THEORY OF HOLOGRAPHY

Thesis by

John Wallace Matthews

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1967

(Submitted May 17, 1967)

ACKNOWLEDGMENTS

The author would like to express his sincere appreciation to Professor Nicholas George for his guidance and encouragement during the course of this research.

The author is greatly indebted to his wife, Jonel, for her constant help and encouragement during the course of this research, and for the effort spent in assisting with the preparation of the manuscript.

The author wishes to thank Mrs. Arlene Wyszomirski for typing portions of the manuscript.

In addition, the author wishes to thank Mrs. Martha Lamson for assistance with the computer programming, and Stan Kistler for his assistance in the processing if the film plates, as well as all the people who have been of assistance in one way or another during the course of this research.

Finally, the financial support of the National Science Foundation, through its fellowship program, is gratefully acknowledged, as well as the indirect financial support of the Electronics Division of the Air Force Office of Scientific Research.

ABSTRACT

Computation of the amplitudes of the diffracted fields which are produced when a reflection hologram or a "thick" transmission hologram is illuminated requires that the 3-dimensional nature of the hologram be accounted for. A general analytical method is formulated for computing the diffracted fields in terms of the initial exposing field, the film characteristics, and the illumination field, taking into account the entire emulsion volume. This method, which is applicable to both transmission and reflection holograms, involves characterizing the emulsion volume by the volume density of scattering particles, with the diffracted field being found by coherently summing the scattered waves, neglecting multiple scattering. The initial exposing field and the illumination field are expressed in the form of a sum of plane or quasi-plane waves, and the diffracted field is expressed as a sum of waves, each of which is found by solving a variation of the same basic problem. This problem consists of computing the directions, amplitudes, and phases of the first-order diffracted waves produced when a 3-dimensional array of scattering particles having a sinusoidal density distribution is illuminated by a plane wave. The solution of this problem is considered, with the directions and phases of the diffracted fields being computed for both transmission and reflection holograms. The amplitudes are computed for the case of transmission holograms and the analytical expressions are evaluated numerically for a number of particular cases to determine the effect of varying different para-

meters on the amplitudes of the diffracted waves. The results are compared with experimental data obtained by making a careful study of different holographic diffraction gratings.

The results of the analytical method described above are compared with the results of the method whereby the hologram is characterized by the transmittance, and it is shown that with respect to the computation of the directions and phases of the diffracted waves, the two methods are equivalent for the case of transmission holograms.

The case where the reference beam is composed of a series of waves (ghost imaging) is considered using both of the above methods, and the translational sensitivity and background noise which arise in this case are investigated. An experiment dealing with translational sensitivity was carried out and the experimental results were found to be in good agreement with the theory.

The duplication of holograms is considered and the duplication process is described in terms of making a hologram of a hologram, rather than in terms of making a contact print. Experimental results are presented to support this point of view and the effects of varying the characteristics of the illumination wave are described. The duplication of both transmission and reflection holograms is dealt with and a simple method for duplicating reflection holograms is proposed and discussed.

TABLE OF CONTENTS

INTRODUCTION

Holography, or wavefront reconstruction, involves the recording of an interference pattern which is generated by the coherent superposition of two or more waves. The basic ideas relating to the holographic process were discovered by Gabor (1) in 1948, who sought to utilize holography to increase the resolution attainable with the electron microscope. A considerable amount of work was done during the 1950's to apply wavefront reconstruction techniques to microscopy, but experimental work was difficult due to the lack of an intense source of coherent radiation in the short wavelength portion of the spectrum*.

The discovery of the gas laser eliminated this difficulty and made practical the use of new experimental techniques. Leith and Upatnieks described the holographic process from the viewpoint of communication theory (2) and went on to demonstrate experimental techniques (3,4) which made practical the application of holography to a wide variety of problems. Before discussing those aspects of holography which will be of interest here, it may be useful to give a brief description of the holographic process.

The holographic process can be described in general terms as a two-step process. In the first step an electromagnetic field**

---

* See reference (68) for a complete bibliography for the period 1948 to 1965.

** Sound waves have also been used (63).

interacts with a sensitive recording device (such as a photographic emulsion layer) and changes its characteristics in some manner. In the second step the recording device (which may have undergone additional processing such as development in the case of film) is illuminated with another electromagnetic field in order to produce the "diffracted" or "reconstructed" field.

The general description of the holographic process given above is not specific enough to point out those aspects which make it uniquely different from other processes which fit this general description, such as ordinary photography. The fundamental idea which is the basis of holography is the utilization of the fact that the intensity distribution in an interference pattern which is generated by two or more coherent electromagnetic waves is a function of the phases of the waves. This idea is utilized in the recording step, where an additional field (called the reference beam) is combined with the "signal" field which is to be recorded on the film plate. The resulting interference pattern which is recorded in effect contains both amplitude and phase information of the signal and reference beams. Thus, by the use of the phenomena of interference it is possible to encode both amplitude and phase information with a recording device such as film which is sensitive only to the intensity of the exposing field.

Discussion and Summary of Text

In the following we shall briefly summarize and discuss the various problems that have been dealt with in this thesis, without attempting to trace the development of similar or related work done

by other researchers, as this is done in some detail within the body of the thesis itself.

In Chapter One we formulate a general analytical method for computing the diffracted field in terms of the initial exposing field, the film characteristics, and the illumination field. The emulsion volume is characterized by the grain density, which specifies the volume density of scattering particles, and the diffracted field is computed by summing the waves scattered by the grains, neglecting multiple scattering. The neglect of the multiple scattering allows us to treat the problem of computing the diffracted field from what might be termed a linear systems approach. We express both the exposing field and the illumination field as a sum of plane or quasi-plane waves and obtain the diffracted field in terms of a sum of plane or quasi-plane waves, whose amplitudes, directions, and phases are computed by solving variations of the same basic problem, that of computing the diffracted fields produced when a three-dimensional array of scattering particles having a sinusoidal density variation is illuminated by a plan wave.

We go on in Chapter One to compute the directions and phases of these diffracted waves, both for transmission and reflection holograms, and demonstrate that when the illumination wave is the reference beam, the signal beam is reconstructed. A comparison is then made of transmission and reflection holograms.

In Chapter Two we compare the results of the analysis of Chapter One with the analysis formulated by Gabor (5), which was

later put in communication theory language by Leith and Upatnieks (2), and show that with respect to computing the directions and phases of the diffracted waves, the two approaches are equivalent.

In Chapter Three we consider the case where the reference beam consists of more than one simple wave, and examine how this affects the reconstruction of the signal beam. If the reference beam is quite complicated then the illumination of the developed hologram plate is accomplished by repositioning the hologram in the experimental setup. The sensitivity of the reconstruction of the signal beam to repositioning errors is described in terms of the analysis developed in Chapter One, and computed for a specific experiment. This experiment was carried out and the experimental results were found to be in good agreement with the theory.

In addition to being sensitive to repositioning errors, holograms with multiple wave reference beams yield a reconstruction of the signal beam that is accompanied by a background noise. This background noise is investigated and a signal to noise ratio is defined and computed.

Of particular interest is the case where we are dealing with Fourier transform holograms with multiple wave reference beams, as they find use in the area of character recognition and complex spatial filtering. We examine translational sensitivity and background noise for this case when we have plane wave and diffuse illumination.

In Chapter Four we extend the analysis of Chapter One to include the computations of the amplitudes of the diffracted waves.

Within the framework of that analysis, it suffices to solve the problem of computing the amplitudes of the two first-order diffracted waves produced when an arbitrary sinusoidal 3-dimensional array of scattering particles within the emulsion layer is illuminated by a plane wave. This is done for an arbitrary illumination plane wave, taking into account attenuation within the emulsion layer and reflection losses at the interfaces. Here, as in Chapter One, multiple scattering is neglected. The analytical expressions are evaluated numerically for certain special cases and these results are compared with experimental data derived from a series of experiments with holographic diffraction gratings. The agreement between experiment and theory was found to be satisfactory.

In Chapter Five the duplication of holograms is described in terms of taking a hologram of a hologram, rather then as making a contact print. The duplication of thick transmission hologram is studied, and the effects of varying the characteristics of the illumination wave are described. Experiments dealing with the duplication of holographic diffraction gratings are described and the experimental results support the point of view taken here. A simple method for duplication of reflection holograms is described and the various factors affecting the production of duplicate reflection are discussed.

In the Sixth and final chapter the results of this thesis are summarized and discussed.

CHAPTER ONE

SCATTERING THEORY OF VOLUME HOLOGRAMS

1.1 Introduction

Gabor, in his classic paper (5) describes the holographic process from the point of view which we shall term as the transmittance approach. Basically, the exposing field is taken to be a complex scalar quantity of the form

$$U = U_o + U_1 = Ae^{i\psi} \quad , \tag{1.1}$$

where U is assumed to be specified in the plane of the film emulsion layer. Variations with depth within the emulsion layer are neglected and the response of the film is characterized by a parameter Γ (the "gamma" of the film). After processing, the developed film plate or "hologram" is assumed to have an amplitude transmittance τ which is proportional to $[UU^*]^\Gamma$, that is

$$\tau = C[UU^*]^\Gamma \quad . \tag{1.2}$$

When the hologram plate is illuminated by another field U_2, the transmitted field U_T is assumed to be given by

$$U_T = C'U_2\tau \quad . \tag{1.3}$$

Equations 1.2 and 1.3 are the fundamental relations used by Gabor and by the majority of investigators engaging in research in the various aspects of holography. The variations in analysis are primarily

concerned with different approaches in the computation of the exposing field U and in the computation of the transmitted or "diffracted" field at an observation point some distance from the film plate. Leith and Upatnieks utilized communication theory to put these calculations in transform language (2), and their approach has been widely used with excellent results.

The neglect of the variations with depth within the emulsion layer (i.e. considering the hologram to be two dimensional) is quite justifiable when the spatial frequencies recorded on the film plate are low, as was usually the case in the early work done by Gabor, Rodgers, and others, prior to the invention of the laser. The invention of the gas laser with its high brightness and relatively long coherence length made practical the use of high spatial frequencies in the experimental configurations first developed and demonstrated by Leith and Upatnieks (2,3,4). In these configurations the reference beam is brought in at some offset angle, with higher spatial frequencies corresponding to larger values of the angle. It was pointed out by Friesem (6) that as the spacing between fringes becomes comparable with the emulsion thickness the film plate can no longer be regarded as two dimensional. In this case the use of the average transmittance becomes questionable as the variation of the transmittance with depth should be accounted for. This does not mean, however, that the transmittance approach as given by equations 1.2 and 1.3 is no longer useful when the spatial frequencies are high. It will be shown in Chapter 2 that with respect to computing the direction and phase of the diffracted waves (produced

when the hologram is illuminated) the transmittance approach yields results which are in agreement with the more general theory which will be presented in this chapter. Where the transmittance approach breaks down is in predicting the amplitudes of the diffracted waves. For example, it has been noted by a number of researchers (7,8,9) that for holograms having high spatial frequencies the brightness of the reconstruction is highly dependent on the angle of illumination.

The first attempt at accounting for the finite thickness of the emulsion layer appears to have been done by Denisyuk (10,11), who was primarily concerned with the "reflection hologram." In this case the interference planes (planes of maximum intensity in the interference pattern generated by the signal and reference beams) are nearly parallel with the emulsion surface and the emulsion layer is considered as thick. P. J. van Heerden, in his investigations of optical information storage in solids (12,13), takes full account of the three-dimensional nature of the recording of interference patterns within a film emulsion layer. He outlines an approach whereby the exposing field is treated as a sum of plane waves, and considers in some detail the recording of the interference pattern of two plane waves on film.

Leith and co-workers (8) have investigated holographic data storage in three-dimensional media, and have analyzed the effect of emulsion thickness on the optical characteristics of the reconstructed images. They treat in detail the case of a hologram of two plane waves (a hologram diffraction grating) noting that the complex spatial distribution of the light from an object can be decomposed into a

spectrum of plane waves.

Offner, in a recent paper (14), considers the special case of a hologram of two point sources and treats the hologram as a diffraction grating whose parameters are a function of position. He then uses generalized grating equations and ray tracing techniques to compute the direction of the diffracted wave in order to examine the reconstruction process. Offner notes that these techniques can be extended to more general holograms.

In this chapter the holographic process is described from a point of view which is closely related to the ideas of van Heerden (13) and Offner (14). The exposing field in the vicinity of the film plate is written as a sum of well-defined wavefronts. The developed film emulsion layer is characterized by the grain density D, which is related to the total electric field $\bar{E}$ within the emulsion layer by a power series in $\bar{E}\cdot\bar{E}^*$. The various interference terms appearing in the resulting expression for D, which yield the real and virtual images, are identified, as are those terms which yield ghost images and higher order images.

The assumption is made that each film grain acts as an independent scatterer when the hologram is illuminated (i.e., we neglect multiple scattering). It is shown that this implies that with respect to computing the direction and phase of the scattered waves the various periodicities or "grating terms" in the expression for D can be considered separately, with the total diffracted field being the linear sum of the waves diffracted by each "grating." The compu-

tation of the amplitudes of the diffracted waves is considered in Chapter Four. In this chapter, general expressions for computing the direction and phase of each wave diffracted by a typical periodicity of D are derived for arbitrary plane wave illumination, in terms of the initial exposing waves which generated that periodicity. These expressions are then used to predict a reconstruction of the original wavefronts (virtual image) when the illumination beam is the reference beam.

The analysis is then extended to the case of the reflection hologram. The reconstruction is described in terms of Bragg reflection from the interference planes within the emulsion layer, as given in the expression for the grain density D. The reconstruction of the original wavefronts (virtual image) is then analyzed and finally, a comparison is made of the reflection and transmission holograms.

1.2 Recording Process

In this section we shall consider the specification of the relevant characteristics of the developed film emulsion layer in terms of the exposing field. The exposing field in the vicinity of the film plate is written in a general form, with no consideration being given to the problem of relating the field at the film plate to the sources which generated the field.

