS AND PROCESSED ONE BLOCK AT A TIME
  DO 500 IBLOCK = 1,16
C BASELINE FOR RECCRDS IN INPUT FILE
  IRLINE = (IBLOCK-1)*32
  DC 200 LINE = 1,32
  IREC = IRLINE + LINE
C READ IN DATA FROM IBLOCK IN FILE 'IN' AND PERFORM HORIZONTAL TRANSFORMS
  CALL READIN( IN, IREC, BLOCK(1,LINE) )
  200 CALL FORT( BLOCK(1,LINE), S, S, 2, IERR )
C TRANSPOSE BLOCKS, WRITE ONTO INTERMEDIATE FILE 99
  DO 490 ISECT = 1,16
  IRMIN = (ISECT-1)*32
  DO 450 LINE = 1,32
  DO 450 ITEM = 1,32
  450 BUFFER(ITEM,LINE) = BLOCK(IRMIN+ITEM,LINE)
  IREC = IRMIN + 2*IBLOCK
  WRITE( 99*IREC-1 ) BUFF1
  490 WRITE( 99*IREC ) BUFF2
  500 CONTINUE
C
C READ IN BLOCKS FROM FILE 99, TRANSPOSE SEGMENTS WITHIN EACH BLOCK
  DO 900 IBLOCK = 1,16
  IRLINE = (IBLOCK-1)*32
  DO 700 ISECT = 1,16
  IREC = IRLINE + 2*ISECT
  READ( 99*IREC-1 ) BUFF1
  READ( 99*IREC ) BUFF2
  IRMIN = (ISECT-1)*32
  DO 700 ITEM = 1,32
  DO 700 ISEG = 1,32
  700 BLOCK(IRMIN+ISEG, ITEM) = BUFFER(ITEM, ISEG)
  DO 800 LINE = 1,32
C PERFORM VERTICAL TRANSFORMS
  CALL FORT ( BLOCK(1,LINE), S, S, 2, IERR)
  IREC = IRLINE + LINE
  800 CALL WRTOU( IOUT, IREC, BLOCK(1,LINE) )
  900 CONTINUE
  RETURN
  END
  SUBROUTINE READIN( IN, IREC, A )
    COMPLEX A(512)
    READ( IN*IREC ) A
C APPLY PHASE FACTOR (-)**(M*N) TO SHIFT ORIGIN OF FFT
C NC = 2 IF ODD, 1 IF EVEN
  NO = 1 + IAND(1, IREC)
  DO 50 N = NO, 512, 2
  50 A(N) = -A(N)
  RETURN
  END
  SUBROUTINE WRTOU( IOUT, IREC, A )
    COMPLEX A(512)
    NO = 1 + IAND(1, IREC)
    DO 50 N = NO, 512, 2
    50 A(N) = -A(N)
    WRITE( IOUT*IREC ) A
    RETURN
  END

```

```
      SUBROUTINE FILTER (M,F,HR)
C
C RETURNS LINE M OF 512 BY 512 FILTER FUNCTION H, FOR A GIVEN RADIAL
C FUNCTION HR
C
      REAL H(512),HR(370)
C ARRAY IS SYMMETRIC ABOUT POINT (257,257)
      DY = ABS( M-257.0 )
      DY2 = DY*DY
C ASSIGN VALUES IN LEFT HALF PLANE EQUAL TO HR VALUE AT NEAREST INTEGER RADIUS
      DO 500 N = 1,257
      DX = 257 - N
      DR = SQRT( DX*DX + DY2 )
      NR = DR + 1.5
500 H(N) = HR(NR)
C
C RIGHT HALF PLANE FROM LEFT HALF DATA
      DO 600 N = 258,512
600 H(N) = H(514-N)
      RETURN
      END
```

```
      SUBROUTINE JINC (HR)
C CIRCLE BLURRING FUNCTION, HR(1) IS CENTER VALUE
      DIMENSION HR(370)
      DATA PI/3.14159265/,RCMAX/5.0/
C TRANSFORM NORMALIZATION FACTOR INCLUDED IN FILTER
      ALF = 1./ (512.*512.)
      C1 = PI*RCMAX/256.
      HR(1) = ALF
      DO 150 NR = 2,370
      DR = NR-1
      PIRC = C1*DR
      BESJ1 : Bessel Function J1
150 HR(NR) = ALF*BESJ1( 2.*PIRO ) / PIRO
      WRITE (6,31) HR
      31 FORMAT (' CIRCLE BLURRING RADIAL FUNCTION'/(1X,1P10E13.3))
      RETURN
      END
```

```
      SUBROUTINE JNCINV(HR)
C INVERSE FUNCTION FOR CIRCLE BLURRING
      REAL HR(370),JINC
      DATA PI/3.14159265/,RCMAX/5.0/
      ALF = 1./ (512.*512.)
      C1 = PI*RCMAX/256.
      C2 = 1.0E-6
      HR(1) = ALF / (1.0 + C2 )
      DO 150 NR = 2,370
      DR = NR-1
      PIRC = C1*DR
      JINC = BESJ1(2.*PIRO)/ PIRO
150 HR(NR) = ALF*JINC/(JINC*JINC + C2 )
      WRITE (6,41) HR
      41 FORMAT (' CIRCLE DEBLURRING RADIAL FUNCTION'/(1X,1P10E13.3))
      RETURN
      END
```

```
//FILTER EXEC FCRTG
//FTC9F001 DD DSN=DRM.FILE,UNIT=SYSSQ,VOL=SER=ACS002,DISP=OLD
//FTS9F001 DD UNIT=SYSSQ,SPACE=(4096,512)
//FCRT DD *
        DEFINE FILE 9(512,1024,L,NXT9)
        REAL H(512),HR(370)
        COMPLEX A(512)

C      CALL FFT2( 9,9 )

C
C      APPLY FILTER FUNCTION
        CALL JNCINV( HR )
        DO 500 M = 1,512
        CALL FILTER( M,H,FR )
        READ( 9'M ) A
        DO 490 N = 1,512
490 A(N) = A(N) * H(N)
500 WRITE ( 9'M ) A

C      CALL FFT2( 9,9 )

C
        M = 257
        READ ( 9'M ) A
        WRITE (6,28) M,A
28 FORMAT (I4/10(1X,2F6.0))
        STOP
        END

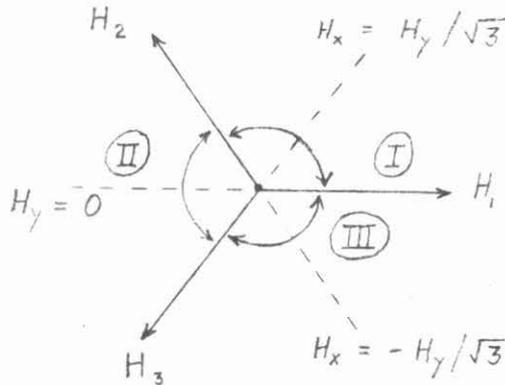
//LOAD DD *
        OBJECT DECKS FFT2 FILTER JNCINV
//STEP9 EXEC FCRTG
//WRTLN01 DD UNIT=TAPE7,VOL=SER=CIT070,LABEL=(2,BLP),DCB=(TRTCH=C),
//          DISP=(CLC,PASS)
//FTC9F001 DD DSN=DRM.FILE,UNIT=SYSSQ,VOL=SER=ACS002,DISP=OLD
//FCRT DD *
C READS DISK FILE AND WRITES REAL VALUES ON VFC TAPE
        DEFINE FILE 9(512,1024,U,NXT9)
        COMPLEX*8 A(512)
        INTEGER*2 LINE(512),LINEN
        NPRINT=0
        DO 300 M=1,512
        READ (9'M) A
        DO 305 N=1,512
        LINEN= A(N)
        IF(LINEN.LT.0) LINEN=-LINEN
        IF(LINEN.GT.255) LINEN=255
305 LINE(N)= LINEN
        CALL WRTLN(LINE,512)
C PRINT OUT SAMPLES FROM EVERY 20TH LINE
        NPRINT=NPRINT-1
        IF(NPRINT.LE.0) GO TO 900
300 CONTINUE
C EOF
        CALL WRTLN(LINE,0)
        STOP
900 NPRINT=20
        WRITE(6,41) (LINE(I),I=1,512,20)
41 FORMAT (/(1X,26I5))
        GO TO 300
        END

//LOAD DD *
        OBJECT DECK WRTLN
//
```

*Write 512 bytes on
tape*

```

SUBROUTINE COMPCN (H, NH1, NH2, NH3)
C SEPARATES COMPLEX FUNCTION H INTO THREE FIXED PHASE COMPONENTS
  COMPLEX H(512)
  INTEGER*2 NH1(512), NH2(512), NH3(512)
  INTEGER*4 HX, HY
  DATA PRCUT3/0.577350269/
  DO 200 N = 1, 512
    HX = REAL( H(N) )
    HY = AIMAG( H(N) ) * PRCUT3
    IF ( HY.LT.0.AND.FX.GT.HY ) GO TO 170
    IF ( FX.LE.-IABS(HY) ) GO TO 160
C 0 <= PHI < 120
  150 NH1(N) = HX + HY
    NH2(N) = 2 * HY
    NH3(N) = 0
    GO TO 200
C 120 <= PHI <= 240
  160 NH1(N) = 0
    NH2(N) = -HX + HY
    NH3(N) = -HX - HY
    GO TO 200
C 240 < PHI < 360
  170 NH1(N) = FX - HY
    NH2(N) = 0
    NH3(N) = -2 * HY
  200 CONTINUE
  RETURN
  END
  
```



$$\tilde{H} = H_0 e^{i\phi} = H_x + i H_y = H_1 + H_2 e^{i2\pi/3} + H_3 e^{i4\pi/3}$$

Given H_x, H_y ; Solve for H_1, H_2, H_3 in three regions:

① $H_y \geq 0$ and $H_x > -H_y/\sqrt{3}$ ($0^\circ \leq \phi < 120^\circ$)

Then: $H_1 = H_x + (1/\sqrt{3})H_y$ $H_2 = (2/\sqrt{3})H_y$ $H_3 = 0$

② $H_x \leq -|H_y|/\sqrt{3}$ ($120^\circ \leq \phi \leq 240^\circ$)

Then: $H_1 = 0$ $H_2 = -H_x + (1/\sqrt{3})H_y$ $H_3 = -H_x - (1/\sqrt{3})H_y$

③ $H_y < 0$ and $H_x > H_y/\sqrt{3}$ ($240^\circ < \phi < 360^\circ$)

Then: $H_1 = H_x - (1/\sqrt{3})H_y$ $H_2 = 0$ $H_3 = -(2/\sqrt{3})H_y$

```
      SUBROUTINE SCALEX (HC,LINE)
C   BINARY SEARCH IN A PRECOMPUTED SCALE TABLE GIVES INTEGER VALUE 0-255
C   FOR EVERY VALUE HC.
      INTEGER*2 LINE(1024)
      INTEGER*4 IBITS(8)/128,64,32,16,8,4,2,1/
      REAL HC(1024),SCALE(256)
      DO 500 N = 1,1024
      HCN = HC(N)
      INDEX = 0
C   BITS OF INDEX DETERMINED BY BINARY SEARCH
      DO 510 I = 1,8
      NEWNDX = INDEX + IBITS(I)
510 IF ( HCN.GE.SCALE(NEWNDX)) INDEX = NEWNDX
500 LINE(N) = INDEX
      RETURN
      ENTRY SCALE1
C   READ IN NONLINEARITY CURVE, 16 VALUES OF LOG DIFFRACTION AMPLITUDE
C   VS DN NUMBER, INCREASING ORDER. GENERATE SCALE TABLE FOR BINARY SEARCH.
      REAL DN(16),TN(16)
      READ (5,10) (CN(I),TN(I),I=1,16)
      WRITE(6,10) (DN(I),TN(I),I=1,16)
10  FORMAT (8F10.3)
      DO 800 I = 1,256
      DNI = I - 1
      TI = YINTERP (DN,TN,DNI,16,1)
800 SCALE(I) = 10.0**TI
      WRITE (6,43) SCALE
43  FORMAT(/' SCALING FUNCTION'/(1X,10F13.6))
      RETURN
      END
```

Linear interpolation routine

```
      SUBROUTINE SCALEQ ( LINE )
C   RESCALES LINE TO COMPENSATE FILM NONLINEARITIES.
      INTEGER*2 LINE(1024),ISCALE(1000)
      DO 100 I = 1,1024
      IF( LINE(I).GT.999 ) LINE(I) = 999
100 LINE(I) = ISCALE( LINE(I) + 1 )
      RETURN
C
      ENTRY SCALE1
C   READ IN NONLINEARITY CURVE, 16 VALUES OF DIFFRACTION INTENSITY VS DN NUMBER,
C   INCREASING ORDER, AND GENERATE 1000 POINT SCALING FUNCTION.
      DIMENSION DN(16),T(16)
      READ (5,10) (DN(I),T(I),I=1,16)
      WRITE(6,10) (DN(I),T(I),I=1,16)
10  FORMAT (8F10.3)
C   RESCALE T ON LINEAR AMPLITUDE SCALE, 0 TO 999
      FACTOR = 999.0 / 10.0**T(16)
      DO 200 I = 1,16
200 T(I) = FACTOR * 10.0**T(I)
      WRITE (6,11) T
11  FORMAT (1X,'AMPLITUDE TRANSMITTANCE FUNCTION'/(1X,8F10.3))
      DO 250 I = 1,1000
      TI = I-1
C   ASSIGN A DN NUMBER FOR EVERY POSSIBLE VALUE OF T
      IDN = YINTERP( T,DN,TT,16,1 )
250 ISCALE(I) = MAX( 0, IDN )
      WRITE (6,12) ISCALE
12  FORMAT(1X,'SCALING FUNCTION'/(1X,25I5))
      RETURN
      END
```

Fast routine

LINE is initially scaled 0-1000

Linear interpolation

```

SUBROUTINE PHASE (A,M)
C APPLIES PHASE DISPERSION FUNCTION TO ARRAY A.
  COMPLEX A(512),W(32,32)
  J = IAND (M,31) + 1
  DO 200 N = 1,512
    I = IAND (N,31) + 1
    200 A(N) = A(N) * W(I,J)
  RETURN

```

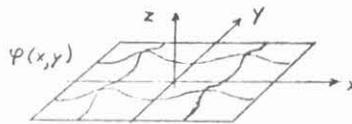
$$\tilde{a}_{nm} = a_{nm} e^{i\varphi_{nm}}$$

```

C
C INITIALIZATION OF PHASE FUNCTION
  ENTRY PHASE1 (PMAX)
  DATA PI/3.14159265/
  DO 300 J = 1,32
    DO 300 I = 1,32
      PIJ = PMAX * COS (PI*I/16.) * COS (PI*J/16.)
    300 W(I,J) = CMPLX ( COS(PIJ),SIN(PIJ) )
  RETURN
  END

```

$$\varphi_{nm} = \varphi_{max} \cos\left(\frac{2\pi n}{32}\right) \cos\left(\frac{2\pi m}{32}\right)$$