1.2.1 Exposing Field

The field which exists in the region of the film plate during exposure of the hologram is taken to be of the form of a sum of well-

defined, regular waves. Plane waves or spherical waves would be examples of such waves, but in general any wave whose wavefronts are sufficiently smooth, such that they can be approximated (locally) by a plane wave, satisfies the conditions being imposed here. It will be assumed that the source used is monochromatic and that the various waves add coherently in the region of space occupied by the film plate. Thus, using complex notation, the field in the region where the film plate is to be placed is written in the form

$$\bar{E} = \sum_n \bar{E}_n \; e^{i(\bar{k}_n \cdot \bar{r} + \phi_n)} \; e^{-i\omega t} \quad . \tag{1.4}$$

In the above expression, $\bar{E}_n$ is a real vector, ϕ_n is a real constant, and the propagation vector $\bar{k}_n$ is given by

$$\bar{k}_n = \frac{2\pi}{\lambda_o} \bar{e}_n \quad . \tag{1.5}$$

Both $\bar{E}_n$ and $\bar{e}_n$ may be functions of $\bar{r}$, but are assumed to be sufficiently slowly varying such that the wavefronts may be considered as (locally) planes*.

In most cases of practical interest (ghost imaging being an exception) one of the well-defined wavefronts in the sum in equation 1.4 has a greater amplitude than the rest and has a direction of propagation significantly different than those of the other terms in the sum. This wave is commonly referred to as the reference beam.

*See section 1.3.2 for a more complete discussion of this point.

It will be convenient to write the term corresponding to the reference beam separately, thus

$$\bar{E} = \bar{E}_o \, e^{i\bar{k}_o \cdot \bar{r}} \, e^{-i\omega t} + \sum_{n=1}^{M} E_n \, e^{i(\bar{k}_n \cdot \bar{r} + \phi_n)} \, e^{-i\omega t} \quad . \qquad (1.6)$$

1.2.2 Recording Media

The characteristics of the developed film plate, or hologram, which was exposed to the field given by equation 1.6, depends on the nature of the film and development process, and upon the field within the emulsion layer during exposure of the plate. The field within the emulsion layer will be of the same form as equation 1.6 but the propagation vectors will have a different direction (and magnitude) due to refraction at the emulsion-air interface. In addition, the vectors $\bar{E}_n$ will be different due to reflection losses at the interface. In the analysis presented here any attenuation or scattering that may take place within the emulsion layer during exposure will be neglected. No notational changes will be made to differentiate between the fields within and without the emulsion, the meaning being clear from the text.

The characteristic of the developed film emulsion which is of interest here is D, the volume density of grains within the emulsion layer. In the case of photographic emulsions utilizing silver halides within a gelatin matrix, the grains referred to are small metallic silver particles of rather complex shape. If the emulsion is bleached

during processing, then these metallic silver particles are replaced by a transparent silver compound having an index of refraction different than that of the surrounding gelatin matrix. In this case the grains referred to would be these transparent silver compounds. The basic property of these grains which is of interest is that they possess either an index of refraction or conductivity which is different from that of the uniform gelatin matrix of the emulsion layer, and thus they act as scatterers of radiation when the hologram is illuminated. The details of the chemical processes which take place within the emulsion layer are discussed at length in several excellent references on photographic chemistry (15,16).

Photographic film is sensitive to the total electric field, rather than to the power density or magnetic field. This is a significant factor when the exposing field is of the form of an interference pattern generated by the coherent superposition of two or more waves, and is discussed more fully in Appendix I.

The grain density D is expressed in terms of the square of the magnitude of the total electric field that existed within the emulsion volume during exposure of the film plate. That is

$$D = C_o + C_1|E|^2 + C_2|E|^4 + \text{---} , \tag{1.7}$$

where C_o, C_1, C_2, --- are constants which depend on exposure time, film characteristics, processing procedures, and the wavelength of the monochromatic exposing field. The quantity $|E|^2$ is the square

of the magnitude of the exposing electric field and is given by

$$|\bar{E}|^2 = \bar{E}.\bar{E}^* \tag{1.8}$$

where $\bar{E}$ is given by equation 1.4. It is a straightforward computation to express the grain density D in terms of the initial exposing field given by equation 1.6 using equations 1.7 and 1.8. These computations are carried out in detail in Appendix II and the results are given below. It is found that D can be written in the form

$$\begin{aligned} D = {} & C_o + C_1 E_o^2 + C_2 E_o^4 + (2C_1 E_o + 4C_2 E_o^3) \sum_n b_n \cos[(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - \phi_n] \\ & + (C_1 + 2C_2 E_o^2) \sum_{n,m} C_{nm} \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m] \\ & + 2C_2 E_o^2 \sum_{n,m} b_n b_m \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m] \\ & + 2C_2 E_o^2 \sum_n b_n^2 \cos[2(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - 2\phi_n] \\ & + 2C_2 E_o^2 \sum_{\substack{n,m \\ n\neq m}} b_n b_m \cos[(2\bar{k}_o - \bar{k}_n - \bar{k}_m)\cdot\bar{r} - \phi_n - \phi_m] \\ & + 4C_2 E_o \sum_{n,p,q} b_n C_{pq} \cos[(\bar{k}_o - \bar{k}_n + \bar{k}_p - \bar{k}_q)\cdot\bar{r} - \phi_n + \phi_p - \phi_q] \\ & + C_2 \sum_{n,m,p,q} C_{nm} C_{pq} e^{i(\bar{k}_n - \bar{k}_m + \bar{k}_p - \bar{k}_q)\cdot\bar{r}} e^{i(\phi_n - \phi_m + \phi_p - \phi_q)} \\ & + \text{higher order terms} . \end{aligned} \tag{1.9}$$

The quantities b_n and C_{nm} are defined by

$$b_n = \frac{1}{|\bar{E}|_o} \bar{E}_o \cdot \bar{E}_n \tag{1.9a}$$

and

$$C_{nm} = \bar{E}_n \cdot \bar{E}_m \quad . \tag{1.9b}$$

In addition to the presence of "grains" distributed throughout the sensitive portion of the volume of the recording media, there may also occur a variation in thickness of the recording media. This has been observed by Rigler (17) for Kodak type 649-f and type So-243 film plates. Rigler reported that reconstructions in reflection could be obtained from such film plates when they were coated with aluminum. Altman (18) discusses relief images on type 649-f plates in more detail and mentions ways in which such images may be enhanced. Urbach and Meier (19) have produced holograms using a "grainless" recording media, where a "phase image" is produced by electrostatically induced deformations of a dielectric surface. It is thus clear that in certain cases the variations in thickness of the recording media may be the significant factor. In other cases, however, such as when photochromic glass (20) is used as the recording media, the variations in thickness will be negligible or non-existent. In the analysis presented here any variations in the thickness of the recording media will be neglected. This is clearly a good approximation for the case where photochromic glass is the recording medium and appears to be a reasonable approxi-

mation for type 649-f film (17).

1.3 Reconstruction with Transmission Type Holograms

In the reconstruction process the developed hologram plate is illuminated with some form of electromagnetic wave, usually the reference beam or some wave of similar characteristics. The problem which will be considered now will be the computation of those properties of the diffracted field which can be computed from a knowledge of the illumination wave and of the hologram film plate as described by the quantity D given by equation 1.9. It is noted that D, the volume density of metallic silver grains within the gelatin matrix of the film emulsion, is not a complete description of the diffracting structure, the film emulsion layer. No attempt has been made to describe the characteristics of the individual grains, such as their size, shape, orientation, etc. Thus it should be expected that only certain characteristics of the diffracted field can be computed from a knowledge of D alone.

Examination of the expression for D shows that the film emulsion layer is a periodic structure, and that these periodicities are specified in the expression for D given by equation 1.9. A knowledge of these periodicities allows the computation of the directions of the various diffracted waves, as well as their phases. A knowledge of the periodicities alone, however, may not be sufficient to provide for the computation of amplitudes, and thus the distribution of power among the diffracted waves. The situation is analogous to the problem of computing the radiation pattern of

an antenna array, where the location of each antenna is known but the individual characteristics of each antenna are unknown. The array factor can be computed from the periodicity but the radiation pattern of the individual antennas remains unknown. In the case under consideration here, we find that a great deal of information can be obtained regarding the amplitudes of the diffracted waves from a knowledge of the periodicities alone. This is discussed further in Chapter Four.

1.3.1 Linearity Hypothesis - Neglect of Multiple Scattering

We shall assume that the field scattered by each grain (metallic silver grain for unbleached gratings, dielectric grain for bleached gratings) is essentially independent of the presence of the other grains within the gelatin matrix of the film emulsion layer (i.e., we neglect multiple scattering). That is, we assume that the field scattered by any given grain is dependent primarily on the illumination field and not to any great extent on the field scattered by the other grains. This would clearly be the case if the amplitude of each scattered wave is small and if the grains are reasonably far apart, and if the emulsion layer is sufficiently thin.

The total scattered field is then just the linear sum of the individual wavefronts scattered by each grain within the emulsion layer. It makes no difference how the terms are grouped in the sum, provided each wavefront is counted once and only once. A convenient way to group terms is to consider all the grains associated with a

particular periodicity term in equation 1.9. The waves scattered by these particular grains will add in phase in certain directions, resulting in the "diffracted waves" which are produced by that periodicity. The calculation of these "diffracted waves" from a given periodicity, neglecting the presence of all other periodicities, is seen to be a logical extension of the stipulation that it is valid to treat the field scattered by each grain independently of all others.

The assumption is made that the amplitudes of the waves diffracted by each periodicity are proportional to the coefficient of the corresponding periodicity term in equation 1.9, and this is merely a statement that in effect says that doubling the number of grains contributing to the diffracted wave doubles the amplitude of the wave, since the individual scattered fields are summed coherently.

1.3.2 Nature of a Particular Periodicity - Restrictions on the Exposing and Illumination Fields

It is recalled that the propagation vectors $\bar{k}_n$ in equation 1.6 are not necessarily constants, but were assumed to be slowly varying in the region where the film plate was to be located. It is thus apparent that the periodicity associated with each cosine interference term in equation 1.9 is in general a function of position over the film plate. The stipulation that the $\bar{k}_n$ vary slowly is defined by requiring that this variation be sufficiently slow such that it is valid to compute the directions and phases of the

diffracted waves assuming the periodicity is (locally) constant. A similar restriction is placed on the illumination wave, namely that it behaves locally as a plane wave, or in the more general case, as a sum of such waves.

It is thus clear that what is required is the general solution of the problem of computing the direction and phase of the waves diffracted by a diffraction grating which was made by recording the interference pattern of two plane waves and which is illuminated with another plane wave. Consideration of this problem shows that it is the periodicity in the plane of the emulsion-air interface that determines the directions of the diffracted waves (along with, of course, the direction of the illumination wave, on the air side of the interface). Physically, this can be explained by a simple consideration of the implications of the assumption that multiple scattering can be neglected. If we consider the grains associated with the particular periodicity under consideration which lie within the infinitesimal layer between z' and $z' + dz'$ (z' defined normal to the emulsion-air interface), then it is clear that the directions in which the waves scattered by these grains add in phase are the same as for those grains that lie in the $z' = 0$ plane. Whether or not the waves scattered from grains within the two "planes" add in phase is a separate question, one which is dealt with in Chapter Four.

The periodicity in the plane of the emulsion-air interface ($z' = 0$) is specified by the fields that existed on either side of

the interface during exposure of the film plate. This is a consequence of the fact that the interference planes must match up at the interface in order to satisfy boundary conditions.

1.3.3 Interference Pattern of Two Plane Waves

In this section we shall compute the orientation and spacing of the intersection of the planes of maximum electric field in an interference pattern (produced by two plane waves) with the $z' = 0$ plane (plane of the emulsion-air interface). Thus, we consider the non-localized interference pattern generated by the two plane waves

$$\bar{E}_1 = \bar{E}_o \, e^{i(\bar{k}_1 \cdot \bar{r} + \phi_1)} \, e^{-i\omega t} \tag{1.10}$$

and

$$\bar{E}_2 = \bar{E}_o \, e^{i(\bar{k}_2 \cdot \bar{r} + \phi_2)} \, e^{-i\omega t} \quad . \tag{1.11}$$

What is of interest is the interference term in $\bar{E} \cdot \bar{E}^*$ where

$\bar{E} = \bar{E}_1 + \bar{E}_2$.

It is readily seen that

$$\bar{E} \cdot \bar{E}^* = 2E_o^2 + 2E_o^2 \cos[(\bar{k}_1 - \bar{k}_2) \cdot \bar{r} + \phi_1 - \phi_2] \quad . \tag{1.12}$$

The interference pattern is characterized by loci of points of maximum $\bar{E} \cdot \bar{E}^*$, which are a set of parallel planes defined by

$$(\bar{k}_1 - \bar{k}_2) \cdot \bar{r} + \phi_1 - \phi_2 = 2\pi M \tag{1.13}$$

where M is an integer. The lines of intersection of these planes with the plane of the emulsion surface, the $z' = 0$ plane, are the loci of points given by equation 1.13 with

$$\bar{r} = \bar{e}_{x'} x' + \bar{e}_{y'} y' \quad . \tag{1.14}$$

Letting ℓ_1', m_1', n_1' be the direction cosines of $\bar{k}_1$ and ℓ_2', m_2', n_2' be the direction cosines of $\bar{k}_2$, with respect to the x', y', z' coordinate system, the loci of the lines of intersection in the $z' = 0$ plane is given by

$$\frac{2\pi}{\lambda_o} (\ell_1' - \ell_2') x' + \frac{2\pi}{\lambda_o} (m_1' - m_2') y' + \phi_1 - \phi_2 = 2\pi M \quad . \tag{1.15}$$

These lines make an angle Φ with the x' axis given by

$$\Phi = \tan^{-1} \left(\frac{\ell_2' - \ell_1'}{m_1' - m_2'} \right) \quad . \tag{1.16}$$

We shall find it convenient to define the xyz coordinate system by a rotation of Φ about the z' axis (see Appendix III for the coordinate transformations and for the transformation equations for the direction cosines). In this coordinate system

$$\ell_1 = \ell_2 \tag{1.17}$$

and thus, in the $z = 0$ plane, equation 1.15 becomes

$$\frac{2\pi}{\lambda_o} (m_1 - m_2) y + \phi_1 - \phi_2 = 2\pi M \quad . \tag{1.18}$$

The periodicity D, or spacing of the intersection lines in the $z = 0$ plane is seen from equation 1.18 to be given by

$$d = \frac{\lambda_o}{|m_1 - m_2|} \quad . \tag{1.19}$$

This displacement δ of these lines, defined by setting $M = 0$ in equation 1.18, is given by

$$\delta = \frac{\lambda_o}{2\pi} \frac{(\phi_2 - \phi_1)}{(m_1 - m_2)} \quad . \tag{1.20}$$

1.3.4 Generalized Grating Equations

Equations 1.16, 1.19, and 1.20 determine the orientation, spacing, and displacement of the lines of intersection of the planes of maximum intensity with the emulsion surface plane. The problem is now one of computing the direction and phase of the diffracted waves when this periodic structure is illuminated by a plane wave of the form

$$\bar{E}_i = \bar{E}_o \, e^{i(\bar{k}_1 \cdot \bar{r} + \phi_i)} \, e^{-i\omega t} \quad , \tag{1.21}$$

where $\bar{k}_i = \frac{2\pi}{\lambda} \bar{e}_i$ and λ is not necessarily equal to λ_o . The directions $\bar{e}_d$ of the diffracted waves are determined by the requirement that the "elementary" waves scattered by each grain in the periodic structure add in phase. This condition can be stated geometrically in terms of the path lengths between two planes, A and

B, defined by

$$\bar{e}_i \cdot \bar{r} = c_1 \tag{1.22}$$

and

$$\bar{e}_d \cdot \bar{r} = c_2 \quad . \tag{1.23}$$

The constant c_1 is taken to be negative so that plane A is on the incident side of the diffracting structure, while c_2 is taken as positive so that plane B is on the transmitted side of the structure. We are interested in L, which is the path length between the planes A and B, going from plane A to a point P_o along $\bar{e}_i$, and then from P_o to the plane B along $\bar{e}_d$. We see that, if $\bar{r}_o$ is the position vector of P_o,

$$L = -c_1 + \bar{r}_o \cdot \bar{e}_i + c_2 - \bar{r}_o \cdot \bar{e}_d \tag{1.24}$$

or, taking P_o to lie in the $z = 0$ plane,

$$L = (\ell_i - \ell_d)\, x_o + (m_i - m_d)\, y_o + c_2 - c_1 \quad . \tag{1.25}$$

We now impose two conditions on L to determine the allowed directions of $\bar{e}_d$. The first of the two conditions is that L be independent of x_o. This assures that there is no net phase shift as we move along a grating line. From equation 1.25 we see that this requires that

$$\boxed{\ell_d = \ell_i} \quad . \tag{1.26}$$

The second of the two conditions is that L change by an integral number of wavelengths $N\lambda$ as y_o changes by d, the periodicity of the recorded interference pattern in the $z = 0$ plane. From equation 1.25 we see that this requires that

$$(m_d - m_i) = \frac{N\lambda}{d} \tag{1.27}$$

or, using equation 1.19, we see that

$$\boxed{(m_d - m_i) = \frac{N\lambda}{\lambda_o} |m_1 - m_2|} \tag{1.28}$$

The z direction cosine of $\bar{e}_d$ is determined from the condition that $\ell_d^2 + m_d^2 + n_d^2 = 1$ and the stipulation that we are considering transmitted waves, which means that n_d will have the same sign as n_i.

Equations 1.26 and 1.28 thus specify the direction cosines of the diffracted waves in terms of the direction cosines of the illumination wave and the two initial exposing waves. The x direction cosines of the initial exposing waves enter implicitly through the definition of the x,y,z coordinate system. The corresponding equations in the x', y', z' coordinate system are found by applying the transformation equations given in Appendix III:

$$\ell'_d - \ell'_i = -\frac{N\lambda}{d} \sin \Phi = -\frac{N\lambda}{\lambda_o} [(\ell'_2 - \ell'_1) \sin^2 \Phi + (m'_1 - m'_2) \sin \Phi \cos \Phi] \tag{1.29}$$

and

$$m_d'-m_i' = \frac{N\lambda}{d} \cos \Phi = \frac{N\lambda}{\lambda_o} [(\ell_2'-\ell_1') \sin \Phi \cos \Phi + (m_1'-m_2') \cos^2 \Phi] \quad . \quad (1.30)$$

Equations 1.26, 1.28, 1.29 and 1.30 are essentially the same as those used by Offner (14) in his work on ray tracing in holography. Offner references the work of Toraldo di Franci (21) who presents these results without reference or derivation.

The phase ϕ_d of the diffracted wave under consideration is found by requiring that the interference pattern generated by the transmitted portion of the illumination wave and the diffracted wave "match up" with the interference pattern recorded on the film plate, in the $z = 0$ plane. The intersection lines of the illumination interference pattern and the $z = 0$ plane are given by

$$\frac{2\pi}{\lambda} (\ell_i - \ell_d) x + \frac{2\pi}{\lambda} (m_i - m_d) y + \phi_i - \phi_d = 2\pi M \quad . \quad (1.31)$$

Recalling that $\ell_i = \ell_d$, the spacing of these lines d_i is given by

$$d_i = \frac{\lambda}{|m_i - m_d|} \quad . \quad (1.32)$$

Using the value of $|m_i - m_d|$ given by equation 1.28, with $N = \pm 1$, it is seen that

$$d_i = \frac{\lambda_o}{|m_1 - m_2|} \quad (1.33)$$

and hence is equal to the periodicity of the recorded interference pattern. We specify ϕ_d by requiring that the loci of minimum field intensity coincide with the loci of maximum grain density in the $z = 0$ plane. Thus we set $y = \delta$ and $M = 1/2$ in equation 1.31, and solve for ϕ_d. We obtain

$$\phi_d = \phi_i + \frac{\lambda_o}{\lambda} \frac{(m_i - m_d)}{(m_1 - m_2)} (\phi_2 - \phi_1) - \pi \quad . \tag{1.34}$$

Now from equation 1.28 we see that

$$\frac{\lambda_o}{\lambda} \frac{(m_d - m_i)}{|m_1 - m_2|} = N \tag{1.35}$$

where N is an integer. Thus, we can write equation 1.34 in the form

$$\phi_d = \phi_i + M(\phi_2 - \phi_1) - \pi \tag{1.36}$$

where M is an integer. We use M rather than N because of the ambiguity in sign due to the fact that we have the absolute magnitude of $m_1 - m_2$ in equation 1.35.

1.3.5 Interpretation of the Terms in the Grain Density Equation

Let us consider the expression for the grain density D given by equation 1.9. According to the theory presented here, the field diffracted by this composite periodic structure is found by summing the fields diffracted by each periodicity, i.e., by the grains

associated with each interference term in equation 1.9. In actuality only a portion of these terms is of interest and it will be possible to ignore most of the higher order terms. A discussion of the significance of the various terms in equation 1.9 is thus of interest at this time.

However, before doing this, it is useful to show that illumination of a particular periodicity with one of the two original waves that formed it yields a diffracted wave having the same direction and phase as that of the other original wave. Thus, let us consider a periodicity term of the form

$$\cos[(\bar{k}_i - \bar{k}_j)\cdot\bar{r} + \phi_i - \phi_j] \tag{1.37}$$

and let the illumination wave be given by

$$\bar{E}_i = \bar{E}_{io} e^{i(\bar{k}_i\cdot\bar{r} + \phi_i - \omega t)} \quad . \tag{1.38}$$

Then, in the xyz coordinate system defined by equation 1.16 with $1 = i$, $2 = j$, it follows from equations 1.17 and 1.26 that

$$\ell_d = \ell_j \quad . \tag{1.39}$$

Application of equation 1.28 yields

$$m_d = m_i + N|m_i - m_j| \tag{1.40}$$

and since N can assume the values ± 1, we see that one of the first order diffracted waves has

$$m_d = m_j \quad . \tag{1.41}$$

The phase of this wave is found from equation 1.34 and is given by

$$\phi_d = \phi_j - \pi \quad . \tag{1.42}$$

Thus, it is clear from equation 1.39, 1.41 and 1.42 that one of the two first-order diffracted waves has the same direction and phase (except for the constant factor of $-\pi$) as that of one of the two original exposing waves, when we illuminate with the other original exposing wave.

Returning to equation 1.9, the constant terms $C_o + C_1E_o^2 + C_2E_o^4 + \ldots$ will constitute a bias, which is required, of course, as it is physically meaningless to have a negative density, and the various cosine interference terms assume negative values. The terms

$$(2C_1E_o + 4C_2E_o^3) \sum_n b_n \cos[(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - \phi_n] \tag{1.43}$$

give rise to the real and virtual images, which are usually what is of interest in the reconstruction process. The real image is associated with one of the first-order diffracted beams for each elemental periodicity in equation 1.43, while the virtual image is associated with the other first order. Because of the sinusoidal variation of density of each of the periodicities, there are no second-order diffracted waves associated with each of the periodicities,* and thus N in equation 1.28 has allowable values of only -1, 0 and +1.

The "second order" real and virtual images which are observed are produced by the first-order waves diffracted by the periodicities associated with the terms

---

* This is discussed in Appendix VI.

$$2C_2E_o^2 \sum_n b_n^2 \cos[2(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - 2\phi_n] \qquad (1.44)$$

The second order images thus arise because the film is not linear in $\bar{E}.\bar{E}^*$, which will in general result in the coefficients C_2, C_3, ... being non-zero. The terms

$$(C_1 + 2C_2E_o^2) \sum_{n,m} C_{nm} \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m]$$

$$+ 2C_2E_o^2 \sum_{n,m} b_n b_m \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m] \qquad (1.45)$$

are of interest in "ghost imaging," and will be discussed in detail in Chapter Three.

The remaining higher order terms contribute little of interest and can usually be neglected due to the small size of their coefficients and usually unfavorable illumination conditions.* Thus, the diffracted fields that will be considered are those diffracted by the periodicities associated with the terms given in 1.43, 1.44 and 1.45. Each of these terms is of the form

$$\text{coefficient} \times \cos[(\bar{k}_1 - \bar{k}_2)\cdot\bar{r} + \phi_1 - \phi_2]$$

---

* By unfavorable illumination conditions we mean the case where the direction of the illumination wave is such that the waves scattered by the grains associated with the particular periodicity under consideration do not add in phase when the entire emulsion volume is taken into account, with the result that the amplitudes of the diffracted waves are quite small. This is discussed in detail in Chapter Four.

and thus equations 1.36, 1.38 and 1.44 can be used to compute the direction and phase of the waves diffracted by each periodicity.

1.3.6 General Illumination

It was shown in the previous section that the terms in equation 1.9 for the grain density D which are of interest can be written in the form

$$D = d_o + \sum_{i,j} d_{ij} \cos[(\bar{k}_i - \bar{k}_j)\cdot\bar{r} + \phi_i - \phi_j] \quad . \qquad (1.46)$$

If the hologram is illuminated by a wave of the form (which satisfies the restrictions mentioned in section 1.3.2)

$$\bar{E}_i = \bar{E}_{io}\, e^{i(\bar{k}_i\cdot\bar{r} + \phi_i)}\, e^{-i\omega t} \qquad (1.47)$$

then according to the analysis presented here, the diffracted field will be of the form

$$\bar{E}_d = \sum_{i,j,N} \bar{F}_{ijN}\, d_{ijN}\, e^{i(\bar{k}_{ijN}\cdot\bar{r} + \phi_{ijN})}\, e^{-i\omega t} \quad . \qquad (1.48)$$

The propagation vectors $\bar{k}_{ijN}$ are determined by applying equations 1.26 and 1.28, while the phase factors ϕ_{ijN} are found from either equation 1.34 or 1.36. The diffracted waves as given by equation 1.48 are taken to be proportional to the coefficient of the corresponding

periodicity term d_{ij} , as was discussed in Section 1.3.1. The real vectors $\bar{F}_{ijN}$ account for the amplitudes and polarizations of the diffracted waves. A rigorous computation of these factors requires the solution of the electromagnetic boundary value problem of the system. This, of course, requires more information about the emulsion layer than is given by the grain density D . The problem of the rigorous computation of $\bar{F}_{ijN}$ will not be considered in this thesis, but certain aspects of computing the amplitudes of the diffracted waves will be considered in Chapter Four.

In the more general case the illumination field may be a sum of terms such as given by 1.47. In this case the diffracted field would be the sum of fields given by 1.48, one such field for each illumination wavefront. This assumes that the field that is scattered by each grain within the emulsion layer when it is illuminated by a number of separate wavefronts is the linear sum of the fields that would be scattered by the grain for each illumination wave taken separately.

1.3.7 Reconstruction of the Original Fields (Virtual Image)

In this section we shall consider the case where the illumination beam $\bar{E}_i$ is identical to the reference beam that was used in exposing the film plate. We saw in Section 1.3.5 that illumination of a particular periodicity with one of the original exposing waves yields the other original exposing wave as one of the two first order diffracted waves. Thus the grains associated with the periodicity terms given by equation 1.43 give rise to a reconstruction of the signal beam when the

illumination wave is the reference beam. In actuality, the reconstructed field will be of the form

$$\bar{E}_d = \sum_j \bar{F}_j \, (2C_1 E_o + 4C_2 E_o^3) b_j \, e^{i(\bar{k}_j \cdot \bar{r} + \phi_j - \pi - \omega t)}$$

$$+ \text{zero order terms} + \text{real image terms} + \text{ghost image terms}$$

$$+ \text{second order terms} + \text{higher order terms} \qquad (1.49)$$

Comparing the diffracted field as given by equation 1.49 with the field that existed at the film plate during exposure of the hologram (equation 1.6), it is seen that except for the unimportant constant phase factor of $-\pi$, the waves given by the sum in equation 1.49 have the same direction and phase as those of the initial exposing field. Thus, except for possible questions regarding the amplitudes of the diffracted waves, one could say that equation 1.49 predicts the reconstruction of the original exposing wavefronts.