```

SUBROUTINE W2048 ( VAL,NEXP )
C RETURNS EXP ( I2PI * NEXP / 2048 )
  DIMENSION C(513)
  COMPLEX VAL
C CCSINE AND SINE TABLE INDICES, MODULO 512
  IC = IAND (MASK1,NEXP) + 1
  IS = 514 - IC
C IPHASE FROM BITS 10 AND 11
  NEXPT = NEXP
  CALL SHIFTR(NEXPT,9)
  IPHASE = IAND(MASK2,NEXPT) + 1
  GO TO (1,2,3,4), IPHASE
  STOP 187
  1 VAL = CMPLX( C(IC), C(IS) )
  RETURN
  2 VAL = CMPLX(-C(IS), C(IC) )
  RETURN
  3 VAL = CMPLX(-C(IC),-C(IS) )
  RETURN
  4 VAL = CMPLX( C(IS),-C(IC) )
  RETURN
  ENTRY WINT
C INITIALIZATION OF CCONSTANTS AND COSINE TABLE
  MASK1 = 512 - 1
  MASK2 = 3
  ALF = 3.14159265 / 1024.0
  DO 100 I = 1,513
    100 C(I) = COS( ALF*(I-1) )
  RETURN
  END

```

Logical AND

Shift right, 9 bits

$$VAL = e^{i2\pi n / 2048}$$

-150-
APPENDIX D

INTERFEROGRAMS OF VARIOUS HOLOGRAM AND MASK SUBSTRATES

A simple and inexpensive, large-aperture interferometer can be built using a hologram, as shown in Figure D1. The hologram is recorded with no test object. The developed hologram, when replaced in the interferometer, causes a mixing of the two waves. The interference pattern is recorded on high contrast sheet film (Kodak 4154, Dev. D-11, 5 min). Expensive optics is not required because the hologram compensates any distortions in the lens. By adjusting the location of one of the pinholes, a prismatic or spherical correction can be made. The hologram used here was a high-efficiency dichromated-gelatin plate.

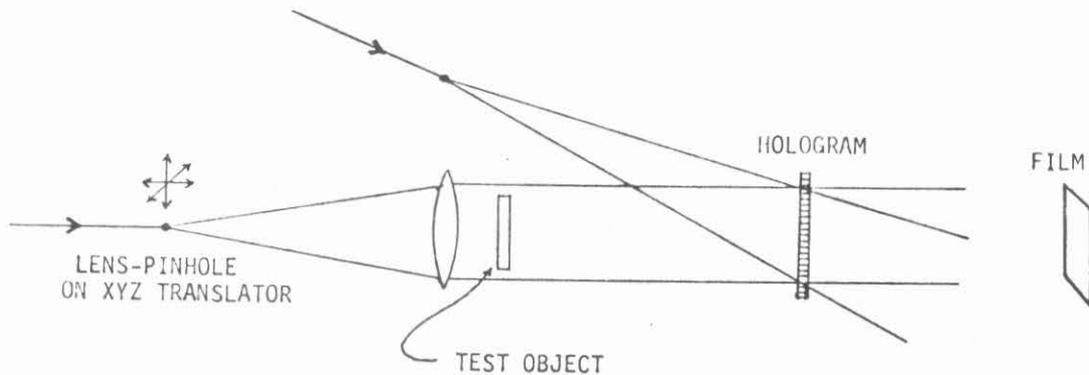


Fig. D1. A simple large-aperture interferometer for making transmission interferograms.

Fringes of maximum contrast were obtained with nearly equal beam ratios.

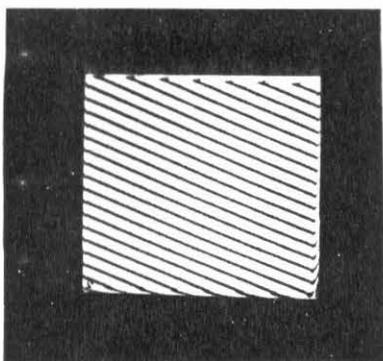
The interferograms in Figure D2 were made using the above interferometer with an argon ion laser ($\lambda = 4880\text{\AA}$). The aperture in each case is 2" square. Figure D2a shows the high-contrast flat fringes obtainable over this aperture. The interferometer was offset slightly to produce the fringes. Interferograms (b) through (i) show the phase distortions introduced by various hologram media and their substrates. The performance of a hologram spatial filter will be severely degraded by such distortions. For application as a spatial filter, spherical distortions will cause only a slight shift in the location of the image plane. Therefore, the interferogram from a best fitting spherical wavefront is shown when this resulted in some improvement.

The interferograms (b) and (c) are for normally processed dichromated gelatin plates (Appendix G), obtained from Kodak 649F 4"x5" holographic plates. Most of the distortion is due to the gelatin layer, as shown by the interferogram for a bare glass plate (d). The gelatin layer on the 0.250" Kodak microflat plates is apparently much more uniform as shown by the interferogram (e) obtained from a bleached hologram.

A simple liquid gate was made by placing a 0.060" glass plate from a Kodak 2"x2" precision flat high resolution plate against the emulsion side of a DCG hologram on 0.040" glass. A drop of xylene fills the space between the plates. The resulting interferograms are shown in (f) and (g). Other commonly used hologram substrates are polyester film bases such as found in (h) Estar film base from 35 mm

Kodachrome II film, and (i) Estar thick base from 4"x5" Kodak Ectapan sheet film.

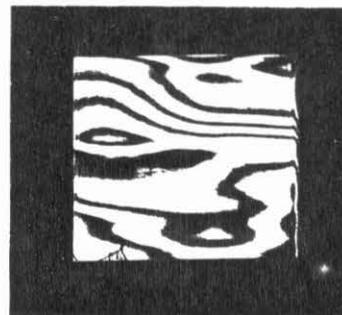
Interferograms (j) through (r) show phase distortions of various possible mask substrates. Spherical distortion is not tolerable in the masks for the holograms of Chapter 3. The Kodak high resolution plates on 2"x2" "precision flat" 0.060" glass show considerable variation in the distortions from one plate to another. Interferograms (j) and (k) are indicative of the best and worst plates in this series. The emulsion on these plates is very flat and most of the distortion arises from the glass plate, as seen by comparison of the same mask with (l) and without (m) the emulsion. Interferogram (n) is for a 1"x3"x0.0375" commercial polished quartz slide (Engelhard Industries). For the hologram mask generation application of Chapter 3, a very thin substrate could tolerate much worse phase distortions. An interferogram (o) for a very thin microscope cover glass is shown (Corning No. 0, 22 mm², 0.11 mm thick). A typical microscope slide is shown in (p): (Corning #2947, 75 mm by 50 mm by 1.2 mm thick). (q) shows a 0.030" thick slide cover glass of the type used for mounting 2" square photographic slides. The effects of a rolling mill are apparent in (p) and (q).



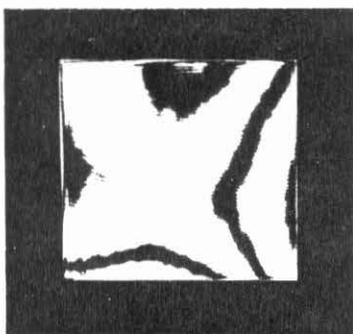
(a) No plate, tilted wavefront



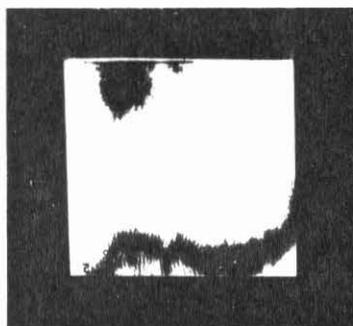
(b) DCG plate, 0.040" glass



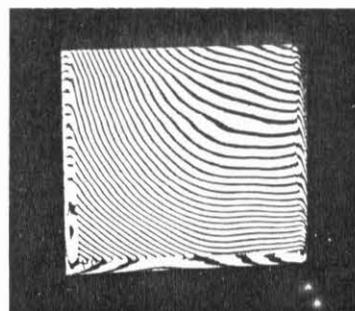
(c) Same as (b). BFS



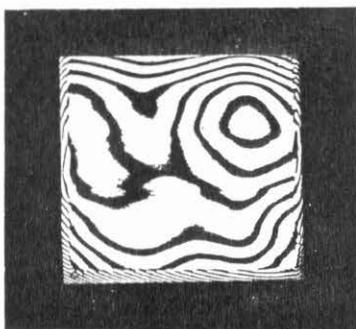
(d) Bare 0.040" glass plate. BFS



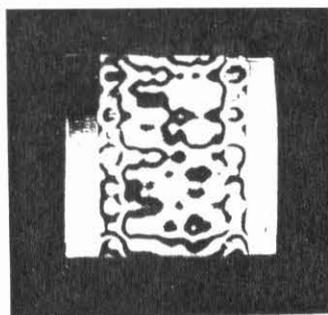
(e) Bleached hologram on microflat. BFS



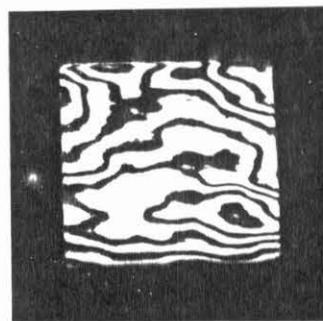
(f) Liquid gate



(g) Same as (f). BFS

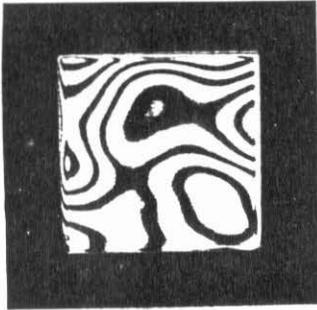


(h) Estar base, 35 mm film

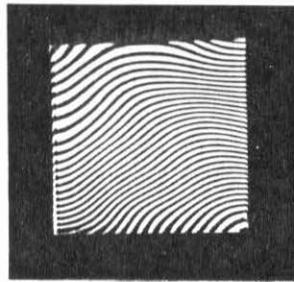


(i) Estar thick base 4x5 sheet film

Fig. D2 Interferograms for various hologram substrates made using the interferometer of Fig. D1. BFS is best fitting spherical wavefront



(j) Mask substrate 1



(k) Mask substrate 2



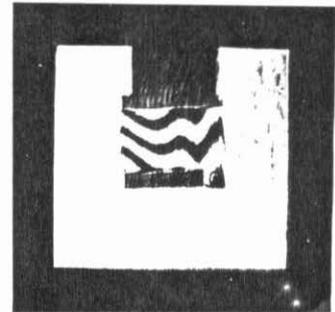
(l) Mask with emulsion



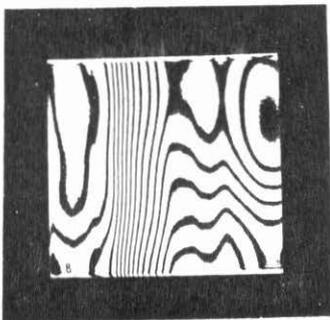
(m) Same without emulsion



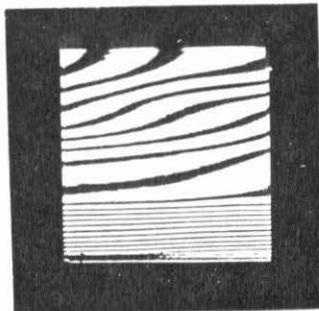
(n) Quartz slide,
0.0375"



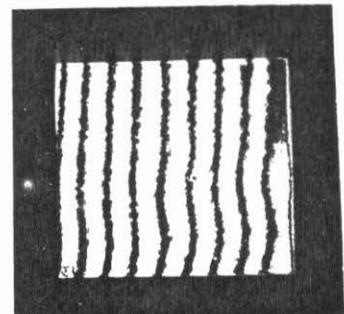
(o) Microscope cover
glass 0.11 mm



(p) Micro-slide (1.2mm)



(q) Slide cover glass



(r) Saran-wrap,
tilted wavefront

Fig. D2. Interferograms for various possible mask substrates

APPENDIX F

Measurement of Noise and Diffraction Efficiency in
Holographic Media

This appendix provides details of the experimental setup and procedures used in generating the noise plots of Chapter 3. Figure F1 is a diagram of the experimental setup. A 4880\AA beam from a power stabilized argon laser is directed at a holographic grating which is synthesized by exposure at the same wavelength in the interferometer of Chapter 3. The intensity of the first order wave is measured with a small aperture photodetector, and the relative amplitude is computed as the square root of the diffraction efficiency. This amplitude is plotted for gratings of various exposure levels in different media.

Noise is measured by moving the detector an angle θ off the diffracted spot. The measured noise may originate from any of three sources: the hologram recording system, the holographic medium, or the noise measurement system. Great care is taken to minimize noise sources in the recording and measurement systems. The gratings are recorded with spherical waves in order to eliminate a lens with its scattering and multiple reflections near the hologram plane in Figure F1. The self-convergent property of the hologram grating will then focus all the diffracted power into a small spot at the detector aperture. The Gaussian profile of the laser beam assures a sharp focus with very low sidelobes.

The recording system is enclosed in a flat black box, and stray light is minimized by careful placement of black screens. Ideally, the hologram should see nothing but two bright points. The major source of noise seems to be scattering from the inside edges of the circular apertures used to define the beams emerging from the pinholes. Without these apertures there would be much more scattering from the lens mount and other components in the system.

A worst case calculation may be made of the expected noise level from a perfect grating in the measurement system of Figure F1. We will assume that each of the diffracted waves carries the same relative noise distribution along with its perfect Gaussian profile. If the grating is strong, then noise from the first-order wave will be dominant. If the grating is weak, then most of the noise will be due to scattering from the straight-through, or zero-order beam. These levels are plotted in Figure 4.6 (Φ_{M_1} and Φ_{M_0}). Φ_{M_0} is the measured scattering from a clean glass plate 10^0 off axis. Φ_{M_1} is 58 dB below the diffracted amplitude of the assumed perfect grating. The 58 dB figure was determined from measurements on the main beam illustrated in Figures F3 through F6. Figure F3 shows the power distribution for the bare beam (at 640 mm from the hologram plane). A Gaussian is plotted for comparison. The plot does not reach the top of the graph at $\theta = 0$ because the beam width exceeds the detector aperture and some power is spilled over. The departure from the Gaussian profile may be due to imperfections in the laser mode, or scattering from any of the optical surfaces which the beam encounters. In order to provide a more realistic profile for the beam, which is assumed to be sharply

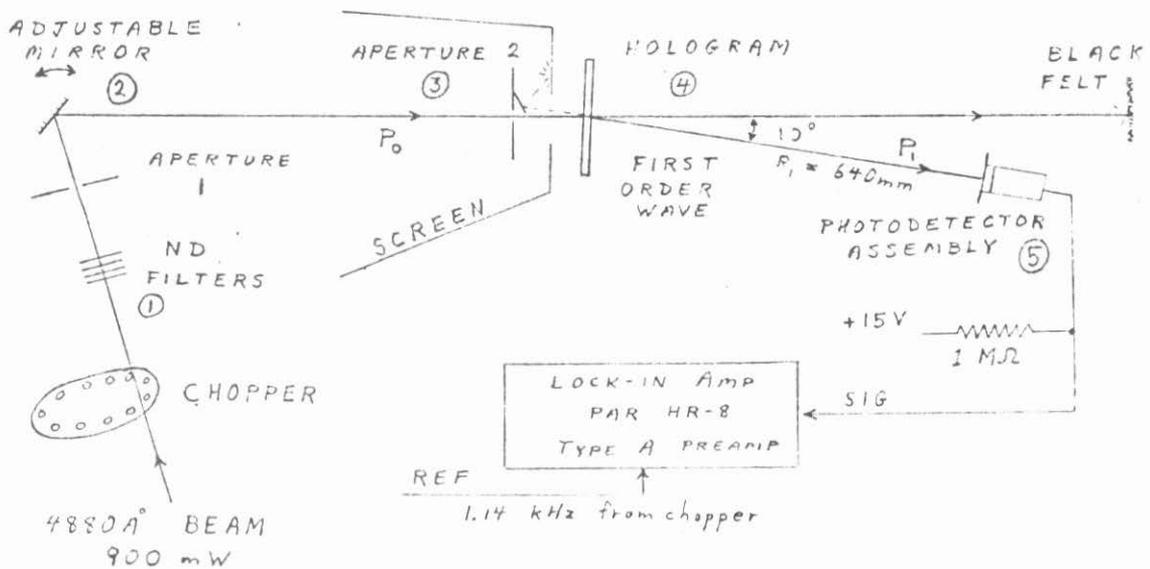
focused by the perfect hologram, lenses were placed just after the mirror in Figure F1. These lenses focus the beam at the detector but also introduce some additional noise, as shown in Figure F4. Placing a glass plate in the hologram holder adds still more noise (Fig. F5). This noise level (at 7.8 μ m) was taken as a worst case expectation for a perfect hologram on a normal glass plate. The scattering from the lenses would not be present in the case of a perfect hologram.

An unexposed (but normally processed) dichromated gelatin plate shows slightly higher noise levels than the bare glass (see Fig. F6). These plates have a milky appearance due to wide angle scattering.

Noise in the recording system would have the effect of lowering the 58 dB figure for a noise free hologram. The highest measured signal-to-noise ratio was 51 dB (DCGI, grating number 6). It is conceivable that the recording system is introducing noise at this level. The validity of the 51 dB measurement can only be established by comparison with some lower noise medium. We can, however, be assured of the validity of the measurements for the more noisy media.

The experimental data were analyzed and results plotted in real time using an HP 9820 calculator and plotter, the program for which is shown in Figure F2. This procedure allows immediate correction of experimental blunders which would otherwise require repetition of the entire experiment.

Fig. F1. Experimental setup used in measurements of hologram noise and diffraction efficiency



NOTES

- ① Schott glass neutral density filters, lightest toward beam. Remove and replace in same order. ND = .30, 1.22, 1.22, 1.24.
- ② Screw adjustments to direct beam to various areas of hologram.
- ③ Aperture 2 is inserted on most sensitive measurements to eliminate stray light. Piece of red plastic catches the reflection off the back surface of the hologram and directs it to a flat black screen.
- ④ Hologram focuses first order wave into photodetector aperture without scattering from additional lenses. Emulsion side faces detector.
- ⑤ Photodetector assembly: Aperture (D = 1/16" diam.), opal glass diffuser, silicon photodiode (UDT PIN 10C). Mounted on x-y translator. Center on first order spot for diffraction efficiency measurement, move 5 mm to side for noise measurement.

TABLE 1: Notes on Experimental Measurements

Diffraction Efficiency

$$\eta = P_1/P_0 = \eta_0 (V_1/V_0) 10^{(ND-ND_0)}$$

V,ND : detector signal voltage, neutral density factor

η_0, V_0, ND_0 : diffraction efficiency, voltage, and density for a calibration grating

Photodetector Solid Angle

$$d\Omega = A/R_1^2 = \pi D^2/4R_1^2$$

$$D = 1/16" = 1.59 \text{ mm}; \quad A = 1.98 \text{ mm}^2$$

$$R_1 = 640 \text{ mm}$$

$$d\Omega = \underline{4.8 \cdot 10^{-6} \text{ sr}}$$

Spatial Frequency

$$v = \sin(\theta)/\lambda \approx \frac{\theta(\text{mr})}{.488} \text{ (lines/mm)}$$

$$\text{for } \lambda = .488\mu$$

Spatial Frequency Aperture

$$\delta v_x \delta v_y = d\Omega/\lambda^2 = 20.2 \text{ (lines/mm)}^2$$

Noise Power Spectral Density (fraction of incident power scattered into solid angle spanned by unit spatial frequency)

$$\Phi = \eta/\delta v_x \delta v_y \text{ (mm)}^2$$

$$\log \Phi = 2 \log \sqrt{\eta} - 1.31$$

Scattering per Unit Solid Angle

$$\Phi_{\Omega} = \eta/d\Omega = \Phi/\lambda^2$$

$$\log \Phi_{\Omega} = \log \Phi + 6.62$$

$$(\text{for } \lambda = .488 \cdot 10^{-3} \text{ nm})$$

Base Level Noise (noise from measurement system with ideal noise-free hologram)

$$\log(\Phi_M) = \log(\Phi_{M_0} + \Phi_{M_1})$$

Φ_{M_0} : noise scattered from main beam

$$\log(\Phi_{M_0}) = -9.45 \text{ (measured with bare glass plate and detector } 10^{\circ} \text{ off axis)}$$

Φ_{M_1} : noise scattered from first order wave

$$= 2 \log \sqrt{\eta} - 1.31 - 5.8$$

(58 dB below first order spot, see Fig. F5)

Fig. F2. Calculator program for HP 9820 used to perform real time analysis and plotting of experimental data.