1.4 Reflection or "White Light" Holograms

In this section we shall consider reflection holograms. It will be seen that they can be treated quite adequately within the framework of the analysis presented in the preceding sections. The analysis is essentially the same until we compute the direction and phase of the diffracted waves, at which point a different method of adding the scattered fields is used.

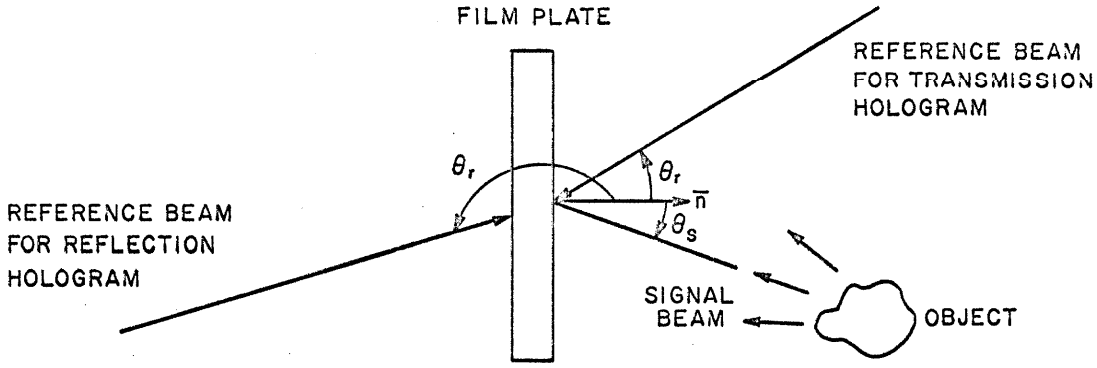
1.4.1 Introduction

The distinction between reflection and transmission holograms is made with regard to whether the reconstructed fields are produced on the same side (reflection) or on the opposite side (transmission) of the film plate with respect to the illumination wave. In the case of the transmission holograms, it was implicitly assumed that the reference beam and signal beam were incident on the film plate from the same side ($\theta_r \leq 90^o$ and $\theta_s \leq 90^o$ in Figure 1.1a.) In the case of the reflection hologram, however, the reference beam is brought in from the other side ($\theta_r > 90^o$ in Figure 1.1a.) When the illumination is done using the reference beam, the reconstructed signal beam (virtual image) is formed in transmission for the transmission hologram in reflection for the reflection hologram (Figure 1.1b,c.)

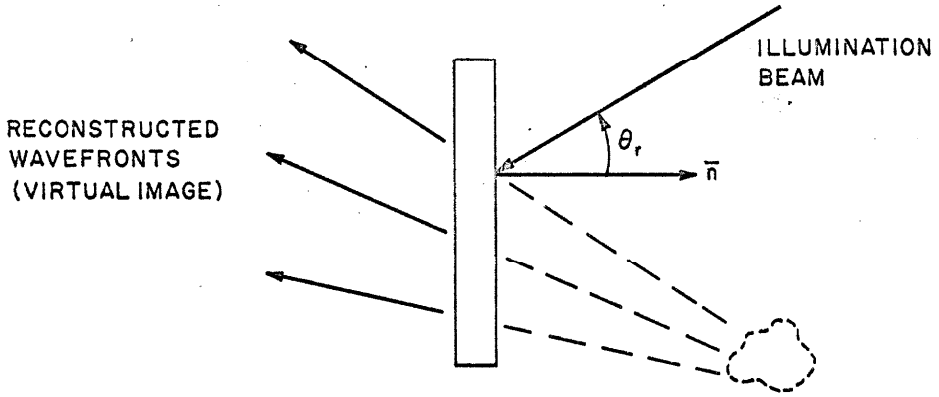
Reflection holograms were first investigated by Denisyuk (10, 11), who described the recording process in terms of the recording of the intensity distribution in the standing wave interference pattern formed by the reference beam and the light scattered by the object. In the reconstruction process, Denisyuk described the reflected field in terms of the waves reflected from the interference planes within the emulsion volume. He noted that the reflection hologram should act as an interference filter, reflecting only those wavelengths which lie in a narrow band about the wavelength used in exposing the plate.* This effect has

---

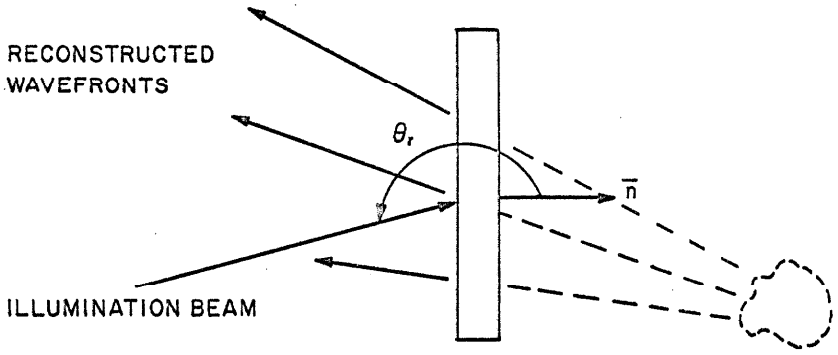
* Except for emulsion shrinkage effects, which shift this wavelength band. This is discussed further in Section 5.5.2.



(a) ARRANGEMENT OF REFERENCE BEAM FOR TRANSMISSION AND REFLECTION HOLOGRAMS



(b) RECONSTRUCTION ARRANGEMENT FOR TRANSMISSION HOLOGRAM



(c) RECONSTRUCTION ARRANGEMENT FOR REFLECTION HOLOGRAM

FIGURE 1.1 TRANSMISSION AND REFLECTION HOLOGRAMS

given rise to the use of the term, "white light hologram" for reflection holograms, due to the fact that a reconstruction can be obtained when the hologram is illuminated with white light.

Denisyuk, who was working with non-laser sources, used the transmitted portion of the reference beam to illuminate the object. The object was placed quite close to the plate in order to keep path length differences small. With the advent of the gas laser with its long coherence length, more efficient and elaborate experimental setups became practical. Stroke and Labryie (22) produced reflection holograms (using a laser as a source) which could be viewed in reflection with illimination provided by a flashlight or the sun. They described the reconstruction in terms of Bragg reflection from the grating-like stratifications within the emulsion layer. Lin et al (23) extended the experimental techniques to the use of two different wavelength lasers and made reflection holograms which yield multicolor reconstructions when viewed using white light. Other work in the area of reflection holograms has been reported by Upatnieks et al (24) and Stroke and Lech (25).

1.4.2 Recording Process

The description of the recording process presented in section 1.2 is sufficiently broad in scope that we may treat the case of reflection holograms without any modifications. Indeed, it is recalled that in section 1.2 there were no assumptions made with respect to the relative directions of arrival of the signal and reference beam wavefronts. Thus the formulation of the exposing fields as given by

equation 1.6 is adequate for application to the case of reflection holograms. The same restrictions that were placed on the nature of the individual waves which are represented by the terms in equation 1.6 are retained here, and thus the grain density of the developed reflection hologram film plate is given by equation 1.9. As before, we neglect any variations in thickness of the emulsion layer and consider both surfaces to be planes.

1.4.3 Linearity Hypothesis

The problem under consideration now is the determination of the direction and phase of the reflected fields that are produced when the reflection hologram is illuminated. The same assumptions that were made in section 1.3.1 are made here, namely that multiple scattering can be neglected. Thus, as was discussed in section 1.3.1, the field scattered by the grains associated with each basic periodicity within the emulsion layer can be computed independently of the presence of all others, and the total field is found by summing the fields scattered by each of the periodicities as specified by equation 1.9.

1.4.4 Standing Wave Interference Pattern of Two Plane Waves

What is of interest is the computation of the direction and phase of the fields scattered by the grains associated with a particular periodicity term in equation 1.9. It is recalled that the propagation vectors of the waves which gave rise to these interference terms were not necessarily constants, and hence the "periodicity" varies as a function of position throughout the emulsion. As before,

we will assume that the spatial variation of the propagation vectors was sufficiently slow such that we may regard the periodicities as (locally) constant. The problem is then reduced to determining the field scattered by the grains associated with an interference pattern generated by two plane waves. This is the approach that was used in the case of the transmission hologram, only in this case the scattered waves will be added in a different manner and hence it will be convenient to use a different coordinate system than the one that was used in section 1.3.3. In addition, we will wish to make the computations with respect to the fields within the emulsion layer.

Thus, let $\bar{E}_1$ and $\bar{E}_2$ be two plane waves which existed in the emulsion layer during exposure of the hologram, where

$$\bar{E}_1 = \bar{E}_o \, e^{i(\bar{k}_1 \cdot \bar{r} + \phi_1)} \tag{1.50}$$

and

$$\bar{E}_2 = \bar{E}_o \, e^{i(\bar{k}_2 \cdot \bar{r} + \phi_2)} \quad . \tag{1.51}$$

As before, we are interested in the interference term in $\bar{E}.\bar{E}^*$, where $\bar{E} = \bar{E}_1 + \bar{E}_2$ and

$$\bar{E} \cdot \bar{E}^* = 2E_o^2 + 2E_o^2 \cos[(\bar{k}_1 - \bar{k}_2) \cdot \bar{r} + \phi_1 - \phi_2] \quad . \tag{1.52}$$

The interference pattern within the emulsion layer is characterized by the loci of points of maximum grain density, which are parallel

planes a distance d' apart, and specified by

$$(\bar{k}_1 - \bar{k}_2)\cdot\bar{r} + \phi_1 - \phi_2 = 2\pi M \qquad (1.53)$$

where M is an integer. It is convenient to carry out the calculations in a coordinate system where the interference planes are perpendicular to the z'' axis. If the x', y', z' system is our standard reference coordinate system, then we define the x'', y'', z'' coordinate system by two coordinate rotations. First we rotate about the z' axis by an angle Φ (given by equation 1.16), to obtain the xyz coordinate system. In this coordinate system, equation 1.53 is of the form

$$(m_1 - m_2)\, y + (n_1 - n_2)\, z + \phi_1 - \phi_2 = 2\pi M \quad . \qquad (1.54)$$

The x'', y'', z'' system is then obtained by a rotation β about the x axis, where

$$\beta = \tan^{-1}\left(\frac{m_2 - m_1}{n_1 - n_2}\right) \quad . \qquad (1.55)$$

In this coordinate system equation 1.53 is of the form

$$(n_1'' - n_2'')\, z'' + \phi_1 - \phi_2 = 2\pi M \qquad (1.56)$$

(the coordinate transformations and transformation equations for the direction cosines are given in Appendix III.) The above equation specifies the planes of maximum grain density within the emulsion layer for the special case where the two exposing fields are plane

waves. If they are not plane waves ($\bar{k}_1$ and $\bar{k}_2$ being slowly varying functions of position) then equations 1.16, 1.55, and 1.56 are applied at the particular location of interest within the emulsion layer with the appropriate values of $\bar{k}_1$ and $\bar{k}_2$ being used.

1.4.5 Bragg Reflection Condition

In this section we wish to determine the field scattered by those grains which are associated with the periodicity defined by equation 1.56 when the illumination field is a plane wave of the form

$$\bar{E}_r = \bar{E}_{ro}\, e^{i(\bar{k}_r \cdot \bar{r} + \phi_r)} \quad . \tag{1.57}$$

Equation 1.57 specifies the illumination field within the emulsion layer, after refraction at the emulsion-air interface.

Each grain, of course, scatters a portion of the illumination wave in essentially all directions. We are only interested in the particular case where the waves scattered by the grains under consideration add in phase in a particular direction. This will occur if the wavelength and direction of the illumination wave are such that Bragg reflection from the planes of constant grain density occurs. The Bragg reflection condition can be simply stated in terms of the spherical coordinate θ_r'' (of $\bar{k}_r$), the illumination wavelength λ_r, and the distance d' between planes of maximum grain density. It is

$$\cos \theta_r'' = \frac{N\lambda_r}{2d'} \quad , \tag{1.58}$$

where N is an integer. If θ_r'' and λ_r are such that the above equation is satisfied, then the waves scattered from the grains in the planes $z'' = c + Md'$ will add in phase in the direction $\bar{e}_d$, whose spherical coordinates are

$$\phi_d'' = \phi_d'' \tag{1.59}$$

and

$$\theta_d'' = \pi - \theta_r'' \tag{1.60}$$

where c is a constant and M is an integer.

It is of interest to express d' in equation 1.58 in terms of the spherical coordinates of the propagation vectors of the two plane waves which generated the periodicity. It is straightforward to show that if

$$\alpha = \cos^{-1} \left(\frac{\lambda^2}{4\pi^2} \bar{k}_1 \cdot \bar{k}_2 \right) \tag{1.61}$$

then

$$d' = \frac{\lambda}{2 \sin (\alpha/2)} \tag{1.62}$$

Now, as a consequence of our choice of the x'', y'', z'' coordinate system

$$\phi_1'' = \phi_2'' \tag{1.63}$$

and

$$\theta_1'' + \theta_2'' = \pi \quad . \tag{1.64}$$

It follows from equations 1.61, 1.63, and 1.64 that

$$\sin\left(\frac{\alpha}{2}\right) = \cos\,\theta_1'' = -\cos\,\theta_2'' \tag{1.65}$$

and hence, using equations 1.58, 1.62, and 1.65, the Bragg reflection condition becomes

$$\cos\,\theta_r'' = \frac{N\lambda_r}{\lambda}\cos\,\theta_1'' = -\frac{N\lambda_r}{\lambda}\cos\,\theta_2'' \quad . \tag{1.66}$$

Equation 1.66 expresses the conditions which are placed on the propagation vector $\bar{k}_r$ of the illumination wave in terms of the two original illumination waves which generated the periodicity under consideration. A similar equation exists for each periodicity term in equation 1.9. (It should be noted that there is in general a different x'', y'', z'' coordinate system associated with each periodicity.)