```
0:      .301+R1;1.52+R2;
1:      2.74+R3;3.98+R4+
2:      ENT "ETA",R10,"M
3:      V",R11,"ND",R12+
4:      R10/R11*10+(-R12
5:      )+BF
6:      SPC 2;PRT "SCALE
7:      FACTORS",R10,R1
8:      1,R12;FLT ;PRT B
9:      ;FXD ;SPC 1+
10:     "NEXT";ENT "ND =
11:     ",A;DSP A;STP +
12:     ENT X,"MV",Y;
13:     SPC 1;PRT X,Y,A+
14:     0.5*(A+LOG (YB))
15:     +Y;IF FLG 1;X/.6
16:     4+X;GTO "B"+
17:     LOG X+X+
18:     "B";PRT X,Y;0+R1
19:     3+
20:     ENT "PLOT POINT
21:     ?",R13+
22:     IF R13=0;GTO "NE
23:     XT"+
24:     LTR X,Y,221+
25:     PLT "X"+
26:     GTO "NEXT"+
27:
28:     14:
29:     "GA";DSP "PEN PO
30:     SITION A";STP +
31:     15:
32:     ENT "XMIN",R8,"X
33:     MAX",R9;SCL R8,R
34:     9,-4.5,0+
35:     16:
36:     AXE R8,-4.5,1,0.
37:     5+
38:     17:
39:     AXE R9,0,0,0+
40:     18:
41:     SPC 3;FXD 3;PRT
42:     "PLOT A",R8,R9;
43:     CFG 1;DSP "SYMB0
44:     L ?";GTO 12;STP
45:     +
46:     19:
47:     "GB";DSP "PEN PO
48:     SITION B";STP +
49:     20:
50:     SCL 0,16,-4.5,0+
51:     21:
52:     AXE 0,-4.5,1,0+
53:     22:
54:     AXE 16,0,0,0+
55:     23:
56:     SPC 3;PRT "PLOT
57:     B";SPG 1;DSP "SY
58:     MBOL ?";GTO 12;
59:     STP +
60:     24:
61:     END +
62:     R349
```

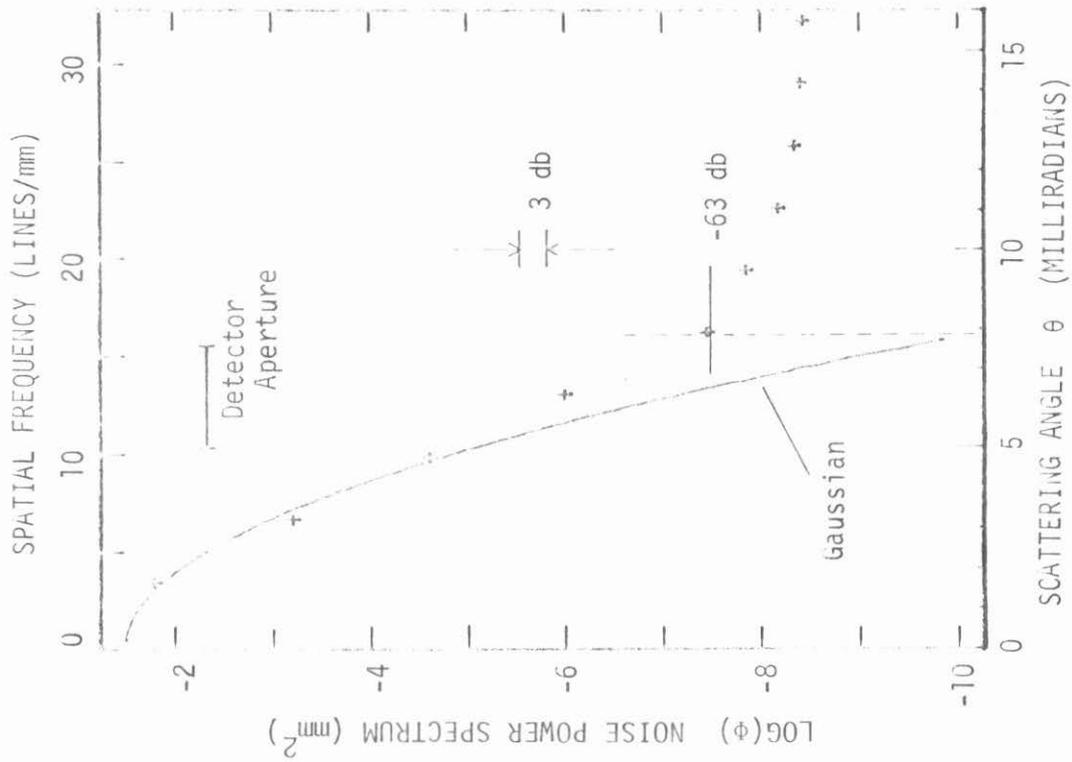


Fig. F3. Power distribution for main beam.

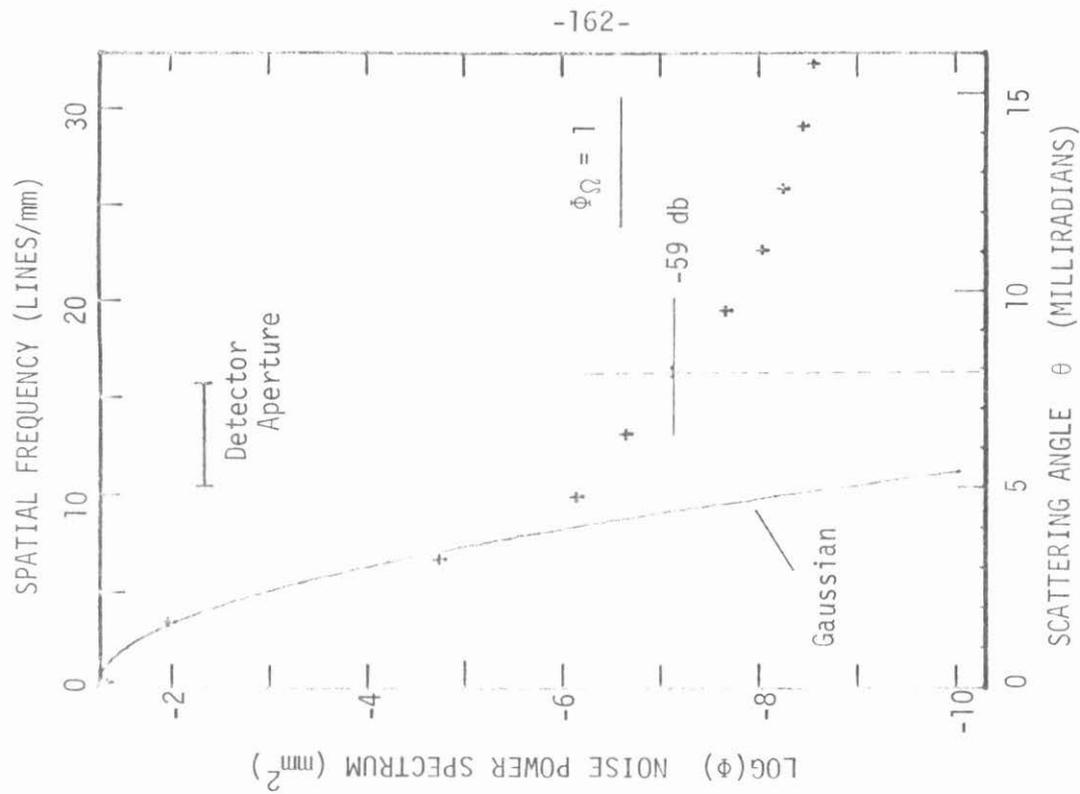


Fig. F4. Main Beam with focusing lenses.

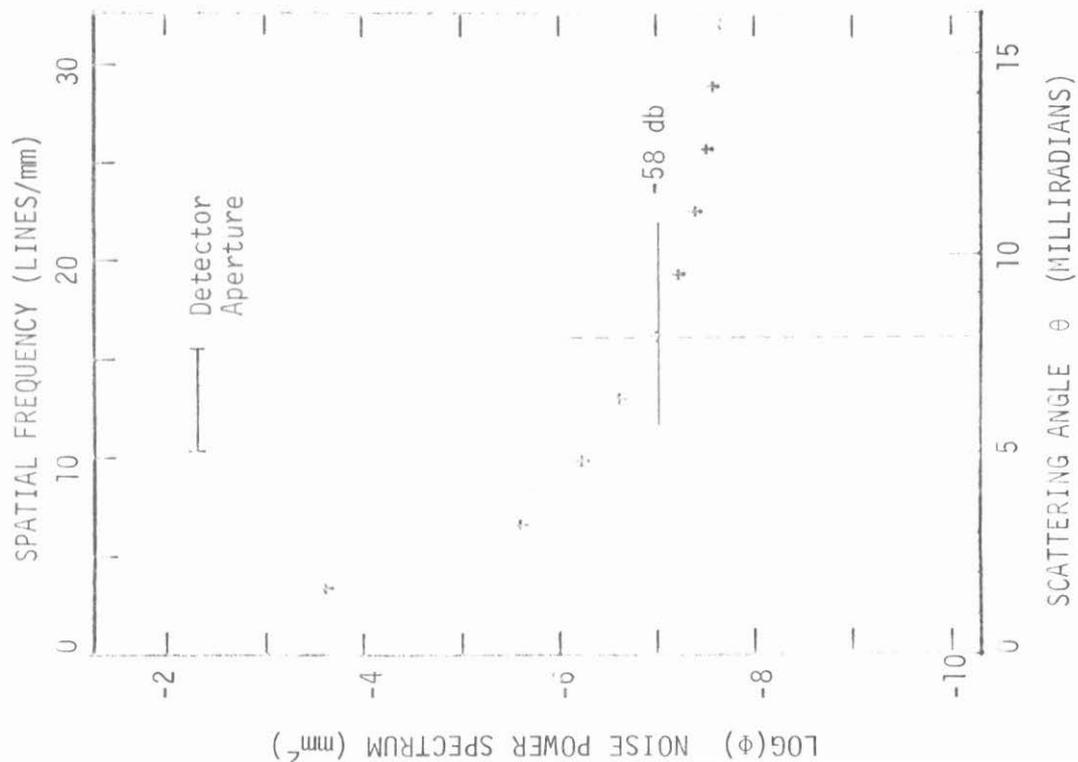


Fig. F5. Main beam with lenses and glass plate in hologram holder.

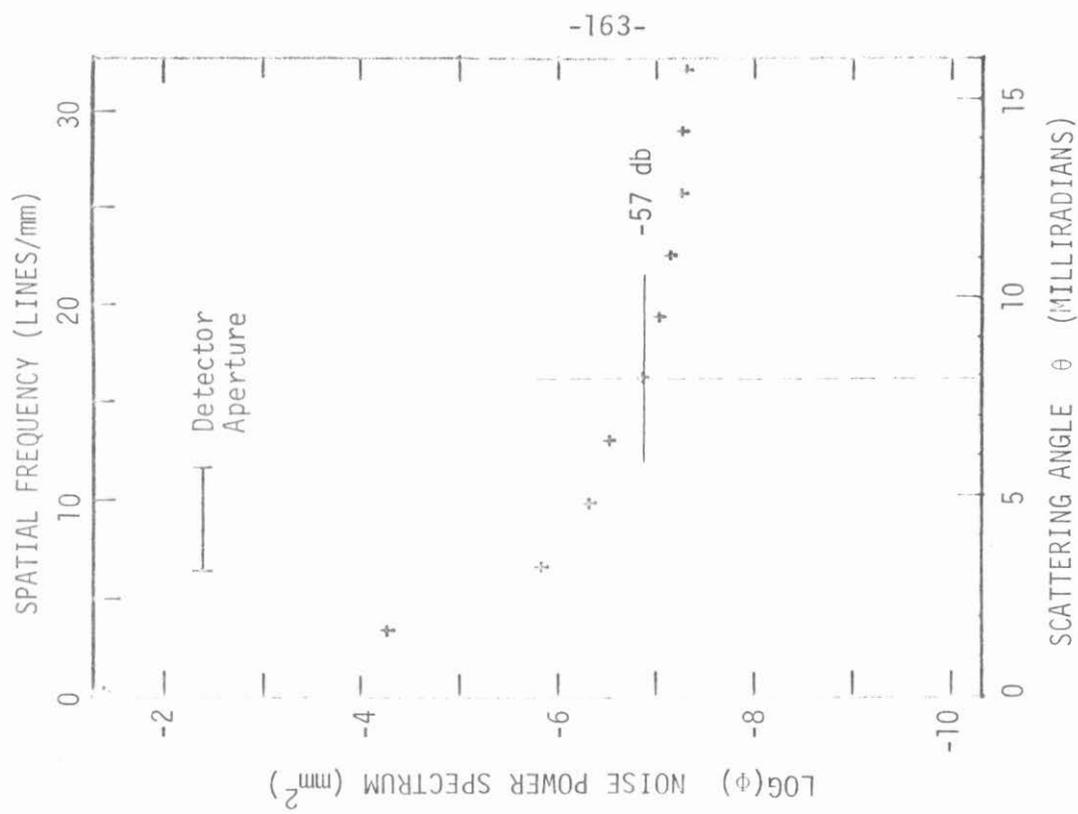


Fig. F6. Main beam, lenses, and unexposed dichromated gelatin plate.

APPENDIX G

Hologram Processing Schedules

The following schedules were used in processing the holograms of this thesis.

The most recent and thorough work on dichromated gelatin seems to be that of D. H. Close et al.^(1,2) The DCG processes used in this thesis are an adaptation of the methods presented by M. Chang⁽³⁾ and D. H. Close. In dichromated gelatin, a compromise must be made between sensitivity and noise. The more sensitive, higher efficiency Process II shows more light scattering, and the plates tend to be slightly milky. The dehydration step (D) in these processes is critical. Rapid dehydration at room temperature gives low-noise, less sensitive holograms. Extending the dehydration process in steps of increasing alcohol concentration and raising the temperature of the alcohol baths will increase the sensitivity and noise. Prehardening of the gelatin in Kodak Rapid Fixer results in lower noise and more stable gelatin layers. Noise and diffraction efficiency data for the two DCG processes listed here are given in Chapter 4.

References

1. D. H. Close, A. Au, A. Graube, "Holographic Lens for Pilot's Head-Up Display", Hughes Research Labs, Tech. Report N62269-73-C-0388, Naval Air Development Center (1974), Table 13 (p.93) "Processing schedule for dye sensitized DCG".
2. D. H. Close, A. Graube, "Materials for Holographic Optical Elements" Hughes Research Laboratories, Tech. Report AFML-TR-73-267 (Oct. 1973).
3. M. Chang, "Dichromated Gelatin of Improved Optical Quality", Appl. Opt. 10, 2550 (1971).

PROCESSING OF DICHROMATED GELATIN HOLOGRAMS*

Process I

- A.
- 1) Fix Kodak 649F plates 10 min. in part A of Kodak Rapid Fixer.
 - 2) Wash in water starting at 21⁰C, rising to 32⁰C over 7 min. period and holding at 32⁰C for 15 min.
 - 3) Cool in room air 1 min (plates horizontal--gelatin is soft).
 - 4) Rinse 30 sec in distilled water with 2 drops per liter Photo-Flo 600.
 - 5) Dry[†] overnight (plates horizontal).
- B.
- 1) Soak 2 min in room temperature water.
 - 2) Harden 10 min in Kodak Rapid Fixer (A and B).
 - 3) Wash 15 min 21⁰C running water.
 - 4) Rinse 30 sec distilled water with Photo-Flo.
 - 5) Dry[†] overnight.
- C.
- 1) Sensitize by soaking 5 min in 7.5% ammonium dichromate with 2 drops per liter Photo-Flo 600.
 - 2) Dry[†] overnight with plates vertical.
- D.
- 1) Wipe crystals from glass side of plate with damp cloth.
 - 2) Expose plates one to five days after sensitization.
 - 3) Wash 5 min in 21⁰C running water.
 - 4) Dehydrate rapidly in 100% isopropyl alcohol, 30 sec with agitation.
 - 5) Dry vertically in flowing dry air.

*Milton Chang, "Dichromated Gelatin of Improved Optical Quality", Applied Optics 10, 2550 (1971).

†Dry plates in dark box with flowing air, humidity set to 32% by saturated CaCl₂ solution.

Process II

Same as process I, except:

- C. 1) 5% ammonium dichromate
- D. 3) 4) 5 min each: Running water (21⁰C), 50% isopropanol (21⁰C),
100% isopropanol (21⁰C).

Processing of 649F Plates^{*}

Develop	5 min	D-19	21 ⁰ C
Stop	30 sec		
Fix	3 min	Rapid Fixer	
Wash	10 min	Running water	
Rinse	30 sec	Deionized water with photoflow	
Dry			

^{*}Kodak Plates and Films for Scientific Photography, publication No.P-315, (Nov. 1973), p. 35d: Table 8, Processing Information Summary.

Bromine Bleaching Process^{*}

Process plate normally as above.

Bleach one hour with bromine vapor in a closed vessel.

^{*}A. Graube, Hughes Research Report No. 484, Feb. 1974, "Advances in Bleaching Methods for Photographically Recorded Holograms" (submitted to Applied Optics).

LLG Bleach Process*

- 1) Develop 649F Plates 5 min, D-19, 21°C.
- 2) Acid stop bath 30 sec
- 3) Rapid Fix 3 min
- 4) Rinse in distilled water 10 min
- 5) Bleach 10 min, 5% cupric bromide.
- 6) Rinse 30 sec, distilled water.
- 7) Desensitize 3 min, one part A, ten parts B.
- 8) Rinse 10 min, distilled water.
- 9) Dry in room air.

Desensitizer Solutions

A

Potassium permanganate 5g
DW to make 1 liter

B

Concentrated sulfuric acid 10 ml
Potassium bromide 40 g
DW to make 1 liter

(Mix solutions immediately before using.)

*H. Lehmann, J. P. Lauer, J. W. Goodman, Applied Optics 9, 1948 (1970),
"High Efficiencies, Low Noise, and Suppression of Photochromic Effects
in Bleached Silver Halide Holography."

Hariharan's Bleach Process for 649-F Plates*

- 1) Expose
- 2) Develop 5 min D-19 with 1.0g/liter sodium thiosulphate
- 3) Rinse 3 min distilled water (2 changes)
- 4) Bleach 5 min (see below)
- 5) Rinse 6 min distilled water (4 changes, add 2 ml/liter photo-flow to the last bath)
- 6) Dry at room temperature and 40-60% RH.

<u>Bleach</u>	<u>Stock Solution A</u>	<u>Stock Solution B</u>
	potassium dichromate 8g	potassium iodide 2g
	conc. sulphuric acid 10 ml	DW to make 1 liter
	DW to make 1 liter	

Mix 1 part A, 1 part B, 8 parts DW, just before use.

*P. Hariharan, et al., Optics Comm. 6, 75 (1972), "Simplified, low-noise processing technique for photographic phase holograms."

HNO₃ Reversal Bleaching*

Immerse 649-F developed (but unfixed) plate in a shallow basin of water. Add 1:1 concentrated nitric acid slowly with constant agitation until the plate clears. Rinse in water and dry.

*M. Chang, N. George, Applied Optics 9, 713 (1970), "Holographic Dielectric Grating: Theory and Practice."

APPENDIX H

Nonlinear Distortions in Holographic Media

In this appendix we analyze some of the distortions in the reconstructed wavefront which result from various nonlinearities in the response of the hologram medium and in the diffraction process. The general approach is to expand the nonlinearity in a Taylor series and derive the magnitude of the first significant intermodulation products. Other analyses of hologram nonlinearities have been presented by Goodman^(1,2) and by Friesem and Zelenka⁽³⁾, using transform methods. The analysis presented here is somewhat simpler and yields the same essential results.

A modulated grating hologram may be formed in any holographic medium, but the most likely applications will involve a thin amplitude hologram, a thin phase hologram, or a thick phase hologram. Each of these three cases is considered in detail.

H.1 Hologram Recording Geometry

In the analysis to follow, we will consider the hologram recording geometry of Fig. H1. The following definitions are to apply:

w_1 : uniform amplitude plane wave (reference wave)

$\tilde{w}_2(x,y)$: complex amplitude of slowly varying wavefront, spatial bandwidth $\ll K/2\pi = \nu_g$ (lines/mm) grating frequency

Fringe spacing: $d = 2\pi/K$. Grating number $K = 2k_a \sin \theta_a$.

Wavenumber $k_a = 2\pi/\lambda_a$, where λ_a and θ_a are the wavelength and incident angle in air. Combining these definitions gives the Bragg law: $\lambda_a = 2d \sin \theta_a$

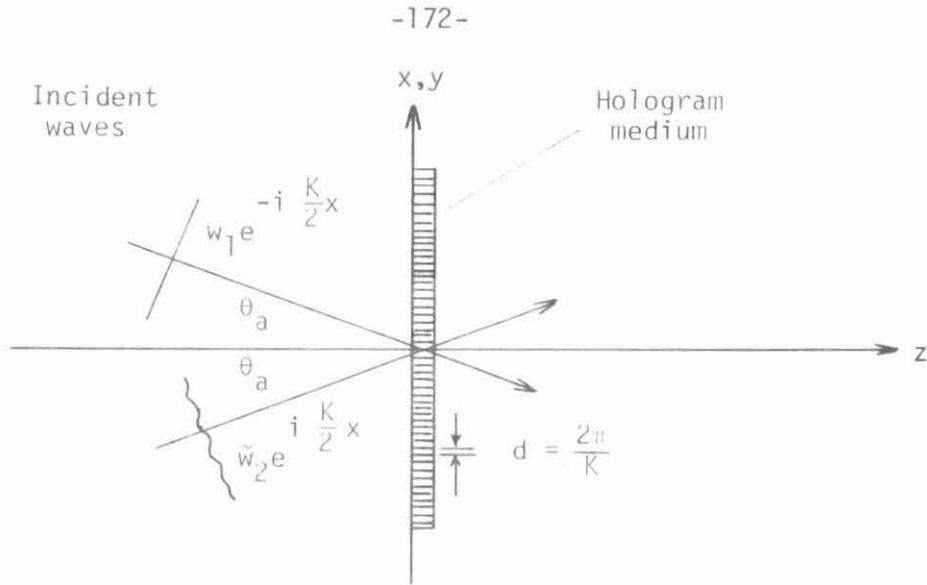


Fig. H1. Hologram Recording Geometry for Analysis of Nonlinearities

τ : exposure time

\sim denotes a complex function

The exposure at the x,y plane (which we assume to be uniform in z if the medium has finite thickness) will be given by

$$\begin{aligned}
 E(x,y) &= \tau |w_1 e^{-i(K/2)x} + \tilde{w}_2 e^{i(K/2)x}|^2 \\
 &= E_0 + \frac{1}{2} \tilde{E}_m e^{iKx} + \frac{1}{2} \tilde{E}_m^* e^{-iKx} \\
 &= E_0 + E_m \cos(Kx + \phi)
 \end{aligned} \tag{H.1}$$

where

$$\begin{aligned}
 E_0(x,y) &= \tau (w_1^2 + |\tilde{w}_2(x,y)|^2) \\
 \tilde{E}_m(x,y) &= 2\tau w_1 \tilde{w}_2(x,y) = E_m(x,y) e^{i\phi(x,y)}
 \end{aligned}$$

The modulation exposure E_m is always less than E_0 in magnitude. In most hologram recording situations, the reference wave is made much stronger than the subject wave. In this case $|\tilde{w}_2| \ll w_1$, and the bias exposure can be considered constant. This will be the assumption in the analysis to follow.

H.2 Thin Amplitude Holograms

The quantity of interest in the case of a thin amplitude hologram is the real amplitude transmittance T_A , which will be a function of exposure (Fig. H2). The medium is assumed to be infinitesimally thin.

The transmittance function can be expanded in a Taylor series about E_0 as follows:

$$T_A(E) = \alpha_0 + \alpha_1 \Delta E + \alpha_2 \Delta E^2 + \alpha_3 \Delta E^3 + \dots \quad (\text{H.2})$$

where $\Delta E = E - E_0 = E_m \cos(Kx + \phi)$

$$\alpha_0 = T_A(E_0), \quad \alpha_1 = T'_A(E_0), \quad \dots, \text{ etc.}$$

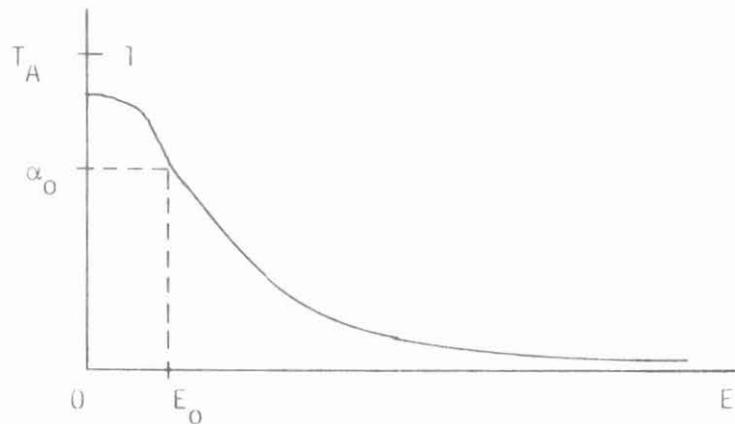


Fig. H2. Amplitude Transmittance vs. Exposure for a Thin Hologram

The coefficients α will depend on the medium, the method of processing, and the bias exposure E_0 . The transmittance function can be converted to a Fourier series as follows:

$$T_A(E) = \sum_{n=0}^{\infty} \alpha_n (\Delta E)^n$$

where

$$\Delta E = \frac{1}{2} \tilde{E}_m(x,y) e^{iKx} + \frac{1}{2} E_m^*(x,y) e^{-iKx}$$

Expand $(\Delta E)^n$ using binomial theorem

$$(\Delta E)^n = \sum_{k=0}^n \binom{n}{k} \frac{1}{2^n} \tilde{E}_m^{n-k} \tilde{E}_m^{*k} e^{i(n-2k)Kx}$$

$$\binom{n}{k} = \frac{n!}{(n-k)! k!}$$

Each term n in the nonlinear transmittance function will diffract a set of waves with phase factors e^{inKx} , $e^{i(n-2)Kx}$, ..., $e^{-i(n-2)Kx}$, e^{-inKx} . Summing all terms for which $n-2k = \ell$ gives

$$T_A(E) = \sum_{\ell=-\infty}^{\infty} \tilde{T}_{\ell} e^{i\ell Kx} \quad (H.3)$$

$$\tilde{T}_{\ell} = \sum_{n=0}^{\infty} \alpha_n \sum_{k=0}^n \delta_{k, \frac{1}{2}(n-\ell)} \binom{n}{k} \frac{1}{2^n} \tilde{E}_m^{n-k} \tilde{E}_m^{*k} \quad (H.4)$$

$$\delta_{k, \frac{1}{2}(n-\ell)} = \begin{cases} 1 & , \quad k = \frac{1}{2}(n-\ell) \\ 0 & , \quad \text{otherwise} \end{cases}$$

Since $T_A(E)$ is real, $\tilde{T}_{-\ell} = \tilde{T}_{\ell}^*$.

The first few terms in the T coefficients are easily evaluated from the above formula.

$$\begin{aligned} T_0 &= \alpha_0 + \frac{1}{2} \alpha_2 E_m^2 + \frac{3}{8} \alpha_4 E_m^4 + \dots \\ \tilde{T}_1 &= \left(\frac{1}{2} \alpha_1 + \frac{3}{8} \alpha_3 E_m^2 + \frac{5}{16} \alpha_5 E_m^4 + \dots \right) \tilde{E}_m(x,y) \\ \tilde{T}_2 &= \left(\frac{1}{4} \alpha_2 + \frac{1}{4} \alpha_4 E_m^2 + \frac{15}{64} \alpha_6 E_m^4 + \dots \right) \tilde{E}_m^2(x,y) \end{aligned} \quad (H.5)$$

We are primarily interested in \tilde{T}_1 , the coefficient of the e^{iKx} transmittance term. Assuming that K is much larger than any spatial frequency in the modulating function $\tilde{E}_m(x,y)$, we can separate the first order wave from the other waves. The hologram transmittance may then be defined by:

$$\tilde{T}_H(E) = \{T_A(E)\}_{BL} \quad (H.6)$$

The symbol $\{ \}_{BL}$ is used to denote spatial band limiting of the enclosed function. This band limiting may be accomplished with apertures in the optical system, or in the case of a thick hologram medium may be a property of the grating itself. With this defini-

tion of hologram transmittance we see:

$$\tilde{T}_H(E) = \tilde{T}_1 e^{iKx} \approx \left(\frac{1}{2} \alpha_1 + \frac{3}{8} \alpha_3 E_m^2 \right) \tilde{E}_m e^{iKx} \quad (H.7)$$

If the hologram is re-illuminated with the wave $w_1 e^{-i(K/2)x}$ (Fig. H1), then the transmitted wave will be

$$\tilde{w}_T(x,y) = \left[\alpha_1 + \frac{3}{4} \alpha_3 E_m^2 \right] w_1^2 \tilde{w}_2(x,y) e^{i(K/2)x} \quad (H.8)$$

The first term in the square brackets represents linear wavefront reconstruction, and the second term represents intermodulation products. Only the first terms in the series have been retained, on the assumption that the α_n coefficients decrease rapidly with n . Additional IM products may arise from modulations of the α coefficients if E_0 is not constant. Making $w_1 \gg |\tilde{w}_2|$ assures that these products will be small.

H.3 Effect of Distortion Products on Image Reconstruction

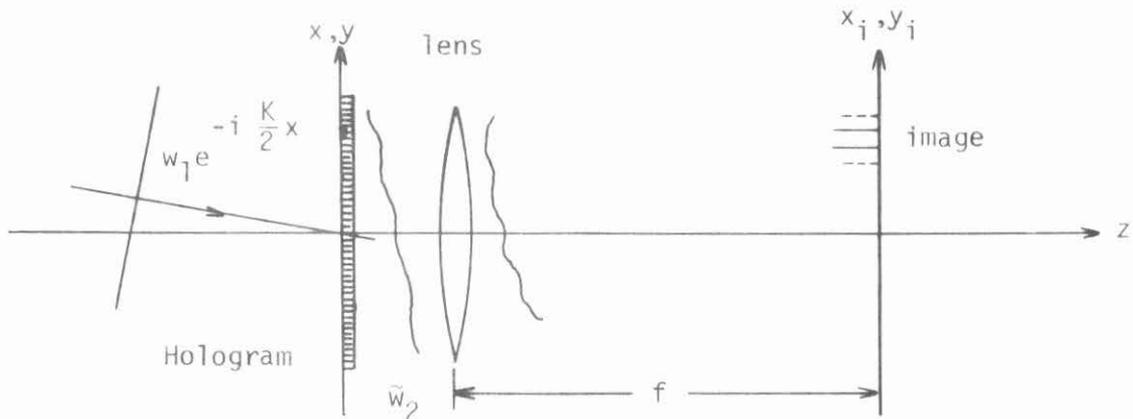


Fig. H3. Image Reconstruction System Showing Ghost Images Due to Intermodulation Products

Consider the system of Fig. H3. An image is projected in the focal plane of a lens placed behind the hologram. If the image is a set of bright spots on a dark background, the effects of IM distortion will be easy to analyze. For example, consider a reconstructed wave with two discrete spatial frequencies,

$$\tilde{w}_2(x,y) = w_2(e^{iK_1x} + e^{iK_2x}) \quad (H.9)$$

$$K_1, K_2 \approx K/2$$

Image spots will occur at

$$x_{1,2} = (\lambda f/2\pi) K_{1,2} \quad (H.10)$$

and intermodulation product spots at

$$(\lambda f/2\pi)(2K_1-K_2), (\lambda f/2\pi)(2K_2-K_1) \quad (H.11)$$

In general, the IM spots will occur at frequencies which are sums and differences of an odd number of image spot frequencies. This explains the absence of α_2 terms in equation H.8. Each of the terms in the hologram transmittance, equation H.7, is on a carrier e^{iKx} . Any second order products will fall in the band at $2K$ or $0K$.

If the image is not a set of distinct spots, but a continuous distribution, the effects of IM distortion are more difficult to analyze. Goodman⁽²⁾ has shown that the effect of nonlinearities on image reconstruction can be represented as multiple self-convolutions of the image amplitude distribution. For a continuous image, IM products will appear as a broad distribution of background light, more concentrated around sharp borders of bright areas. If the bright areas are distinct and widely separated, ghost images may appear at locations corresponding to the interaction of several of the bright areas. Photographs of reconstructed images showing these effects are presented in the papers of Goodman⁽²⁾ and of Friesem⁽³⁾.

A measure of the severity of IM distortion for the bright spot images may be taken as the intensity ratio of IM spots to image spots. From equation H.8 we find this ratio to be:

$$R = \frac{9}{16} \left(\frac{\alpha_3}{\alpha_1} \right)^2 E_m^4 \quad (\text{H.12})$$

4.4 Thin Phase Holograms

The modulated grating hologram may find application as a thin phase, or surface relief hologram. The physical characteristics of this type make it especially suitable for mass manufacture by an embossing technique⁽⁴⁾. The thin phase hologram also finds application in optical memories as a page composer in which information is recorded as surface deformations in a thermoplastic film⁽⁵⁾.

Distortion products in the thin phase hologram arise from two sources. Nonlinearities in the response of the medium to exposure introduce distortions in the recording. The diffraction process itself also exhibits nonlinearities which will cause distortions in the reconstruction. An analysis of these nonlinearities for relief phase holograms recorded in photoresists has been presented by Bartolini⁽¹²⁾.

Assume a recording geometry as in Fig. H1. The surface deformations of the processed hologram will result in phase modulation of an incident wavefront and diffraction of several waves corresponding to the various orders of the diffraction grating (Fig. H4).

The process of making a surface relief pattern from the exposure distribution may involve photoresists, thermoplastics, or whatever, but in general it may be characterized by a phase shift vs. exposure curve

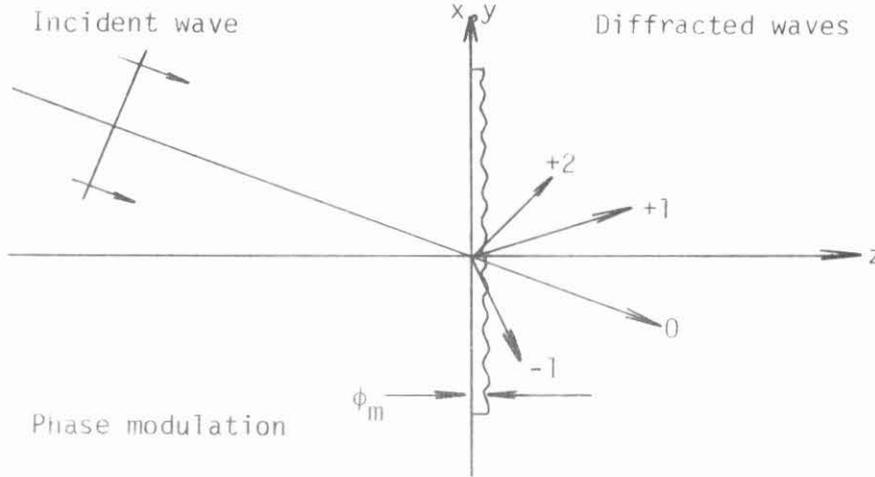


Fig. H4. Wavefront Reconstruction from a Thin Phase Hologram

as in Fig. H5. The phase modulation function is written as a Taylor series,

$$\phi_m(E) = \alpha_0 + \alpha_1 \Delta E + \alpha_2 \Delta E^2 + \alpha_3 \Delta E^3 + \dots$$

$$\Delta E = E - E_0 = \frac{1}{2} \tilde{E}_m e^{iKx} + \frac{1}{2} \tilde{E}_m^* e^{-iKx} \quad (\text{H.13})$$

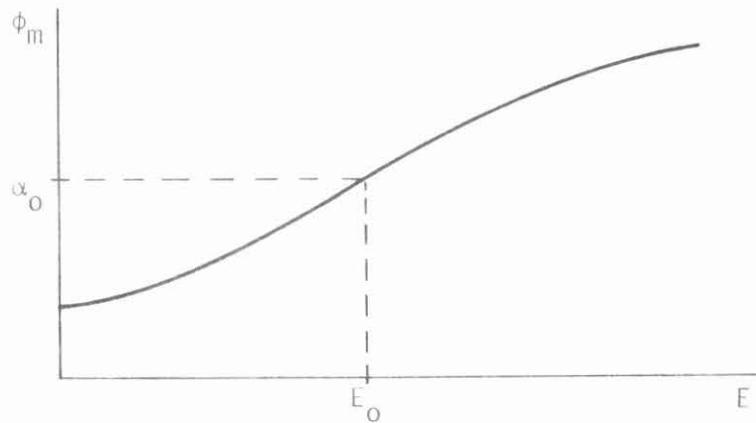


Fig. H5. Phase Shift vs. Exposure for a Thin Phase Hologram

This can be converted to a Fourier series as before:

$$\begin{aligned} \phi_m(E) &= \phi_0 + \frac{1}{2} \tilde{\phi}_1 e^{iKx} + \frac{1}{2} \tilde{\phi}_1^* e^{-iKx} + \frac{1}{2} \tilde{\phi}_2 e^{i2Kx} + \dots \\ \phi_0(x,y) &\approx \alpha_0 + \frac{1}{2} \alpha_2 E_m^2(x,y) \\ \tilde{\phi}_1(x,y) &\approx (\alpha_1 + \frac{3}{4} \alpha_3 E_m^2) \tilde{E}_m(x,y) \end{aligned} \quad (H.14)$$

The complex amplitude transmittance is

$$\tilde{T}_A(E) = e^{i\phi_m(E)} \quad (H.15)$$

The amplitudes of the various diffracted orders may be computed with the formulas

$$\begin{aligned} e^{i\phi \cos \theta} &= \sum_{n=-\infty}^{\infty} i^n J_n(\phi) e^{in\theta} \\ e^{i\phi \sin \theta} &= \sum_{n=-\infty}^{\infty} J_n(\phi) e^{in\theta} \end{aligned} \quad (H.16)$$

which are valid for any complex ϕ ⁽⁶⁾. The Fourier exponential series (equation H.14) is equivalent to a sine-cosine series

$$\begin{aligned} \phi_m(E) &= \phi_0 + \tilde{A}_1 \cos(Kx) + \tilde{B}_1 \sin(Kx) + \tilde{A}_2 \cos(2Kx) + \tilde{B}_2 \sin(2Kx) \\ &\quad + \dots \end{aligned} \quad (H.17)$$

with

$$\begin{aligned} \tilde{A}_1 &= \frac{1}{2} (\tilde{\phi}_1 + \tilde{\phi}_1^*) \\ \tilde{B}_1 &= \frac{i}{2} (\tilde{\phi}_1 - \tilde{\phi}_1^*) \dots \text{etc.} \end{aligned}$$

The complex transmittance is then

$$\begin{aligned}
 T_A(E) &= e^{i\phi_0 + i \sum_{\ell=1}^{\infty} [\tilde{A}_{\ell} \cos(\ell Kx) + \tilde{B}_{\ell} \sin(\ell Kx)]} \\
 &= e^{i\phi_0} \prod_{\ell=1}^{\infty} \left[e^{i\tilde{A}_{\ell} \cos(\ell Kx)} e^{i\tilde{B}_{\ell} \sin(\ell Kx)} \right] \\
 &= e^{i\phi_0} \prod_{\ell=1}^{\infty} \left[\sum_{n, n'} i^n J_n(\tilde{A}_{\ell}) J_{n'}(\tilde{B}_{\ell}) e^{i(n+n')\ell Kx} \right] \quad (H.18)
 \end{aligned}$$

Each term ℓ in the Fourier series expansion of the nonlinearity diffracts a whole set of waves with angular spacing $\ell K/k_a$. Only the $\ell = 1$ term contributes to the first order wave. The hologram transmittance as defined in equation H.6 then becomes

$$\begin{aligned}
 \tilde{T}_H(E) &= \{\tilde{T}_A(E)\}_{BL} \\
 &= e^{i\phi_0} \sum_{n+n'=1} i^n J_n(\tilde{A}_1) J_{n'}(\tilde{B}_1) e^{i(n+n')Kx} \quad (H.19)
 \end{aligned}$$

To explore the behavior of this function we may assume $\tilde{\phi}_1$ real ($\phi = 0$ in equation H.1). This is equivalent to a small shift in the coordinate x and does not alter the physics of the problem. With this assumption

$$\begin{aligned}
 \tilde{A}_1 &= \phi_1, \quad \tilde{B}_1 = 0 \\
 J_{n'}(0) &= \begin{cases} 1 & n' = 0 \\ 0 & n' \neq 0 \end{cases} \\
 \tilde{T}_H(E) &= i e^{i\phi_0} J_1(\phi_1) e^{iKx} \quad (H.20)
 \end{aligned}$$

This J_1 Bessel function behavior is a well known property of the diffracted wave amplitude in a thin phase hologram⁽⁷⁾. A sketch of the function is shown in Fig. H6. The nonlinearity is inherent in the diffraction process and would cause distortions even if the medium were perfectly linear.

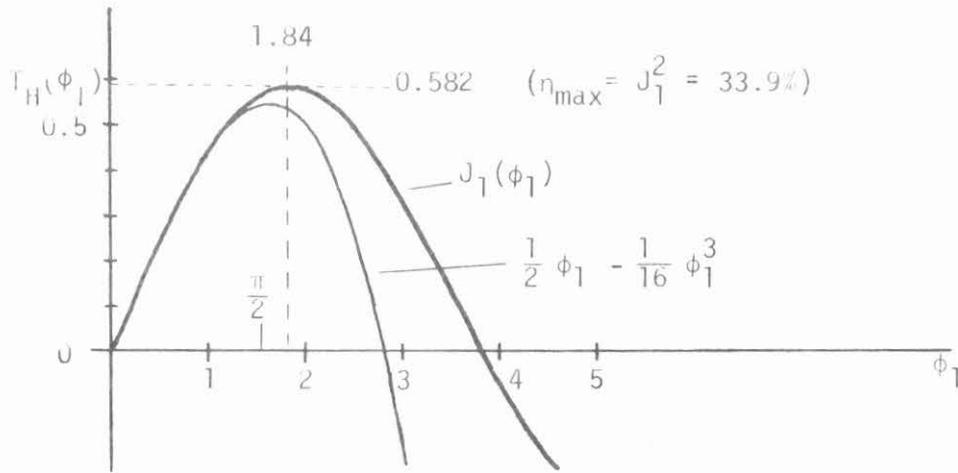


Fig. H6. Amplitude of First Order Diffracted Wave in a Thin Phase Hologram

The Bessel function can be approximated by⁽⁸⁾

$$J_1(\phi) \approx \frac{1}{2} \phi - \frac{1}{16} \phi^3 + \frac{1}{384} \phi^5 + \dots, \quad \phi \ll 1 \quad (\text{H.21})$$

Expanding T_H for small ϕ_1 gives

$$\begin{aligned} T_H(E) &= J_1(\phi_1) \approx \frac{1}{2} \phi_1 - \frac{1}{16} \phi_1^3 \\ &\approx \frac{1}{2} (\alpha_1 + \frac{3}{4} \alpha_3 E_m^2) E_m \\ &- \frac{1}{16} (\alpha_1^3 + \frac{9}{4} \alpha_1^2 \alpha_3 E_m^2 + \dots) E_m^3, \quad \alpha_3 E_m^2 \ll \alpha_1 \end{aligned}$$

$$T_H(E) \approx \underbrace{\left[\frac{1}{2} \alpha_1 \right]}_{\text{linear term}} + \underbrace{\left[\frac{3}{8} \alpha_3 E_m^2 - \frac{1}{16} \alpha_1^3 E_m^2 \right]}_{\text{distortion products}} E_m(x,y) \quad (\text{H.22})$$

The α_3 term in the distortion products is the contribution from the medium nonlinearity (see equation H.7), and the α_1^3 term arises from the diffraction process nonlinearity. The amplitudes of the two terms may add or cancel depending on the signs of the α coefficients.

For a hologram with a few bright spots in the image, the intensity ratio for IM spots to signal spots is (from equation H.22)

$$R = \left[\frac{3}{4} \frac{\alpha_3}{\alpha_1} - \frac{1}{8} \alpha_1^2 \right]^2 E_m^4 \quad (\text{H.23})$$

Comparison of this formula with equation H.12 shows the similarity of the terms due to medium distortion and the added effect of diffraction distortion.

H.5 Thick Phase Holograms

The high diffraction efficiency and angular selectivity of the thick phase hologram make it a good candidate for such applications as spatial filtering and laser machining⁽⁹⁾. In a material such as dichromated gelatin⁽¹⁰⁾ a modulated grating hologram can generate a wavefront with very high diffraction efficiency, negligible higher order waves, and very low noise.

As in the case of the thin phase hologram, distortion arises from nonlinearity of the medium response during exposure and from nonlinearity in the diffraction process during reconstruction. To analyze these

nonlinearities we assume a recording geometry as in Fig. H1 with an exposure distribution given by equation H.1 and the appropriate restrictions on the bandwidth and amplitude of wave \tilde{w}_2 . We assume a medium of thickness T with negligible absorption of the exposing light. The exposure will then be uniform in z . The refractive index in the medium after development will be related to the exposure by

$$n(E) = \alpha_0 + \alpha_1 \Delta E + \alpha_2 \Delta E^2 + \alpha_3 \Delta E^3 + \dots$$