It is of interest to determine if there is a particular $\bar{k}_r$ which will satisfy equation 1.66 for a significant number of periodicity terms in equation 1.9. An examination of equation 1.9 shows that the terms which yield the virtual image in transmission all have the propagation vector $\bar{k}_o$ in common, and the interference terms are all of the form of the interference term in equation 1.52.

It follows from equation 1.66 that if the illumination wave is the same as either of the two original waves, then equation 1.66 is satisfied (with $|N| = 1$). Furthermore, it follows from equations 1.59, 1.60, 1.63 and 1.64 that illumination by one of the two original waves yields the other as the reflected wave. The phase factor ϕ_d for the reflected wave is specified by requiring that the planes of minimum intensity in the interference pattern generated by the illumination beam and the reflected beam coincide with planes of maximum grain density. For example, consider the interference term generated by the reference beam and the j^{th} signal wave. The recorded interference pattern is, from equation 1.56,

$$(n_o'' - n_j'')\, z'' - \phi_j = 2\pi M \quad . \tag{1.67}$$

When we illuminate with the reference beam, the loci of points of minimum electric field intensity is given by

$$(n_o'' - n_j'')\, z'' - \phi_d = 2\pi(M + \tfrac{1}{2}) \tag{1.68}$$

and hence

$$\phi_d = \phi_j - \pi \quad . \tag{1.69}$$

Thus, except for the unimportant constant phase factor $-\pi$, the phase of the reflected wave equals that of the other original illumination wave, and hence we can say that illumination of the periodicity by one of the two waves which produced it yields a "reflected wave" whose direction and phase equals that of the other wave.

It thus follows that when we illuminate the hologram with the reference beam, Bragg reflection will result in reflected waves whose directions and phases are the same as those of the original signal beam, and hence a reconstruction (virtual image) is produced.

1.5 Comparison of Transmission and Reflection Holograms

There are a number of fundamental differences between the two types of holograms that warrant discussion here. First of all, while the directions and phases of the "diffracted" or "reflected" fields are determined by requiring that the individual scattered waves add in phase, there is a basic difference in the way in which we "group terms" in summing these scattered waves. In the case of the transmission hologram the directions of the diffracted waves (i.e., the directions in which the individual scattered waves add in phase) are determined from the periodicities in the plane of the emulsion surface, and these directions are expressed by what could be termed "generalized diffraction grating equations." On the other hand, in the case of reflection holograms, the individual scattered waves add in phase when the Bragg reflection condition is satisfied with respect to the planes of constant grain density within the emulsion layer.

It can be immediately seen that with respect to the determination of the directions of the "diffracted" or "reflected" waves, the variation of grain density with depth is of primary importance in the case of the reflection hologram, but of only secondary importance in transmission holograms of the type considered here. Indeed, we could

let the emulsion thickness approach zero in a transmission hologram and still obtain a reconstruction, while this would result in the complete disappearance of the reconstruction in the case of the reflection hologram.

It was seen in Section 1.3.4 that there were two directions in which the waves scattered by the grains associated with a particular periodicity of the transmission hologram add in phase. This is analogous to the two first orders produced by a diffraction grating, and in the case of the transmission hologram, these two directions correspond to the real and virtual images. The situation is quite different in the case of the reflection hologram, where, as was seen in Section 1.4.5, there is only a single direction in which the scattered waves from a particular periodicity add in phase. Thus, only a single image is formed by a reflection hologram, and as was seen in Section 1.4.5, when the illumination wave is the reference beam, the single image is the virtual image. This does not, however, preclude the formation of a real image. Denisyuk (10,11) explains the conditions placed on the exposing and illumination beams that must be satisfied in order for the single reconstructed image to be real.

Perhaps the most striking difference between the two types of holograms is the difference in reconstructions for the case when the illumination wave has a wide range of spectral components (the geometrical characteristics being the same as that of the reference beam). In the case of the transmission hologram, each spectral component yields diffracted waves whose directions are specified by equations

1.26 and 1.28. It is seen from equation 1.28 that the directions are a function of λ, and thus each wavelength will yield an image displaced some amount in angle with respect to the images produced by the other spectral components in the illumination beam. The end result is a blurred image, the blurring becoming worse as the spectral range of the illumination beam is increased.

The situation is quite different in the case of the reflection hologram, as the Bragg reflection condition (equation 1.66) must be satisfied in order to obtain any reconstruction. If the illumination beam has the geometrical characteristics of the reference beam but has a range of spectral components, only those wavelengths which satisfy equation 1.66 will yield scattered waves which will add in phase, and hence give a reconstruction. The other spectral components do not produce images and hence do not degrade the reconstructed image. Thus, one can illuminate a reflection hologram with "white light" and still obtain a reconstruction of reasonable quality.

There are a number of practical considerations that modify the previous statements regarding reflection holograms. In practice, film plates such as Kodak 649-f high resolution plates are used in making either transmission or reflection holograms, and the emulsion layer is typically 15μ or less in thickness. In the case of the reflection hologram this means that there will be only a limited number of interference planes contributing to the Bragg reflection phenomena. The radiation pattern or "array factor" associated with each periodicity in the reflection hologram is thus much less directive than in the case of

the transmission hologram, where the periodicity is in the plane of the emulsion surface and hence many more basic periods are included in the diffracting structure. The result is that the reflection hologram will yield a reconstruction over a band of wavelengths, this band being narrower for thicker emulsions. In addition there is much less sensitivity to the angle of incidence of the illumination wave. An additional consideration is the fact that emulsion shrinkage may occur during processing of the film plate, and this will result in the shifting of the wavelength band for reconstruction of the reflection hologram to shorter wavelengths. Experimental data as well as an analytical treatment of the wavelength sensitivity of reflection holograms can be found in the works of Denisyuk (10,11) and Fleisher et al (27). Although the authors of (27) don't specifically consider holograms, their work is directly applicable and includes a worthwhile amount of information concerning processing of high resolution film plates.

1.6 Summary and Discussion

The analytical description of the holographic process that has been formulated in this chapter takes into account the three-dimensional nature of the recording media by characterizing the emulsion volume by the volume density of scattering particles, the grain density. A general film response is allowed for by expressing the grain density as a power series in $\bar{E}\cdot\bar{E}^*$, where $\bar{E}$ is the total electric field.

By expressing the exposing and illumination fields in the form of a sum of plane or quasi-plane waves and neglecting multiple

scattering we have been able to handle the problem of computing the diffracted fields by using what could be termed a linear systems approach. This involves computing the waves diffracted by the grains associated with each periodicity term in the expression for the grain density, and then summing (coherently) these waves to obtain the total diffracted field. The problem of computing the diffracted field is thus reduced to solving variations of the same basic problem, that of computing the direction, amplitude and phase of each of the two first-order diffracted waves that are produced when a "volume diffraction grating" is illuminated with a plane wave. We solve this problem for the general case, considering the directions and phases in this chapter and the amplitudes in Chapter Four (for transmission holograms only).

It should be noted that we have been dealing with the values of the fields in the immediate vicinity of the hologram plate and have not considered the problem of relating the exposing and illumination fields to the sources that generate them or the problem of computing the diffracted field at some distant observation point. These problems can be dealt with by utilizing Fresnel-Kirchhoff diffraction theory (as discussed in Chapter Two) or perhaps by using some sort of geometrical optics or rays tracing technique, as was done by Offner (14).

In general we shall deal primarily with the fields in the vicinity of the hologram plate, except where we are considering a specific problem where the field in some other region is of particular interest, as is the case with Fourier transform holograms. We shall see

that a knowledge of the field in the vicinity of the hologram plate is sufficient to provide the solutions of a number of problems which are of interest.

CHAPTER TWO

TRANSMITTANCE DESCRIPTION OF WAVEFRONT RECONSTRUCTION

2.1 Introduction

In this chapter we shall consider the transmittance description of wavefront reconstruction as developed by Gabor (5). We will see that although the use of the concept of amplitude transmittance becomes questionable as the spatial frequencies involved become high (i.e., when the emulsion must be regarded as "thick"), the analytical formulation still remains valid. Furthermore, we will show that with respect to computing the directions and phases of the diffracted waves the transmittance approach and the approach formulated in Chapter One are equivalent. We will then compare the two approaches.

2.2 Transmittance Approach

The term "transmittance approach" as used here refers to a particular formulation of the problem of determining the diffracted field produced when the hologram is illuminated. The diffracted field is to be specified in terms of the initial exposing fields, the film characteristics, and the illumination field. The fields referred to are specified in the immediate vicinity of the film plate, and the problem of relating the exposing field to the sources which generate it and the problem of computing the diffracted field at some distant observation point are considered as separate problems.

2.2.1 Analytical Formulation

The following formulation is due to Gabor (5) and has been successfully used by many researchers to treat many different problems in holography. Letting the $z = 0$ plane coincide with the surface of the film emulsion layer, the exposing field U is taken to be of the form ($e^{-i\omega t}$ suppressed)

$$U(x,y) = U_1(x,y) + U_2(x,y) \tag{2.1}$$

where U_1 and U_2 are complex scalar quantities which we identify as the reference beam and signal beam respectively. It is convenient to write

$$U_1(x,y) = A_1(x,y)\, e^{i\psi_1(x,y)} \tag{2.2}$$

and

$$U_2(x,y) = A_2(x,y)\, e^{i\psi_2(x,y)} \tag{2.3}$$

where A_1, ψ_1, A_2 and ψ_2 are real functions.

In the recording process the amplitude transmittance τ of the developed film emulsion layer is taken as the quantity which specifies the characteristics of the developed film plate which are of interest. It is generally assumed that τ is real and can be found in terms of the intensity UU^* from the characteristic curve for the film. Furthermore, if one assumes that the amplitudes of U_1 and U_2 (and the exposure time) are chosen so as to have the exposure be in the "linear" range of the characteristic curve, then the amplitude

transmittance τ can be expressed in the form

$$\tau = b_o + (UU^*)^{\Gamma/2} \tag{2.4}$$

where b_o is a constant and Γ is the "gamma" of the film. It is convenient to take $\Gamma = 2$ as this simplifies the algebra. This can be justified by noting that if we make the amplitude of the reference beam much greater than that of the signal beam, $(UU^*)^{\Gamma/2}$ can be expanded in a binomial series, the first order terms corresponding to the case $\Gamma/2 = 1$. That is, writing

$$(UU^*)^{\Gamma/2} = [A_1^2 + A_2^2 + 2A_1A_2 \cos(\psi_1 - \psi_2)]^{\Gamma/2}$$

or

$$(UU^*)^{\Gamma/2} = A_1^2 \left[1 + \frac{A_2^2}{A_1^2} + \frac{2A_2}{A_1} \cos(\psi_1 - \psi_2)\right]^{\Gamma/2} \tag{2.5}$$

and assuming $A_2/A_1 \ll 1$, we see

$$(UU^*)^{\Gamma/2} \simeq A_1^2 \left[1 + \frac{A_2^2}{A_1^2} + \frac{\Gamma A_2}{A_1} \cos(\psi_1 - \psi_2)\right]$$

or

$$(UU^*)^{\Gamma/2} \simeq A_1^2 + A_2^2 + \Gamma A_1A_2\cos(\psi_1 - \psi_2) \quad . \tag{2.6}$$

Thus, taking $\Gamma = 2$ and suppressing the constant b_o, the amplitude transmittance of the hologram plate is of the form

$$\tau = A_1^2 + A_2^2 + A_1A_2\, e^{i(\psi_2 - \psi_1)} + A_1A_2\, e^{-i(\psi_2 - \psi_1)} \quad . \tag{2.7}$$

Thus, by definition, if we illuminate the hologram with a wave U_3, the transmitted wave U_T will be given by

$$U_T = \tau U_3 \quad . \tag{2.8}$$

Writing

$$U_3 = A_3 e^{i\psi_3} \tag{2.9}$$

equation 2.8 becomes, using equation 2.7,

$$U_T = (A_1^2 + A_2^2)\, A_3 e^{i\psi_3} + A_1 A_2 A_3 e^{i(\psi_2 - \psi_1 + \psi_3)} + A_1 A_2 A_3 e^{-i(\psi_2 - \psi_1 - \psi_3)} \tag{2.10}$$

The first term is the zeroth order term, the second term is the virtual image term and the third term is the real image term.

2.2.2 Variations with Depth

The choice of the amplitude transmittance as the quantity which characterizes the developed film emulsion layer implies that we neglect or "average out" any variations in the z directions. This is clearly a good approximation when the thickness of the emulsion layer is considerably less than the period of the highest spatial frequency component in UU^* . However, it is not a good approximation in those cases where high spatial frequencies are involved, such as with the two-beam method (2) which is extensively used because it provides an angular separation of the images. For example, the emulsion layer on typical high resolution film plates used in holography may be from 5 to 15 microns thick, and from equation 1.62 it is seen that with a

wavelength of 6328A° and an angle of 7° between two plane waves the period of the resulting interference pattern will be of the order of 5 microns. One could go on and mention examples where the interference fringe planes are inclined with respect to the z axis and where averaging over z would "wash out" the interference pattern. In such cases the strict interpretation of τ in equation 2.8 as the actual amplitude transmittance would clearly rule out the validity of this approach. However, it will be seen shortly that equation 2.8 is valid even in such cases, but the interpretation of τ must be different.

This should not be too surprising as it was seen in Chapter One that the directions and phases of the "diffracted" waves are determined by the periodicity in the $z = 0$ plane. In the expression for τ as given by equations 2.4 or 2.7, it is the fields in the $z = 0$ plane that are used, and it is precisely the fact that the variations with z are not included that allows the use of equation 2.8 in those cases where the concept of amplitude transmittance breaks down.