$$\Delta E = E - E_0 = E_m \cos(Kx + \phi) \quad (\text{H.24})$$

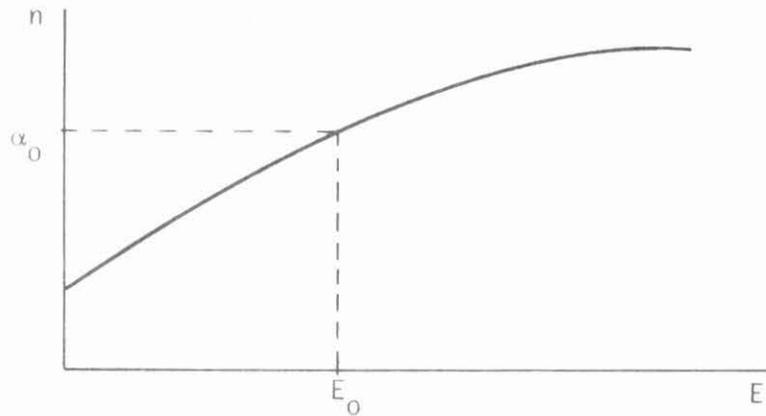


Fig. H7. Relation between Refractive Index and Exposure for a Thick Phase Hologram

The Taylor series is converted to a Fourier series as in the previous cases.

$$n(E) = n_0 + \frac{1}{2} \tilde{n}_1 e^{iKx} + \frac{1}{2} \tilde{n}_1^* e^{-iKx} + \frac{1}{2} \tilde{n}_2 e^{i2Kx} + \frac{1}{2} \tilde{n}_2^* e^{-i2Kx} + \dots \quad (\text{H.25})$$

$$n_0 \approx \alpha_0 + \frac{1}{2} \alpha_2 E_m^2$$

$$\tilde{n}_1(x,y) \approx (\alpha_1 + \frac{3}{4} \alpha_3 E_m^2) \tilde{E}_m(x,y)$$

If $\tilde{w}_2 = w_2 e^{i\phi}$, then $\tilde{n}_1 \approx n_1 e^{i\phi}$ and

$$n(E) \approx n_0 + n_1 \cos(Kx + \phi) \tag{H.26}$$

The wavefront reconstruction process is shown in Fig. H8.

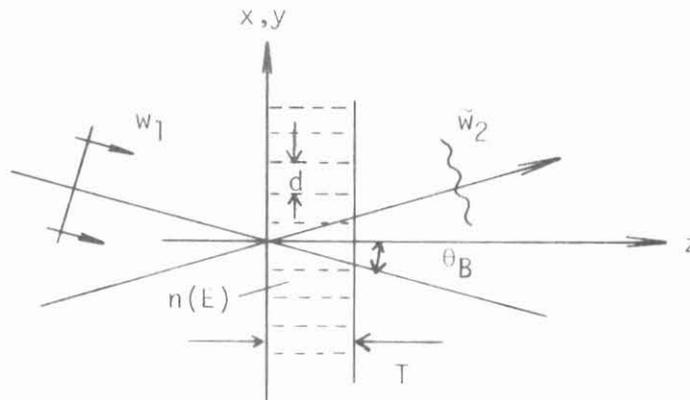


Fig. H8. Wavefront Reconstruction from a Thick Phase Hologram

When re-illuminated with w_1 , this hologram will diffract a wave in the w_2 direction corresponding to the +1 order. This is the only wave which satisfies the Bragg condition, so all other diffracted orders may be ignored. All terms in the $n(E)$ series (equation H.25) except the n_1 term may be ignored for the same reason. No energy is diffracted unless the waves phase match with the grating planes throughout the medium.

The effective hologram transmittance for incident waves at the Bragg angle θ_B (in the medium) is shown from an analysis of coupled wave theory to be⁽¹¹⁾ (see Appendix K)

$$\tilde{T}_H(x,y) = i \sin\left[\frac{\pi T}{\lambda_a \cos \theta_B} n_1(x,y)\right] e^{i(Kx + \phi)} \quad (H.27)$$

This equation was derived using the following assumptions:

- 1) Grating must be thick enough that coupled wave theory is valid. Thickness parameter:

$$Q = 2\pi\lambda_a T/n_0 d^2 \geq 10$$

- 2) $n_1 \ll n_0$
- 3) Index modulation fringes perpendicular to surface, no absorption.
- 4) Incident wave at Bragg angle.

An estimate of the amplitude of the distortion products in equation H.27 can be derived by inserting the series for n_1 into the power series for the sine function and keeping only the lowest order terms in E_m .

$$n_1(E) \approx \alpha_1 E_m + \frac{3}{4} \alpha_3 E_m^3, \quad \alpha_3 E_m^2 \ll \alpha_1 \quad (H.28)$$

$$\begin{aligned} T_H(E) &= \sin\left[\frac{\pi T}{\lambda_a \cos \theta_B} n_1\right] \\ &= \sin(\gamma n_1) \approx \gamma n_1 - \frac{1}{6}(\gamma n_1)^3 \end{aligned}$$

$$\approx \underbrace{\gamma \alpha_1 E_m}_{\text{signal wave}} + \underbrace{\left(\frac{3}{4} \gamma \alpha_3 - \frac{1}{6} \gamma \alpha_1^3 \right) E_m^3}_{\text{distortion products}} \quad (\text{H.29})$$

The intensity ratio of these lowest order IM products to the "signal" wave is

$$R = \underbrace{\left[\frac{3}{4} \frac{\alpha_3}{\alpha_1} - \frac{1}{6} \left(\frac{\pi T}{\lambda_a \cos \theta_B} \right)^2 \alpha_1^2 \right]^2}_{\substack{\text{medium} \\ \text{nonlinearity}} \text{ diffraction nonlinearity}} E_m^4 \quad (\text{H.30})$$

This formula may be compared to equation H.23 for the thin phase hologram and equation H.12 for the thin amplitude hologram. We notice in all three cases the same term due to the medium recording nonlinearity. The diffraction nonlinearity introduces additional distortion which may add to or cancel part of the medium distortion depending on the signs of the α coefficients. Normally $\alpha_1 > 0$ and $\alpha_3 < 0$, so the effects are additive. The absence of E_m^2 terms in the distortion products from the diffraction nonlinearity (equation H.29, H.22) follows from the absence of quadratic terms in the expansions of the J_1 and sine functions.

In the derivation of all these formulas, only the lowest order terms in E_m have been retained in any product or series expansion. We must assume a well behaved Taylor series (α_n rapidly decreasing with larger n) and $\alpha_3 E_m^2 \ll \alpha_1$. If E_m is large, the distortions will be severe and the approximations no longer valid. Also, the bias

exposure will not be constant, and the resulting modulations of the α coefficients may produce several more terms in the distortion products. The approximations will be good if w_1 is a constant amplitude wave with $w_2 \ll w_1$.

References

1. J. W. Goodman, Introduction to Fourier Optics (McGraw-Hill, 1968) p. 230, "Effects of Film Nonlinearities".
2. J. W. Goodman, G. R. Knight, "Effects of Film Nonlinearities on Wavefront Reconstruction Images of Diffuse Objects", J. Opt. Soc. Am. 58, 1276 (1968).
3. A. A. Friesem, J. S. Zelenka, "Effects of Film Nonlinearities in Holography", Appl. Opt. 6, 1755 (1967).
4. R. Bartolini, et al., "Embossed Hologram Motion Pictures for Television Playback", Appl. Opt. 9, 2283 (1970).
5. R. J. Collier, C. B. Burckhardt, L. H. Lin, Optical Holography, Ch. 10.10, "Photoconductor-Thermoplastic Films" (1971).
6. Harold Wayland, Complex Variables Applied in Science and Engineering, Van Nostrand (1970). Formulas 3.26 follow directly from equation 8.77 (p. 220).
7. R. J. Collier, C. B. Burckhardt, L. H. Lin, Optical Holography, p. 225 (Academic Press, 1971).
8. G. N. Watson, A Treatise on the Theory of Bessel Functions, Sect. 2.11, eq. 2 (Cambridge University Press, 1966).
9. A. Engel, J. Steffen, G. Herziger, "Laser Machining with Modulated Zone Plates", Appl. Opt. 13, 269 (1974).
10. R. K. Curran, T. A. Shankoff, "The Mechanism of Hologram Formation in Dichromated Gelatin", Appl. Opt. 9, 1651 (1970).
11. R. J. Collier, C. B. Burckhardt, L. H. Lin, ibid, Chapter 9.6.1.
12. R. A. Bartolini, "Characteristics of Relief Phase Holograms Recorded in Photoresists", Appl. Opt. 13, 129 (1974)

APPENDIX K

Summary of the Results of Kogelnik's Coupled Wave Theory
for a Phase Transmission Hologram

This appendix provides a brief summary of the derivation of the diffraction efficiency for very thick, phase transmission holograms, and provides a bridge between the equations of Kogelnik and equations used in this thesis. A more complete presentation of this topic, which is very complicated algebraically, can be found in Kogelnik's original paper⁽¹⁾ or in several textbooks.^(2,3)

In this discussion we assume a non-absorbing unslanted grating of thickness T , as shown in Figure K1.

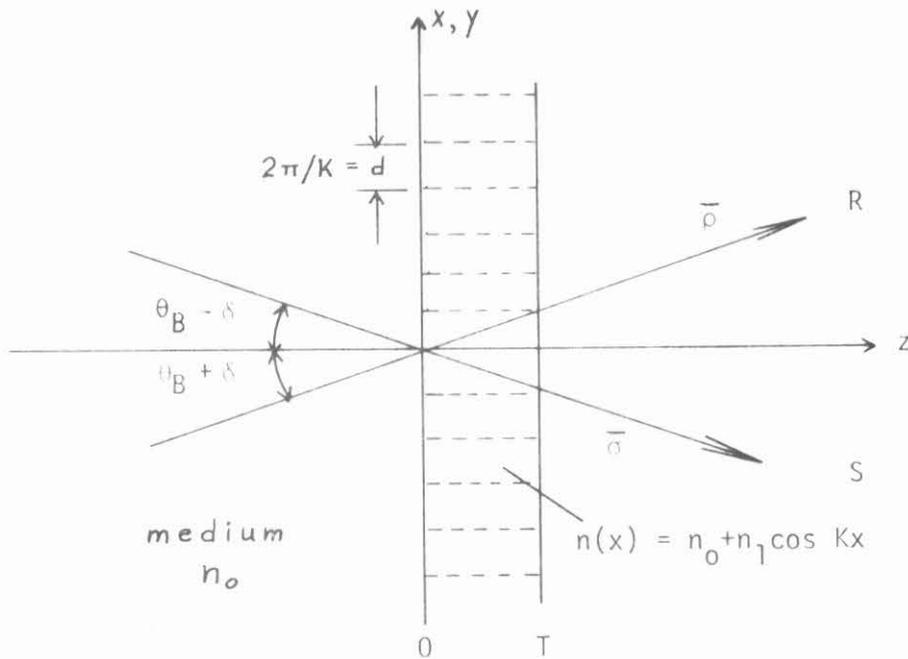


Fig. K1. Interaction of coupled waves in a thick phase hologram

Wave equation

$$(\nabla^2 + k^2) E_y(x, z) = 0 \quad (K.1)$$

$$k^2 = k_0^2 [n_0 + n_1 \cos(Kx)]^2$$

$$k_0 = 2\pi/\lambda_a \quad n_1 \ll n_0$$

$$K = 2\pi/d$$

Assume a solution with only two coupled waves:

$$E_y(\vec{r}) = [R(z) e^{i\vec{\rho} \cdot \vec{r}} + S(z) e^{i\vec{\sigma} \cdot \vec{r}}] e^{-i\omega t} \quad (K.2)$$

$\vec{\rho}$: incident wavevector, near Bragg angle θ_B ($\delta \ll 1$)

$\vec{\sigma} = \vec{\rho} - Kx$ (grating condition is satisfied)

Insert equation K.2 into K.3, set coefficients of $e^{i\vec{\rho} \cdot \vec{r}}$ and $e^{i\vec{\sigma} \cdot \vec{r}}$ separately equal to zero. The second-order differential equations are reduced to first-order by assuming slowly varying solutions and neglecting R'' and S'' terms. The resulting differential equations in R and S can be solved with the boundary conditions $R(0) = 1$, $S(0) = 0$. The amplitude of the diffracted wave from a phase grating of thickness T is shown to be:⁽⁴⁾

$$S(T) = i(C_R/C_S)^{1/2} e^{i\xi} \sin(v^2 + \xi^2)^{1/2} / (1 + \xi^2/v^2)^{1/2} \quad (K.3)$$

$$v = \pi n_1 T / \lambda_a (C_R C_S)^{1/2}$$

$$\xi = \delta K T \sin(\phi - \theta_B) / 2C_S$$

ϕ : grating angle (90° in this case)

C_R, C_S : obliquity factors

For an unslanted grating with an incident wave near the Bragg angle,

the obliquity factors, $C_R \approx C_S \approx \cos \theta_B$, and the above formulas can be simplified. The diffraction efficiency is then:

$$\eta_1 = |S(T)|^2 = \sin^2(v^2 + \xi^2)^{1/2} / (1 + \xi^2/v^2) \quad (K.4)$$

$$v = \pi n_1 T / \lambda_a \cos \theta_B \quad (K.5)$$

$$\xi = \pi \delta / \delta_0 \quad (\delta_0 = d/T) \quad (K.6)$$

$$\theta_B = \arcsin(\lambda_a / 2n_0 d) \quad \text{Bragg angle} \quad (K.7)$$

Summary of assumptions and approximations leading to equation K.4

- 1) Equation K.4 applies to a thick, unslanted phase grating with no absorption.
- 2) Grating must be thick enough that two-wave theory applies.
Thickness parameter $Q = 2\pi\lambda_a T / n_0 d^2 > 10$.
- 3) Index modulation is sinusoidal with $n_1 \ll n_0$.
- 4) Deviation from Bragg angle $\delta \ll 1$.

At the Bragg angle ($\delta = 0$) equation K.4 reduces to

$$\eta_1 = \sin^2[\pi n_1 T / \lambda_a \cos \theta_B] \quad (K.8)$$

The thickness of gratings is most easily found by measuring the angle at which the diffraction efficiency drops to zero. This first null occurs when

$$(v^2 + \xi^2) = \pi^2$$

The cutoff angle is then

$$\delta_c = \delta_0 [1 - (v/\pi)^2]^{1/2} \quad (K.9)$$

The angular bandwidth of the hologram is seen to decrease slightly with

increasing diffraction efficiency. At $\eta = 100\%$, $\nu = \pi/2$, and $\delta_c = .87 (d/T)$.

References

1. H. Kogelnik, "Coupled Wave Theory for Thick Hologram Gratings", Bell Syst. Tech. J. 48, 2909 (1969).
2. R. J. Collier, C. B. Burckhardt, L. H. Lin, Optical Holography, (Academic Press, 1971), Ch. 9, "Diffraction from Volume Holograms".
3. W. T. Cathey, Optical Information Processing and Holography, (Wiley, 1974), Sect. 6.5, "Volume Recorders and Modulators".
4. H. Kogelnik, Ibid., eq. (42), p. 2922.

APPENDIX N

A Simple One-Lens Optical Processor Using Holographic
Spatial Filters

This appendix presents an analysis of the processor discussed in Chapter II. The analysis is based on the Rayleigh-Sommerfeld scalar diffraction integral⁽¹⁾

$$\tilde{a}_2(x_2, y_2) = \frac{1}{2\pi} \iint_{\Sigma} \tilde{a}_1(x_1, y_1) (e^{ikr_{12}/r_{12}})(1/r_{12} - ik) \cdot \cos(\hat{n} \cdot \hat{r}_{12}) dx_1 dy_1 \quad (N.1)$$

$\tilde{a}_1(x_1, y_1)$: field at a point in the aperture Σ

$\tilde{a}_2(x_2, y_2)$: field at a point in a plane to the right of Σ

r_{12} : distance from point 1 to point 2

\hat{n} : unit vector normal to plane 1

$k = 2\pi/\lambda$

With the usual paraxial approximations and Fresnel zone approximations:

$$\cos(\hat{n} \cdot \hat{r}_{12}) \approx 1$$

$$(1/2\pi)(1/r_{12} - ik) \approx -i/\lambda$$

$$(e^{ikr_{12}/r_{12}}) \approx (e^{ikd/d}) e^{ik[(x_1-x_2)^2 + (y_1-y_2)^2]/2d} \quad (N.2)$$

d : distance from plane 1 to plane 2

and ignoring constant phase factors $-ie^{ikd}$, the diffraction formula can be written:

$$\tilde{a}_2(x_2, y_2) = \frac{1}{\lambda d} e^{i\phi_d(x_2, y_2)} \iint_{\Sigma} \{ \tilde{a}_1(x_1, y_1) e^{i\phi_d(x_1, y_1)} \} \cdot e^{-i2\pi(x_1 x_2 + y_1 y_2)/\lambda d} dx_1 dy_1 \quad (N.3)$$

$$\phi_d(x, y) = k(x^2 + y^2)/2d \quad (\text{quadratic phase factor, focal distance } d)$$

Equation N.3 is a form very convenient for computation using the Fast Fourier Transform (Appendix C).

Consider the processor (Fig. N1), with an object transparency, amplitude transmittance $t_o(x_o, y_o)$. The amplitude at the input plane will be:

$$\tilde{a}_1(x_o, y_o) = a_o t_o(x_o, y_o) e^{-i\phi_{do}(x_o, y_o)} \quad (N.4)$$

Application of N.3 yields the amplitude incident on the filter:

$$\tilde{a}_f(x_f, y_f) = (a_o/\lambda d_o) e^{i\phi_{do}(x_f, y_f)} \tilde{T}(\xi, \eta) \quad (N.5)$$

$$\tilde{T}(\xi, \eta) = \mathcal{F}\{t_o(x_o, y_o)\}$$