2.2.3 Equivalence of the Scattering Theory Approach and the Transmittance Approach

In this section we shall show that with respect to computing the directions and phases of the diffracted waves the transmittance approach as given by equations 2.7 and 2.8 is equivalent to the approach given in Chapter One for transmission holograms. In particular we shall consider the special case where U_1, U_2, and U_3 are plane waves, and we shall show that for this case equations 2.7 and 2.8 yield the

equations given in Chapter One for determining the direction and phase of the diffracted waves.

Hence, let us take

$$U_1 = A_1 e^{i(\bar{k}_1 \cdot \bar{r} + \phi_1)} \tag{2.11}$$

$$U_2 = A_2 e^{i(\bar{k}_2 \cdot \bar{r} + \phi_2)} \tag{2.12}$$

and

$$U_3 = A_3 e^{i(\bar{k}_3 \cdot \bar{r} + \phi_3)} \tag{2.13}$$

where A_1, A_2, and A_3 are constants.

It is convenient to express the quantities $\bar{k}_1 \cdot \bar{r}$, $\bar{k}_2 \cdot \bar{r}$ and $\bar{k}_3 \cdot \bar{r}$ in the x,y,z coordinate system defined as in Section 1.3.3 (equation 1.16) so that the x direction cosines of $\bar{k}_1$ and $\bar{k}_2$ are equal (the $z = 0$ plane coinciding with the emulsion surface). Thus, in the $z = 0$ plane

$$\bar{k}_1 \cdot \bar{r} = \frac{2\pi}{\lambda_o} (\ell_1 x + m_1 y) \tag{2.14}$$

$$\bar{k}_2 \cdot \bar{r} = \frac{2\pi}{\lambda_o} (\ell_2 x + m_2 y) \tag{2.15}$$

$$\bar{k}_3 \cdot \bar{r} = \frac{2\pi}{\lambda} (\ell_3 x + m_3 y) \tag{2.16}$$

Let us consider the second term in equation 2.10, which is the term which corresponds to the virtual image if U_1 is identical as the reference beam. It is clear that since we have taken U_1, U_2, and

U_3 to be plane waves, the diffracted waves given by equation 2.10 will also be plane waves. Thus the factor $\psi_2 - \psi_1 + \psi_3$ will be of the form (in the $z = 0$ plane)

$$\psi_2 - \psi_1 - \psi_3 = \frac{2\pi}{\lambda} (\ell_4 x + m_4 y) + \phi_4 \quad . \tag{2.17}$$

But ψ_1, ψ_2, and ψ_3 are specified by equations 2.14, 2.15, and 2.16, hence (recalling $\ell_1 = \ell_2$)

$$\frac{2\pi}{\lambda} (\ell_4 x + m_4 y) + \phi_4 = \frac{2\pi}{\lambda} (\ell_3 x + m_3 y) + \frac{2\pi}{\lambda_o} (m_2 - m_1) y + \phi_2 - \phi_1 + \phi_3 \quad . \tag{2.18}$$

If this equation is to be valid for all values of x and y, the x and y coefficients must be equal. That is

$$\ell_4 = \ell_3 \tag{2.19}$$

and

$$m_4 - m_3 = \frac{\lambda}{\lambda_o} (m_2 - m_1) \quad . \tag{2.20}$$

Now equations 2.19 and 2.20 are identical to the equations derived in Section 1.3.3 that specify the directions of the diffracted waves which occur when a particular periodicity is illuminated by a plane wave. That is, equation 2.19 is identical to equation 1.26 and equation 2.20 is identical to equation 1.28 for the case $|N| = 1$ (virtual image). Similar results are obtained for the real image term in equation 2.10, the only difference being a change of sign of the

right hand side of equation 2.19 which corresponds to the case where the opposite sign is taken in equation 1.38.

The phase factor ϕ_4 for the virtual image is found to be

$$\phi_4 = \phi_3 + \phi_2 - \phi_1 \tag{2.21}$$

while the phase factor ϕ for the real image is found to be

$$\phi = \phi_3 - \phi_2 + \phi_1 \ . \tag{2.22}$$

Except for the absence of the constant factor $-\pi$, equations 2.21 and 2.22 are identical to equation 1.36. It was seen in Section 1.3.4 that the factor of $-\pi$ is a consequence of the boundary conditions requiring that the total $\bar{E}$ field be a minimum in a region of maximum grain density. We can put this in the context of the transmittance theory by noting that in Section 1.3.4 we are dealing with a "negative." If we were to make a "positive" (assuming that the spatial frequency is sufficiently low that it is meaningful to speak in such terms) then we would in effect shift the fringe pattern by 1/2 period and the factor of $-\pi$ would no longer be present.

We have thus demonstrated, for the special case of plane waves, that with respect to computing the directions and phases of the diffracted waves, the transmittance approach and the approach described in Chapter One are equivalent. It is clear that this equivalence will also hold when the fields are of the form of a sum of plane or quasi-plane waves, or when the fields can be expressed as a continuous distribution of such waves.

Having demonstrated the equivalence of the two approaches with respect to the directions and phases of the diffracted waves, it is logical to ask whether any such equivalence exists with respect to the amplitudes of the diffracted waves. The formulation of the analysis as given in Chapter One was done so as to take into account the entire emulsion layer. Using this formulation, the problem of predicting the reconstruction efficiency and the distribution of amplitudes in the diffracted waves will be dealt with in Chapter Four. The transmittance approach, as outlined in Section 2.2.1, does not take into account the entire emulsion layer, but instead deals only with the fields in the $z = 0$ plane. Thus, we would expect that the two approaches should yield different predictions when the emulsion is "thick" (i.e., high spatial frequencies, with period comparable to the emulsion thickness) but should agree in the limit as the ratio of emulsion thickness to minimum fringe spacing approaches zero.

This can be seen to be the case as follows: In the scattering theory approach the grain density is taken to be the quantity which specifies the developed film emulsion layer. Any particular diffracted wave is identified as being the result of coherent scattering by a certain number of grains within the emulsion layer which are associated with a particular periodicity term in equation 1.9. The amplitude of this diffracted wave is taken to be proportional to the amplitude of the illumination wave and proportional to the total number of grains N associated with the corresponding periodicity term in equation 1.9. Examination of the periodicity terms in equation 1.9

which yield either the real or virtual images shows that N is proportional to the product of the amplitudes of the signal beam and reference beam. In terms of the example used in the first part of this section (where the fields are given by equations 2.11, 2.12 and 2.13), this means that according to the scattering theory approach, the amplitude of the diffracted waves (real or virtual images) will be proportional to $A_1A_2A_3$, just as is predicted by the transmittance approach, as can be seen from equation 2.10. However, the constant of proportionality for the real and virtual images is the same according to the transmittance approach but may be different according to the scattering theory approach. In fact, this "proportionality constant" is not a constant at all, but a factor which is a function of the illumination wave geometry, emulsion characteristics, and other factors as discussed in Chapter Four. It will be shown in Chapter Four that this factor may differ by orders of magnitude between the real and virtual image. These differences are due to the different conditions for having the fields scattered by the grains in the different planes z = constant add in phase. In the limiting case where the emulsion thickness goes to zero this effect clearly becomes unimportant and the two approaches yield the same result. The same conclusion is reached in the case where the emulsion thickness remains appreciable, but where the spatial frequencies decrease to the point where the ratio of emulsion thickness to minimum fringe spacing approaches zero. In this case angular separation of the two images is sufficiently small and the "width" of the orientation sensitivity curve (see Chapter Four)

is sufficiently large that there is no appreciable difference in the amplitudes of the two images.

2.3 Comparison of the Two Approaches

As a basis of comparison of the two approaches, we shall consider the range of applicability of the two approaches and the ease with which calculations can be made using them. The concept of the amplitude transmittance was clearly justified in the early work of Gabor and others where the spatial frequencies were low. The analytical formulation of this approach has been shown to remain valid even for high spatial frequencies, but we have seen in such a case that the interpretation of the quantity τ (as given by equation 2.4 or 2.7) as the amplitude transmittance may not be correct. Instead, we must view τ as specifying the periodicities in the plane of the emulsion surface.

In comparison, the formulation given in Chapter One takes into account the entire volume of the emulsion layer, and allows, in a very general way, the treatment of both transmission and reflection holograms. By considering the coherent scattering by the grains within the emulsion layer we were able to show that the directions and phases of the diffracted waves are specified by the periodicities within the emulsion layer. In the case of the transmission hologram we saw that the directions and phases of the diffracted waves are specified by the periodicities in the plane of the emulsion surface. This verifies in physical terms the validity of the application of the transmittance

approach in those cases involving high spatial frequencies.

In addition, the formulation given in Chapter One provides a convenient basis for the investigation of the dependence of the reconstruction efficiency on the film characteristics and processing procedures as well as for the determination of the relative amplitudes of the diffracted waves. The two methods, as we have seen above, yield similar predictions for the amplitudes of the diffracted waves when the ratio of emulsion thickness to minimum fringe spacing is small. However, when this is not the case, the approach of Chapter One is clearly the better of the two methods. If the problem is not one of computing the relative amplitudes of the waves in the two images, but rather that of computing the relative amplitudes of the waves associated with one of the images (say the virtual image, with the illumination beam being the reference beam), then the transmittance approach may be satisfactory, as variations in the "proportionality constant" between waves in the same image may be small.

We have seen that with respect to calculations involving the directions and phases of the diffracted waves, the two methods are equivalent. In the scattering theory approach, the fields at the film plate are assumed to be in the form of a sum of quasi-plane waves. The diffracted field is then given as a sum of such waves, where the direction and phase of each wave at each point of the film plate is computed according to the formalism developed in Chapter One. In the transmittance approach, on the other hand, the exposing and illumination fields are written in a very simple form (equations 2.2, 2.3 and 2.9)

and the diffracted field is found using the simple relationship given by equation 2.8 or 2.10.

It is often the case, however, that part of the over-all problem is that of relating the exposing fields to the sources which generate them and computing the diffracted field at some distant observation point when the hologram is illuminated. Then, depending on the problem under consideration, one or the other of the two methods may be more useful. For example, consider the case where the exposing fields consist of a well defined reference beam plus the light scattered from the surface of a diffusely reflecting object. It would clearly be quite difficult to specify $A_2(x,y)$ and $\psi_2(x,y)$ in equation 2.3. On the other hand, the scattered field could be represented as arising from a number of point sources distributed over the surface of the object, and thus would be of the form of the field given by equation 1.4. The diffracted field could then be computed in a straightforward (but lengthy) manner for any arbitrary illumination field which can be put in the form of a sum of quasi-plane waves.

In contrast, an example where the transmittance formulation is most useful would be for the case where the exposing field consists of the light diffracted by a two-dimensional transparency plus a plane or spherical wave reference beam. Fresnel-Kirchhoff diffraction theory could then be used both for specifying the exposing field at the film plate and for computing the field at some observation point when the hologram is illuminated. These calculations can be put in transform language and discussed within the framework of communication theory (2).

Calculations of these types, using the transmittance formulation, have formed the basis of the majority of hologram investigations thus far. We shall have occasion to use both approaches in the following chapter.

Finally, we note that what might prove to be the most useful aspect of the approach developed in Chapter One is that it allows us to gain a good deal of insight into the holographic process by examining in detail a very simple type of hologram, the holographic diffraction grating, formed by recording an interference pattern which is generated by two plane waves.

CHAPTER THREE

MULTIPLE WAVEFRONT REFERENCE BEAM HOLOGRAPHY - GHOST IMAGING

3.1 Introduction

The concept of multiple reference beam holography or "ghost imaging" as used here refers to the case where the reference beam is no longer a simple, well-defined wave (such as a plane or spherical wave), but instead consists of a discrete sum or continuous distribution of such waves. In particular, these waves may arise from a portion of the object itself.

The first investigation of such a case was made by P. J. van Heerden who, in a paper developing the theory of the intensity filter (12), predicted that when a planar Fourier transform hologram (or "intensity filter") is illuminated with a portion of the original exposing field, a reconstruction of the remainder of the exposing field is obtained. In the optical system considered by van Heerden, this reconstructed field is brought to a focus to form an image of the original transparency, and this image was referred to as a "ghost image."

The first experimental observation of such a "ghost image" appears to have been made by Stroke et al (35), who were investigating the effect of "extended" sources on the resolution attainable in Fourier transform holography. This was followed shortly by a series of experiments by Collier and Pennington (36,37), who verified van Heerden's original prediction that a translation of the portion of

the object generating the "reference beam" during the illumination results in a corresponding translation of the ghost image, which remains in register with the image of the object fragment. They also presented experimental results showing the formation of a ghost image with a diffusely illuminated transparency as well as with a diffusely reflecting object. In the latter case, they observed that repositioning of the hologram plate was critical, and that things had to remain as they were during the exposure in order to obtain a reconstruction of the "ghost image."

In this chapter we shall be interested in examining the mechanism of the formation of the ghost image (i.e., a reconstruction of the signal beam when a complicated reference beam is used), both from the point of view taken in Chapter One and from the point of view taken in Chapter Two. We shall examine the effect of errors in repositioning of the hologram plate (the "translational sensitivity") as well as the background noise that arises when the reference beam is no longer a single wavefront.

We shall begin by considering the case where the reference beam is of the form of a sum of plane or quasi-plane waves, such as would arise, for example, if the reference beam was generated by the light from a diffusely reflected object. After computing the effect of repositioning errors and background noise, we shall describe an experiment in which the power in the reconstructed signal beam is measured as a function of repositioning error, or translation of the hologram from its original position. The experimental results are

then compared with the results obtained by applying the analysis presented here.

We shall then go on to consider ghost imaging with Fourier transform holograms, and investigate the translational sensitivity and background noise both for plane wave illumination (of the transparency) and for diffuse illumination.