$$(\xi, \eta) = (x_f/\lambda d_o, y_f/\lambda d_o)$$

$$t_o = 0 \quad \text{outside } \Sigma$$

Filter transmittance:

$$\tilde{h}(x_f, y_f) e^{-iKx_f - ik(x_f^2 + y_f^2)/2f} \quad (N.6)$$

$$f = [1/d_o + 1/d_i]^{-1} \quad (\text{focal length})$$

At the image plane:

$$\tilde{a}_2(x_2, y_2) = \frac{1}{\lambda d_i} e^{i\phi_{di}(x_2, y_2)} \iint \{ \tilde{a}_f(x_f, y_f) \tilde{h} e^{-iKx_f - i\phi_f} e^{i\phi_{di}} \} \cdot e^{-i2\pi(x_f x_2 + y_f y_2)/\lambda d_i} dx_f dy_f$$

ϕ_f is chosen so that the spherical phase factors inside the integral cancel.

$$\phi_{do} - \phi_f + \phi_{di} = k(x_f^2 + y_f^2)(1/d_o - 1/f + 1/d_i)/2 = 0 \quad (N.7)$$

Then:

$$\tilde{a}_2(x_2, y_2) = \frac{a_o}{\lambda^2 d_i d_o} e^{i\phi_{di}} \iint \{ \tilde{T}(\xi, \eta) \tilde{H}(x_f, y_f) e^{-iKx_f} \} \cdot e^{-i2\pi(x_f x_2 + y_f y_2)/\lambda d_i} dx_f dy_f$$

Define:

Magnification factor: $M = d_i/d_o$

New coordinates centered on image:

$$x_i = x_2 + K\lambda d_i/2\pi \quad y_i = y_2 \quad (N.8)$$

Amplitude at the image plane is then:

$$\tilde{a}_i(x_i, y_i) = \frac{a_o}{M} e^{i\phi_{di}} \iint \{ \tilde{T}(\xi, \eta) \tilde{H}(\lambda d_o \xi, \lambda d_o \eta) \} e^{-i2\pi(x_i \xi + y_i \eta)/M} d\xi d\eta$$

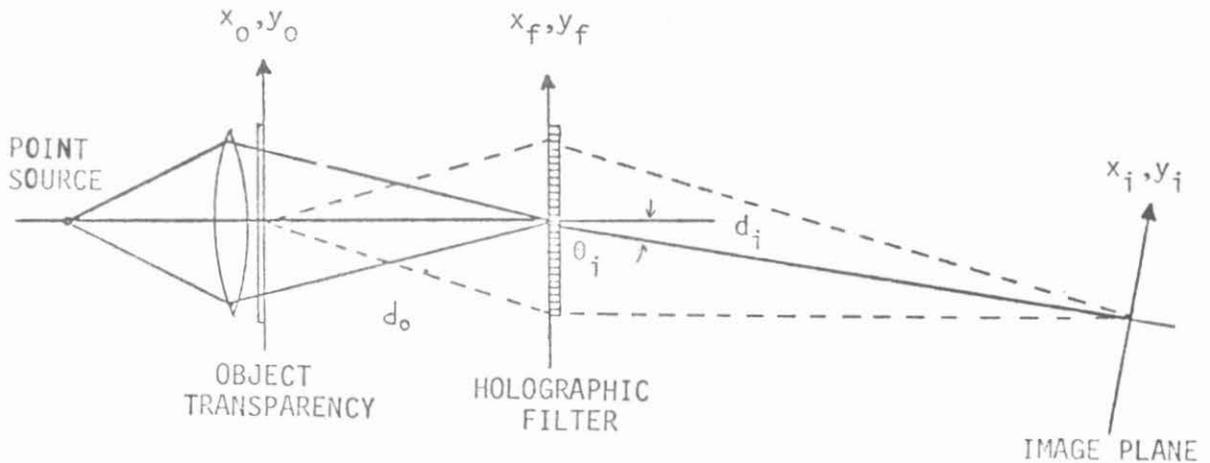


Fig. N1. Coherent optical processor with a holographic spatial filter.

The image field can be written as a convolution:

$$\tilde{a}_i(x_i, y_i) = \frac{a_0}{M} e^{i\phi_{di}(x_2, y_2)} \tilde{h} * t_0\left(-\frac{x_i}{M}, -\frac{y_i}{M}\right) \quad (N.9)$$

$$\mathcal{F}\{\tilde{H}(\lambda d_0 \xi, \lambda d_0 \eta)\} = \tilde{h}\left(\frac{x_i}{M}, \frac{y_i}{M}\right) \quad \text{filter impulse response}$$

$$\mathcal{F}\{T(\xi, \eta)\} = t_0\left(-\frac{x_i}{M}, -\frac{y_i}{M}\right) \quad \text{inverted magnified image}$$

The intensity distribution at the image plane is:

$$I_i(x_i, y_i) = (I_0/M^2) |\tilde{h} * t_0(-x_i/M, -y_i/M)|^2 \quad (N.10)$$

The phase factor $e^{ik(x_2^2 + y_2^2)/2d_i}$ in equation N.9 causes a slight tilt in the image plane. In most simple analyses this plane is assumed to be perpendicular to the z-axis, but the focal plane in any simple imaging system is actually a curved surface. Ray trace calculations can be used to map the location of image surfaces in holographic lens systems. (2,3)

References

1. J. W. Goodman, Introduction to Fourier Optics, (McGraw-Hill, 1968), Ch. 3, "Foundations of Scalar Diffraction Theory".
2. J. N. Latta, "Computer-Based Analysis of Holography Using Ray Tracing", Appl. Opt. 10, 2698 (1971).
3. D. H. Close, A. Au, A. Graube, "Holographic Lens for Pilot's Head-Up Display", Hughes Research Labs, Tech. Report N62269-73-C-0388, Naval Air Development Center (Aug. 1974).

APPENDIX P

Relation of Phase Shift to Density in an Optical Medium

Phase shift of the wave transmitted through a photographic emulsion can be the result of surface height modulations or bulk index variations. The recent synthesis of computer-generated, on-axis holograms on Kodachrome II film⁽¹⁾ has stimulated interest in the underlying mechanism producing the phase shift in such media. An understanding of this mechanism will perhaps enable us to make better holograms through more precise control of the phase of the transmitted wavefront.

This appendix presents the derivation of a formula for that part of the phase shift due to the refractive index variations resulting from dye absorption in the emulsion layer. Numerical calculations using this formula show that for a typical commercial film (Kodachrome II) the expected phase shift is much less than one wavelength.

P.1 Derivation of the Dispersion Integral

The phase shift in an emulsion layer due to refractive index modulation by the dye can be calculated using the Kramers-Kronig relations:

$$A(\omega) = -\frac{1}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{B(\omega') d\omega'}{\omega' - \omega} \quad (\text{P.1})$$

$$B(\omega) = \frac{1}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{A(\omega') d\omega'}{\omega' - \omega} \quad (\text{P.2})$$

where P.V. denotes the principal value of the integral, and $A(\omega) + iB(\omega) = \tilde{H}(\omega)$ an analytic function with no poles in the lower half ω plane. A derivation of these relations is given in the text

by Yariv⁽²⁾. The application of these relations in such diverse fields as electrical network theory⁽³⁾, dispersion of X-rays⁽⁴⁾, and dispersion in optical media⁽⁵⁾ is an indication of the fact that the analyticity of the functions representing these various quantities follows from some very general assumptions about the physical nature of these quantities. In general, a passive, causal, linear system can be represented by a transfer function $\tilde{H}(\omega)$ with no poles in the lower half plane.⁽⁸⁾

If the impulse response of the system $h(t)$ is real, then $\tilde{H}(\omega)$ is hermitian ($A(\omega)$ even and $B(\omega)$ odd function of ω), and the above relations can be put in a form more suitable for computation. Multiply the integrands by $(\omega' + \omega)/(\omega' - \omega)$ and eliminate terms with odd symmetry.

$$A(\omega) = -\frac{2}{\pi} \text{P.V.} \int_0^{\infty} \frac{B(\omega')\omega' d\omega'}{\omega'^2 - \omega^2} \quad (\text{P.3})$$

$$B(\omega) = \frac{2}{\pi} \text{P.V.} \int_0^{\infty} \frac{A(\omega')\omega d\omega'}{\omega'^2 - \omega^2} \quad (\text{P.4})$$

The functions A and B can be related to density and phase shift in an optical medium. The wave equation for a linear, homogeneous medium is

$$(\nabla^2 + \tilde{k}^2) \tilde{a}(\vec{r}) = 0 \quad (\text{P.5})$$

$$\tilde{k}^2 = \omega^2 \mu \epsilon(\omega) - i\omega \mu \sigma(\omega) = \omega^2 \mu \tilde{\epsilon}(\omega)$$

$$e^{i\omega t} \quad \text{time dependence assumed}$$

Assume a solution of the form:

$$\begin{aligned} \tilde{a}(z) &= a_0 e^{-i\tilde{n}(\omega)k_0 z} \\ k_0 &= \omega\sqrt{\mu\epsilon_0} \\ \tilde{n}(\omega) &= n_r(\omega) + in_i(\omega) = (\tilde{\epsilon}(\omega)/\epsilon_0)^{1/2} \end{aligned} \quad (P.6)$$

The transmittance function for an emulsion layer of thickness ℓ may be defined as

$$\begin{aligned} \tilde{T}(\omega) &= \tilde{a}(\ell)/a_0 e^{-ik_0 \ell} = e^{-i(\tilde{n}(\omega) - 1)k_0 \ell} \\ &= e^{-i(n_r(\omega) - 1)k_0 \ell + n_i(\omega)k_0 \ell} \end{aligned} \quad (P.7)$$

The phase delay (in wavelengths) and density are related to $\tilde{n}(\omega)$ by

$$\phi(\omega) = \frac{k_0 \ell}{2\pi} (n_r(\omega) - 1) \quad (P.8)$$

$$D(\omega) = -\log_{10} |\tilde{T}|^2 = -\frac{2}{\ln(10)} k_0 \ell n_i(\omega) \quad (P.9)$$

Let us now consider a very thin layer $\ell = \delta$, $k_0 \delta \ll 1$.

$$\tilde{T}(\omega) \approx 1 - \underbrace{i(\tilde{n}(\omega) - 1)k_0 \delta}_{\tilde{T}_S(\omega)} \quad (P.10)$$

The wave transmitted by the medium is just the superposition of the original incident wave and the wave scattered by the medium. Since $\tilde{T}_S(\omega)$ is the Fourier transform of a real, finite impulse response of a passive system, it satisfies the conditions for application of equation P.4.

$$\begin{aligned}\tilde{T}_s(\omega) &= A(\omega) + iB(\omega) = n_i(\omega)k_0\delta + i(1 - n_r(\omega))k_0\delta \\ &= \left[-\frac{\ell n(10)}{2} D(\omega)\right] + i[-2\pi\phi(\omega)]\end{aligned}\quad (P.11)$$

Substitution of A and B into equation P.4 gives

$$\phi(\omega) = \frac{\ell n(10)}{2\pi^2} \text{P.V.} \int_0^{\infty} \frac{D(\omega')\omega d\omega'}{\omega'^2 - \omega^2}\quad (P.12)$$

which relates the phase delay in wavelengths to the spectral density of the layer. Since both density and phase shift are linearly proportional to thickness ℓ , this result, which was derived for a thin layer, is valid also for layers of arbitrary thickness.

P.2 Phase Shift for a Photographic Emulsion

The phase shift for a monochromatic wave passing through a color photographic film can be written as $\phi(\omega) = \phi_s(\omega) + \phi_c(\omega) + \phi_D(\omega)$ where $\phi_s(\omega)$ is the surface relief phase shift, $\phi_D(\omega)$ arises from the dye absorptions, and $\phi_c(\omega)$ is the remaining phase shift due to other materials which may be present in the emulsion and to variations in film base thickness and index.

The phase shift due to dye absorption may be calculated using equation P.12 if the absorption spectrum of the dye is known. This equation has been evaluated numerically by the method detailed in Section P.3, using spectral density data for the blue absorbing layer of Kodachrome II ⁽⁶⁾. The results are plotted in Figure 2.10. The small x's on the density curve are data points, except for the

point at 300 nm which is a linear extrapolation from the preceding data points. The original Kodak data has a peak density near 1. The data plotted here have been arbitrarily scaled up to represent a more heavily dyed layer.

We can see from this calculation that even with a peak density of 2.5 in the blue, this particular dye will produce a phase shift of only a fraction of a wavelength for red light. The observed phase shifts of several wavelengths must therefore be due to surface relief modulation or index modulation by some other material in the emulsion layer. The extinction of these holograms when immersed in an index matching fluid indicates that the phase shift is almost entirely a surface effect⁽⁷⁾. It would be nice to have a hologram of this type which could be immersed in a liquid gate to eliminate noise and aberrations resulting from surface imperfections in the emulsion and base. The dye used in commercial color films is by no means optimum for this application. A dye with a sharp cutoff in the absorption near the wavelength of interest might produce a much better phase characteristic, since the phase shift is roughly proportional to the slope of the absorption curve at any given wavelength.

P.3 Numerical Evaluation of the Dispersion Integral

We wish to evaluate the following integral:

$$\phi(\beta) = \frac{\ln 10}{2\pi} \text{P.V.} \int_0^{\infty} \frac{D(\beta')\beta \, d\beta'}{\beta'^2 - \beta^2}$$

where ϕ is phase delay (in wavelengths), P.V. denotes the principal value of the integral, D is the measured spectral density curve,

β is a parameter proportional to frequency ω , or inversely proportional to wavelength λ .

The integral is evaluated in two sections as follows:

$$\phi(\beta) = c_1(Q_1 + Q_2) + \text{error}$$

$$c_1 = (\lambda n 10/2\pi^2)$$

$$Q_1 = \int_{\beta_1}^{\beta-\epsilon} \frac{D(\beta')\beta \, d\beta'}{\beta'^2 - \beta^2}$$

$$Q_2 = \int_{\beta+\epsilon}^{\beta_N} \frac{D(\beta')\beta \, d\beta'}{\beta'^2 - \beta^2}$$

$$\begin{aligned} \text{error} &= c_1 \text{ P.V.} \int_{\beta-\epsilon}^{\beta+\epsilon} \frac{D(\beta')\beta \, d\beta'}{\beta'^2 - \beta^2} \\ &\approx c_1 \text{ P.V.} \int_{-\epsilon}^{\epsilon} \frac{D(\beta+x)dx}{2x} \approx \epsilon D'(\beta) \\ &\approx (c_1/2) [D(\beta+\epsilon) - D(\beta-\epsilon)] \end{aligned}$$

β_1 and β_N represent the lower and upper limits of the density curve $D(\beta)$. The size of the excluded region ϵ in the K.K. integral is chosen so as to minimize numerical error. Too small an ϵ will result in error in the evaluation of Q_1 and Q_2 , while too large an ϵ will make the $\epsilon D'$ error large.

The following is a listing of the program and data deck used in generating the plot of Figure 2.10.

```
// SET TIME=1,IOC=10
// EXEC FORTG
//SYSPLTN DD SYSCUT=N
//FORT DD *
C KRAMERS-KRONIG PHASE SHIFT PLOTTING PROGRAM
C PROGRAM PLOTS PHASE-SHIFT VS. WAVELENGTH FOR A PHOTOGRAPHIC EMULSION
C WITH A GIVEN DENSITY VS. WAVELENGTH CURVE. PHASE SHIFT IS CALCULATED
C USING THE KRAMERS-KRONIG INTEGRAL.
C
C DIMENSION X(200),PHI(200),INTVAL(10),WN(20),TITLE(3),W1(6),
C DELTW(6),EPSLON(6)
C EXTERNAL F
C COMMON B,B2,DO,ND,BN(20),DN(20)
C PHASE IN MULTIPLES OF ONE WAVELENGTH
C1 =ALCC(10.)/(2.*3.14159*3.14159)
C READ IN INITIAL DATA
1 READ (5,10) ND,TITLE,DFACTR
WRITE (6,10) ND,TITLE,C1
10 FORMAT (I10,10X,3A4,20X,E15.5)
IF (ND.EQ.0) STCP
C READ IN INTEGRATION RANGE SPECIFICATIONS
READ (5,12) (W1(I),DELTW(I),EPSLON(I),I=1,6)
WRITE (6,12) (W1(I),DELTW(I),EPSLON(I),I=1,6)
12 FORMAT (3(1X,F9.1,F5.1,F10.5))
C DENSITY CURVE, LONGEST WAVELENGTH FIRST, NANOMETERS
READ (5,11) (WN(I),DN(I),I=1,ND)
WRITE (6,11)(WN(I),DN(I),I=1,ND)
11 FORMAT (4(1X,F9.0,1X,F9.3))
DO 90 I = 1,ND
90 DN(I) = DFACTR*DN(I)
C DRAW AXES
CALL VLABEL ( 4.0,5.0,1000.,200.,8.0,8.0,0,0,0,'(F6.0)',5)
CALL VLABEL (3.0,1.0,1100.,200.,9.0,9,'WAVELENGTH (NANOMETERS)',
C 25,0,'(F6.0)',5)
CALL VLABEL (12.5,1.0,0.0,3.0,6.0,6,'DENSITY',-7,1,'(F6.1)',6)
CALL VLABEL (3.5,2.5,-1.0,1.0,5.0,4,'PHASE DELAY (WAVELENGTHS)',
C 25,1,'(F6.1)',6)
C BETA = WAVENUMBER PER MICRON
DO 100 I = 1,ND
100 BN(I) = 1000./WN(I)
BN1 = BN(1)
BNN = BN(ND)
C BEGIN CALCULATION OF PHASE SHIFT CURVE
NSECT = 0
C NEXT SECTION
150 NSECT = NSECT + 1
ERR = 0.0
DW = DELTW(NSECT)
IF (DW.EQ.0) GO TO 800
C EXCLUDED REGION OF KK INTEGRAL
EPS = EPSLON (NSECT)
N = 0
C NEXT POINT
160 N = N + 1
W = W1(NSECT) - (N-1)*DW
IF (W.LT.W1(NSECT+1)) GO TO 700
B = 1000./W
B2 = B*B
DO = 0.0
IF (BN1.LT.B.AND.B.LT.BNN) DO = D(B)
C LOWER HALF INTEGRAL
IF (B-EPS.LE.BN1) GO TO 201
```