3.2 Holograms with Multiple Wavefront Reference Beams

In this section we shall consider the case where the reference beam is more complex than the simple single wavefront reference beam considered previously, as is the case in "ghost imaging" or holography with "extended sources". We shall use the approach outlined in Chapter One, namely where we treat the exposing fields as sums of locally plane waves. This approach is in principle essentially that used by van Heerden (13) in his treatment of ghost imaging in his paper on information storage in three-dimensional media, the basic difference being that we shall allow for a (slow) variation of the propagation vectors across the film plate.

We shall be concerned with the case where the illumination beam is the reference beam, and shall examine in detail the effects of slight changes in the reference beam (such as caused by an error in repositioning the film plate). Furthermore, we will consider the nature of the reconstructed field (virtual image) and examine "noise" or "distortion" terms which generally arise when a multiple wave reference beam is used. We shall treat this "background noise" both from the point of view taken in Chapter One and from the "transmittance" point

of view, and will show that the results are equivalent, as is to be expected.

3.2.1 Recording Process

We are now interested in the case where the "reference beam" is no longer a single plane or spherical wave but is rather a sum of such waves. The field at the film plate is still of the form of equation 1.4, but now the first P terms are identified as the reference beam and the last $M-P$ terms as the signal beam. The field at the film plate is thus written in the form ($e^{-i\omega t}$ suppressed)

$$\bar{E} = \sum_{j=1}^{P} \bar{E}_j \, e^{i(\bar{k}_j \cdot \bar{r} + \phi_j)} + \sum_{m=P+1}^{M} \bar{E}_m \, e^{i(\bar{k}_m \cdot \bar{r} + \phi_m)} \quad . \qquad (3.1)$$

With $\bar{E}$ in this form the expression for the grain density D given by equation 1.9 becomes

$$D = C_o + C_1 \sum_{j=1}^{P} \sum_{k=1}^{P} C_{jk} \cos[(\bar{k}_j - \bar{k}_k) \cdot \bar{r} + \phi_j - \phi_k]$$

$$+ C_1 \sum_{n=P+1}^{M} \sum_{m=P+1}^{M} C_{nm} \cos[(\bar{k}_n - \bar{k}_m) \cdot \bar{r} + \phi_n - \phi_m]$$

$$+ 2C_1 \sum_{j=1}^{P} \sum_{n=P+1}^{M} C_{jn} \cos[(\bar{k}_j - \bar{k}_n) \cdot \bar{r} + \phi_j - \phi_n]$$

$$+ \text{ higher order terms.} \tag{3.2}$$

Since we will be illuminating the hologram with the reference beam (or a beam very similar to it) we shall be primarily interested in those terms in equation 3.2 which correspond to interference between reference beam waves and signal beam waves. These terms are those included in the third double sum in equation 3.2. The first and second double sums in equation 3.2 correspond to interference between the various reference beam waves and the various signal beam waves, respectively. We will assume that the geometry of the problem is such that the diffracted waves produced when these periodicities are illuminated with the reference beam are either separated in angle from the reconstructed signal beam or else are sufficiently weak (due to unfavorable illumination conditions) such that they may be neglected.

Thus, under these conditions, we shall write

$$D = D_v + D_r \tag{3.3}$$

where D_v includes those terms which are significant in the reconstruction process when the illumination beam is the reference beam and hence is given by

$$D_v = 2C_1 \sum_{j=1}^{P} \sum_{n=P+1}^{M} C_{jn} \cos[(\bar{k}_j - \bar{k}_n)\cdot\bar{r} + \phi_j - \phi_n] \quad , \tag{3.4}$$

and D_r represents the remaining terms.

3.2.2 Reconstruction of the Signal Beam

We are interested in the case where the illumination beam is the reference beam, and is thus given by

$$\bar{E}_r = \sum_{j=1}^{P} \bar{E}_j \, e^{i(\bar{k}_j \cdot \bar{r} + \phi_j)} \quad . \tag{3.5}$$

We will consider only one of the two first-order diffracted waves produced when each periodicity term in equation 3.4 is illuminated by one of the reference beam waves of equation 3.5. These will correspond to the virtual image. The other first-order diffracted waves correspond to the real image which is assumed to be separated from the virtual image as well as being considerably weaker, and thus will be neglected.

Hence there will be $P^2(M-P)$ diffracted waves which must be considered. $P(M-P)$ of these waves contribute to the reconstruction of the signal beam, and the remaining $P(P-1)(M-P)$ waves constitute a background noise.

Let us consider those waves that contribute to the reconstruction of the signal beam, leaving the others to be considered in detail in Section 3.2.4. Thus, let us consider the jth illumination wave in equation 3.5,

$$\bar{E}_{r_j} = \bar{E}_j \, e^{i(\bar{k}_j \cdot \bar{r} + \phi_j)} \quad . \tag{3.6}$$

We recall from Chapter One that when the illumination wave has the same $\bar{k}$ and ϕ as one of the two initial waves which yielded the

interference term under consideration, then one of the two first-order diffracted waves from this periodicity has its propagation vector $\bar{k}$ and phase ϕ equal to that of the other wave (except for a constant phase factor of $-\pi$). Thus the terms

$$2C_1 \sum_{n=P+1}^{M} C_{jn} \cos[(\bar{k}_j - \bar{k}_n)\cdot\bar{r} + \phi_j - \phi_n]$$

yield the diffracted waves

$$\bar{E}_{d_j} = 2C_1 \sum_{n=P+1}^{M} \bar{F}_{jn}(\bar{E}_j\cdot\bar{E}_n)\, e^{i(\bar{k}_n\cdot\bar{r} + \phi_n - \pi)} \tag{3.7}$$

when the illumination wave is given by equation 3.6. In writing equation 3.7 we replaced C_{jn} by $\bar{E}_j\cdot\bar{E}_n$ (equation 1.9b) and made use of the factor $\bar{F}$ defined in Chapter One, and in addition we neglected the diffracted waves corresponding to the real image, as discussed previously. If we now consider all the illumination waves as given by equation 3.5, then we obtain

$$\bar{E}_d = \sum_{j=1}^{P} \bar{E}_{dj} \tag{3.8}$$

or

$$\bar{E}_d = 2C_1 \sum_{j=1}^{P} \sum_{n=P+1}^{M} \bar{F}_{jn}(\bar{E}_j\cdot\bar{E}_n)\, e^{i(\bar{k}_n\cdot\bar{r} + \phi_n - \pi)} \quad . \tag{3.9}$$

We shall assume that

$$2C_1 \bar{F}_{jn}(\bar{E}_j\cdot\bar{E}_n) \tilde{=} CE_j\bar{E}_n \tag{3.10}$$

and hence equation 3.9 becomes

$$\bar{E}_d = C \sum_{j=1}^{P} \sum_{n=P+1}^{M} \bar{E}_n E_j \, e^{i(\bar{k}_n \cdot \bar{r} + \phi_n - \pi)} \quad . \tag{3.11}$$

Upon comparing equations 3.11 and 3.1 we see that equation 3.11 represents a reconstruction of the signal beam, there being P terms contributing to each wave in the signal beam.

3.2.3 Translation Sensitivity

Equation 3.11 shows that the reconstruction of each signal wave front consists of a superposition of terms, which add in phase provided that the hologram film plate is illuminated by the "reference beam" as given by equation 3.5. In general this requires that the hologram be repositioned in exactly the same place where it was during the exposure of the hologram, and that the hologram setup remain unchanged. If this repositioning is not done accurately, the result will be the appearance of a phase factor $e^{i\phi_j}$ under the summation sign in equation 3.11. This can, as was observed by van Heerden (13), result in the disappearance of the reconstruction, as the sum may then average out to zero.

To put these statements on a more quantitative basis, let us consider a translation of the developed hologram film plate by an amount $\bar{r}_o$ in the x',y' plane. We shall consider the reconstruction of one of the original signal waves, which is of the form

$$\bar{E}_m \, e^{i(\bar{k}_m \cdot \bar{r} + \phi_m)} \, e^{-i\omega t} \quad . \tag{3.12}$$

In the absence of any translation, the jth wavefront of the illumination reference beam interacts with the periodicity term

$$C_{jm} \cos[(\bar{k}_j - \bar{k}_m) \cdot \bar{r} + \phi_j - \phi_m] \tag{3.13}$$

to produce a diffracted wave of the form of equation 3.12. The effect of the translation of the film plate is to make the phase factor ϕ_m a function of j :

$$\phi_m \rightarrow \phi_{mj} \quad . \tag{3.14}$$

The phase factor ϕ_{mj} may be specified by requiring, as before, that the interference pattern generated by the transmitted portion of the illumination wave and the diffracted wave of interest match up with the recorded interference pattern corresponding to equation 3.13. This is done in Appendix III. A much easier way of specifying the phase factor ϕ_{mj} can be found by noting that a translation of the film plate is equivalent to an equal and opposite translation of both the source and the observer (or detector). Since we are considering each of the wavefronts to be locally a plane wave, the specification of ϕ_{jm} at any particular point on the film plate reduces to the problem considered by van Heerden (13), who points out that a displacement of the source by an amount $\bar{r}_o$ introduces a phase shift of $\bar{k} \cdot \bar{r}_o$ to a plane wave

with propagation vector $\bar{k}$. Since we are considering the source and observer to be fixed, the total phase shift the observer sees in the diffracted wave $\bar{k}_m$ is

$$\phi_{jm} = (\bar{k}_j - \bar{k}_m)\cdot\bar{r}_o \quad , \tag{3.15}$$

where $\bar{k}_j$ is the wave vector of the illumination wave and $\bar{r}_o$ is the displacement of the film plate. Thus, equation 3.11 becomes

$$\bar{E}_d = C \sum_{j=1}^{P} \sum_{m=P+1}^{M} E_j \, e^{i(\bar{k}_j - \bar{k}_m)\cdot\bar{r}_o} \, \bar{E}_m \, e^{i(\bar{k}_m\cdot\bar{r} + \phi_m - \pi - \omega t)} \quad . \tag{3.16}$$

We observe that the factor

$$E_j \, e^{i\bar{k}_j\cdot\bar{r}_o}$$

can be removed from under the summation over m , and thus each of the reconstructed signal beam waves is multiplied by the same factor η , where

$$\eta = \sum_{j=1}^{P} E_j \, e^{i\bar{k}_j\cdot\bar{r}_o} \quad . \tag{3.17}$$

It is clear that if the E_j are essentially equal and if the $\bar{k}_j$ vary sufficiently such that $\bar{k}_j\cdot\bar{r}_o$ ranges over $0 \to 2\pi$, then η will be essentially zero and we will have no reconstruction.

We should bear in mind that $\bar{r}_o$ is a constant vector and that the $\bar{k}_j$ may be functions of position, with the result that η

may vary over the film plate. Let us consider the case where

$$\bar{r}_o = \Delta x' \, \bar{e}_{x'} \quad , \tag{3.18}$$

then

$$\eta = \sum_{j=1}^{P} E_j \, e^{i\frac{2\pi}{\lambda} \ell' \Delta x'} \quad . \tag{3.19}$$

An estimate of the translation $\Delta x'_v$ required to make $\eta \to 0$ is found by assuming that the E_j are equal and the direction cosines ℓ' are uniformly distributed over the range ℓ'_{min} to ℓ'_{max}. Then

$$\Delta x'_v \simeq \frac{\lambda}{\ell'_{max} - \ell'_{min}} \quad . \tag{3.20}$$

We will have occasion to apply the above equation as well as equation 3.19 in Section 3.3, and we shall see that these equations yield predictions that are in good agreement with experimental results.

3.2.4 Background Noise

We view as background noise all diffracted waves other than those corresponding to a reconstruction of the signal beam as given by equation 3.11, which are not separated in angle from the signal beam. By separated in angle we mean having a propagation vector $\bar{k}$ whose direction is significantly different from those of the signal waves. We will retain the assumption made in Section 3.2.2 that the geometry of the experiment is such that we can neglect the waves diffracted from

all periodicities associated with the terms D_r in equation 3.3, either because they are separated in angle from the signal beam or because their amplitude is sufficiently small.

This leaves the periodicity terms given in equation 3.4. We have seen that these terms give rise to a reconstruction of the original signal beam when the hologram is illuminated with the reference beam. There are $P(M-P)$ periodicity terms in equation 3.4, where P and $M-P$ are the number of waves in the reference and signal beams, respectively. When the hologram is illuminated with the P waves of the reference beam, each periodicity will yield P diffracted waves (we are only considering one of the two first order diffracted waves), but only one of these P waves contributes to the reconstruction of the signal beam. The other $P-1$ waves constitute a background noise, and thus the ratio of signal terms to noise terms is

$$\frac{N_s}{N_n} = \frac{1}{P-1} \quad . \tag{3.21}$$

The noise waves can be computed in a straightforward manner in the same way that the signal field is computed, by using equations 1.16, 1.26, 1.28 and 1.34.

The fact that the reconstructed signal beam is accompanied by a background noise can also be seen from the transmittance approach. We recall that equation 2.10 specified the transmitted field in the $z' = 0^+$ plane, in terms of the exposing and illumination fields. In applying this equation, we shall find it convenient to write the

reference beam U_1 and signal beam U_2 in the form of a sum of plane waves. Thus we write

$$U_1 = \sum_{j=1}^{P} A_j \, e^{i\psi_j} \tag{3.22}$$

and

$$U_2 = \sum_{k=P+1}^{M} B_k \, e^{i\psi_k} \quad , \tag{3.23}$$

where the A_j and B_k are real constants and the ψ_j and ψ_k are of the form

$$\psi = \ell' x + m' y + \emptyset \quad . \tag{3.24}$$

Since we are interested in the case where the illumination beam is the reference beam, we set

$$U_3 = U_1 \quad . \tag{3.25}$$

We recall from Section 2.2.3 that an arbitrary field can be represented as a continuous spectrum of plane waves. We can view our choice of writing the reference beam and signal beam as discrete sums of plane waves as either a restriction to fields of this form or as an approximation to the more general case.