```

      BL = AMIN1(B-EPS,PNN)
      Q1 = SIMSON (F,BN1,BL,1.0E-3,ITEST,INTVAL)
      IF (ITEST.EQ.0) GO TO 202
      WRITE (6,23) N,W,INTVAL
23  FORMAT (22H NUMERICAL ERROR, STEP,14,18H Q1  WAVELENGTH =,
      C   F7.1,14H  INTERVAL = ,10I2)
      GO TO 202
201 Q1 = 0.0
202 CONTINUE
C   UPPER HALF INTEGRAL
      IF (B+EPS.GE.BNN) GO TO 301
      BL = AMAX1 (B+EPS,BN1)
      Q2 = SIMSON (F,BL,BNN,1.0E-3,ITEST,INTVAL)
      IF (ITEST.EQ.0) GO TO 302
      WRITE (6,24) N,W,INTVAL
24  FORMAT (22H NUMERICAL ERROR, STEP,14,18H Q2  WAVELENGTH =,
      C   F7.1,14H  INTERVAL = ,10I2)
      GO TO 302
301 Q2 = 0.0
302 CONTINUE
C   ERROR ESTIMATE
      ERR1 = ABS ( D(B+EPS)-D(B-EPS) )
      ERR = AMAX1(ERR1,ERR)
      X(N) = W
500 PHI(N) = C1*(Q1+Q2)
      GO TO 160
700 NW = N-1
      ERR = C1*ERR/2.0
      WRITE (6,22) (X(I),PHI(I),I = 1,NW)
22  FORMAT (/5(1X,F10.1,F10.3))
      WRITE (6,21) ERR
21  FORMAT (1H0,24H MAXIMUM ERROR ESTIMATE ,E10.3/1H1)
C   PLOT PHASE SHIFT CURVE
      CALL PLOTXY (Nh,X,PHI,1400.,-100.,-2.0,+2.0,0,1,0,0,TITLE)
      GO TO 150
C   PLOT DENSITY CURVE
800 CALL PLCTXY (ND,WN,DN,1400.,-100.,-0.5,4.5,1,2,4,1,TITLE)
      GO TO 1
      END
      FUNCTION F(BX)
C   INTEGRAND OF KK INTECRAL
      COMMON B,B2,DO
      F = D(BX)*B / (BX*BX-B2)
      RETURN
      END
      FUNCTION D(X)
C   LINEAR INTERPOLATION IN DENSITY CURVE
      COMMON B,B2,DO,ND,BN(20),CN(20)
      D = YNTERP (BN,DN,X,ND,1)
      RETURN
      END
//DATA  DD      *
      18
      1000.  10.   K.K.PLOT
      200.   0.0   0.01  550.   2.0   0.005  300.  10.0  0.01
      580.   .010  550.   .022  533.   .042  523.   .068
      516.   .098  508.   .150  496.   .250  480.   .462
      467.   .670  458.   .765  452.   .822  448.   .832
      444.   .846  440.   .848  434.   .846  424.   .805
      412.   .735  300.   0.00
      END
//

```

Symbol Definitions

CI:	Numerical constant
ND:	Number of points in density curve
TITLE:	Alphanumeric to appear in upper right corner of plot
DFACTR:	Scale factor for density curve

Plot may be broken up into as many as six sections with different integration parameters for each section.

W1(I):	Wavelength at start of I th section
DELTW(I):	Plotting increment for I th section
EPSLON(I):	Size of excluded region in K.K. integral
WN:	Array of wavelengths for density curve
DN:	Measured densities
W:	Wavelength (nanometers)
B:	Frequency parameter β
BX:	Integration variable β'
Q1:	Lower half integral
Q2:	Upper half integral
X:	Plotted wavelength array
PHI:	Plotted phase shift array

Subroutines^{*}

VLABEL: Draws axes and labels

SIMSON: Numerical integration of function F over specified
 interval using Simpson's Rule

PLOTXY: General two-dimensional plotting routine

YINTERP: Performs linear and higher order interpolation in a
 given data array

* Documentation available at Caltech computing center.

References

1. D. C. Chu, J. R. Fienup, J. W. Goodman, "Multiemulsion On-Axis Computer Generated Hologram", Applied Optics 12, 1386 (1973).
2. A. Yariv, Quantum Electronics (John Wiley, 1967); Appendix 1: "The Kramers-Kronig Relations".
3. H. W. Bode, Network Analysis and Feedback Amplifier Design (Van Nostrand, 1945), Ch. 14: "Relation between Real and Imaginary Components of Network Functions".
4. R. L. Kronig, "On the Theory of Dispersion of X-rays", J. Opt. Soc. Amer. 12, 547 (1926).
5. R. W. Ditchburn, Light, 2nd ed. (Blackie & Son, Glasgow, 1963), Ch. 15: "The Electromagnetic Theory of Absorption and Dispersion", Appendix 19B: "Kramers-Kronig Dispersion Relations".
6. Kodachrome II Film (Daylight)/(5029), Process K-12, Spectral Dye Density, Eastman Kodak Co., 7/26/72.
7. D. C. Chu, personal communication.
8. L. D. Landau, E. M. Lifshitz, Electrodynamics of Continuous Media, (Pergamon, 1960), Sect. 62, "The Relation between the Real and Imaginary Parts of $\epsilon(\omega)$ ".