Now according to equation 2.10, the transmitted field U_T is given by

$$U_T = U_1(U_1 U_1^* + U_2 U_2^*) + U_1^2 U_2^* + U_1 U_1^* U_2 \quad . \tag{3.26}$$

We are interested in the last term in the above equation, which corresponds to the reconstruction of the signal beam. Now

$$U_1 U_1^* U_2 = \sum_{\substack{j=1 \\ k=1}}^{P} A_j A_k \, e^{i(\psi_j - \psi_k)} \sum_{\ell=P+1}^{M} B_\ell \, e^{i\psi_\ell} \tag{3.27}$$

and we see that if the signal beam U_2 is to be reconstructed without distortion or background noise we must have $P = 1$, which corresponds to the case where we have a single plane wave for the reference beam. We observe that we can write equation 3.27 in the form

$$U_1 U_1^* U_2 = \sum_{\substack{j=1 \\ k=1 \\ j\neq k}}^{P} A_j A_k \, e^{i(\psi_j - \psi_k)} \sum_{\ell=P+1}^{M} B_\ell \, e^{i\psi_\ell} + \sum_{j=1}^{P} A_j^2 \sum_{\ell=P+1}^{M} B_\ell \, e^{i\psi_\ell} . \tag{3.28}$$

The first group of terms corresponds to the background noise mentioned previously, while the second group corresponds to the reconstruction of the signal beam.

It is of interest to compare the power in the signal beam to that in the noise beam. A convenient measure is the signal to noise ratio S/N, which is seen from equation 3.28 to be of the form (the factors $U_2 U_2^*$ cancelling)

$$S/N = \frac{\left(\sum_{j=1}^{P} A_j^2 \right)^2}{\left(\sum_{\substack{j=1 \\ k=1 \\ j\neq k}}^{P} A_j A_k \, e^{i(\psi_j - \psi_k)} \right) \left(\sum_{\substack{j=1 \\ k=1 \\ j\neq k}}^{P} A_j A_k \, e^{-i(\psi_j - \psi_k)} \right)} . \tag{3.29}$$

We observe that we can write the denominator N in the form

$$N = \sum_{\substack{jk\ell m \\ j\neq k \\ \ell\neq m}}^{P} A_j A_k A_\ell A_m \, e^{i(\psi_j - \psi_k - \psi_\ell + \psi_m)} \quad . \tag{3.30}$$

It is of interest to write out separately those terms in equation 3.30 for which the exponent is zero. This occurs when the $j = \ell$ and $k = m$ (the cases where $j = k$ and $\ell = m$ not being allowed). There are P^2 such terms, and we can write N in the form

$$N = \left(\sum_{j=1}^{P} A_j^2 \right)^2 + \sum_{\substack{jk\ell m \\ j\neq k \\ j\neq \ell \\ \ell\neq m \\ 1\neq m}}^{P} A_j A_k A_\ell A_m \, e^{i(\psi_j - \psi_k - \psi_\ell + \psi_m)} \quad . \tag{3.31}$$

We recall from equation 3.24 that the ψ_i are functions of position on the film plate. It is clear that if the propagation vectors $\bar{k}_i$ and or phase factors ϕ_i have sufficient variations that at any given observation point (x,y) the ψ_j will vary over the range $0 \to 2\pi$, and hence the contribution of the second factor in equation 3.31 can be neglected, as it will essentially average out to zero. In this case the signal to noise ratio is unity:

$$S/N = 1 \quad . \tag{3.32}$$

This can be interpreted in a simple way by recalling that there are P waves that contribute the reconstruction of each of the signal beam waves. These waves add in phase so that the power in each of these

signal waves goes as P^2 . On the other hand there are $P(P-1)$ corresponding noise waves, but these don't add in phase and hence their power goes as $P(P-1)$, with the result that the signal to noise ration goes as

$$\frac{P^2}{P(P-1)} \to 1$$

for large P .

3.3 Translation Sensitivity Experiment

In this section we shall describe an experiment where the total power in the reconstructed signal beam is measured as a function of the error in repositioning the hologram plate. The analysis of Section 3.2.3 is applied to compute the quantity being measured and experimental and computed values are compared, and found to be in excellent agreement.

3.3.1 Description and Analysis of the Experiment

The experiment consists of taking a hologram using a multiple wave reference beam, and then measuring the total power in the reconstructed signal beam as a function of the displacement of the developed film plate from the position it occupied during the exposure of the hologram. For reasons of experimental convenience, the signal beam was taken to be a single converging wave. This corresponds to the special case of $M = P + 1$ in equation 3.1. It follows from equation 3.16 that the reconstruction of each wavefront in the signal beam can be

considered separately with respect to the effects of plate translation, and thus there is no loss of generality incurred by this choice of the signal beam.

The reference beam is produced by illuminating a diffusely reflecting surface with a collimated laser beam, the scattered light forming the reference beam. The reference beam is thus of the form given in equation 3.1.

The actual geometry used in the experiment is shown in Figure 3.1. The converging wave which forms the signal beam is incident on the film plate at an angle of incidence of 24°, and comes to a focus behind the film plate at $(x',y',z') = (-11", 0, -24.7")$. It exposes an elliptically shaped area centered at the origin. A converging beam was used because it provides a convenient means for the measurement of the total power in the reconstruction of the signal beam, and for the discrimination against background noise.

The reference beam is provided by the light scattered from a magnesium oxide powder layer sandwiched between two microscope slides. The magnesium oxide layer is illuminated by the laser beam (which was divided into two beams by a beam splitter) which is passed through a 3/32" diameter aperture before striking the oxide layer at an angle of 26° 30' as shown in Figure 3.1. The reference beam can thus be assumed to arise from a large number of point sources distributed over an elliptically shaped area in the $x' = -1/2$ inch plane. The amplitude of each point source depends on the variation in amplitude across the cross section of the illuminating beam, and will be dealt with later when

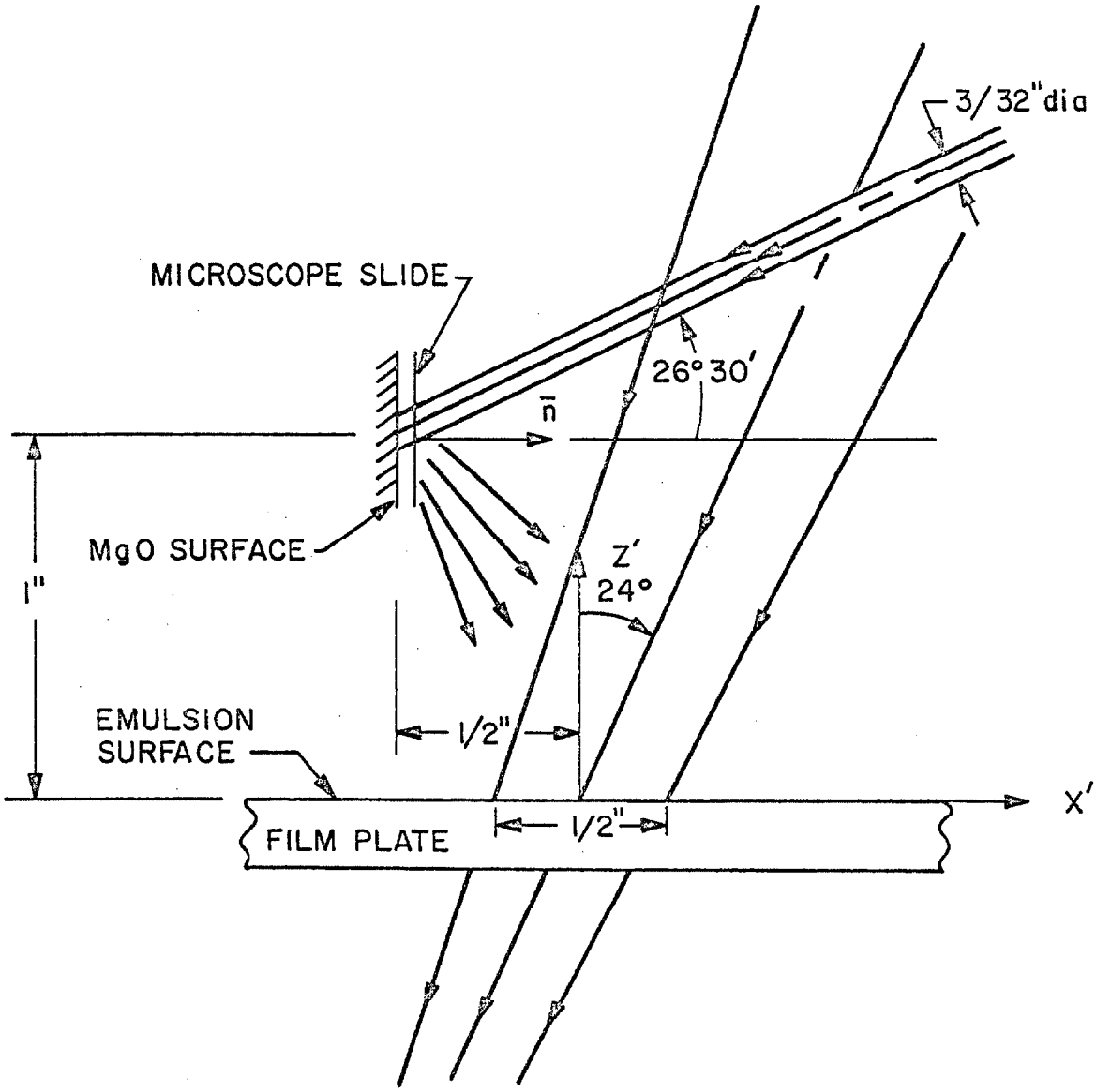


FIGURE 3.1 GEOMETRY OF TRANSLATION SENSITIVITY EXPERIMENT

required in the calculations. The point sources have random phases.

The light scattered by the reference beam source points will illuminate most of the film plate for $x' > -1/2$, but the only area of interest is that area illuminated by both signal and reference beams since this is the only area which will yield a reconstruction. This consists of the elliptically shaped area which is illuminated by the signal beam. The displacement $\Delta x_v'$ in the x' direction that the developed hologram film plate must be moved in order to have the reconstruction vanish is a function of the observation point in the hologram area. That is, if we were to observe the reconstruction through a small aperture placed at some observation point $(x',y',0)$ in the hologram area, the value of $\Delta x_v'$ will vary with $(x',y',0)$. In this particular problem it is the total power in the reconstruction that is measured, which is the sum of the power passing through each element of area of the region which constitutes the hologram.

An estimate of the $\Delta x_v'$ can be obtained by neglecting the variation of the y' direction and applying equation 3.20. Thus, to obtain an estimate of $\Delta x_v'$ (max) , equation 3.20 is applied to the case where the source points which generate the reference beam all lie along the line between (-.5, 0, .948) and (-.5, 0, 1.052). The observation point which yields the maximum value of $\Delta x_v'$ is at (-.25, 0, 0).

The angles between the x' axis and the lines which join the two extremities of the source line and the observation point are

$$\theta_1 = \tan^{-1} \left(\frac{.948}{.25}\right) = 75^\circ\ 14' \qquad (3.33)$$

$$\theta_2 = \tan^{-1}\left(\frac{1.052}{.25}\right) = 76^\circ\ 38' \tag{3.34}$$

and hence

$$\ell' \text{ max} = \cos(75^\circ\ 14') = .25488 \tag{3.35}$$

$$\ell' \text{ min} = \cos(76^\circ\ 38') = .23118 \quad . \tag{3.36}$$

Using the above values in equation 3.20, with $\lambda = .6328\mu$, we find as an estimate for $\Delta x'_v$

$$\Delta x'_v \simeq 26.7 \text{ microns} \quad . \tag{3.37}$$

As will be seen later, this estimate is in quite good agreement with experimental measurements.

The quantity which is measured in the experiment is the total power P_n in the reconstructed signal beam as a function of the displacement of the film plate from its original position, normalized with respect to the power obtained when the displacement of the film plate is zero. P_n is given by

$$P_n = \frac{\int_{x'}\int_{y'} \bar{E}_d(\Delta x')\cdot\bar{E}_d^*(\Delta x')dx'dy'}{\int_{x'}\int_{y'} \bar{E}_d(0)\cdot\bar{E}_d^*(0)dx'dy'} \tag{3.38}$$

where the integration over the $x'y'$ plane extends over the area of the film plate where both the signal and reference beam wavefronts

existed during the exposure of the hologram.

The diffracted field $\bar{E}_d$ is given by equation 3.16, with

$$\bar{r}_o = \Delta x' \bar{e}_{x'} \quad . \tag{3.39}$$

The factor η, defined by equation 3.17, then becomes

$$\eta(\Delta x') = \sum_{j=1}^{P} E_j \, e^{i\frac{2\pi}{\lambda} \ell_j' \Delta x'} \quad . \tag{3.40}$$

Recalling that we have only a single signal beam wave, we see that $M = P + 1$ in equation 3.16 (only one term in the summation over m), and hence, using equations 3.16 and 3.40, equation 3.38 becomes

$$P_n(\Delta x') = \frac{\int_{x'} \int_{y'} \bar{E}_m \cdot \bar{E}_m \, \eta(\Delta x') \, \eta^* (\Delta x') dx' dy'}{\int_{x'} \int_{y'} \bar{E}_m \cdot \bar{E}_m \, \eta(0) \, \eta^* (0) dx' dy'} \quad . \tag{3.41}$$

Observation of the developed hologram plates has shown that the amplitude of the signal wave $\bar{E}_m$ is essentially constant over the area of integration, and thus it is a reasonable approximation to remove $\bar{E}_m \cdot \bar{E}_m$ from under the integral sign in both the numerator and denominator of equation 3.41. Thus

$$P_n(\Delta x') = \frac{\int_{x'} \int_{y'} \eta(\Delta x') \, \eta^* (\Delta x') dx' dy'}{\int_{x'} \int_{y'} \eta(0) \, \eta^* (0) dx' dy'} \quad . \tag{3.42}$$

The quantity E_j in the expression for η given by equation 3.40 is the amplitude at the point $(x', y', 0)$ of the field scattered by the jth scattering point on the magnesium oxide surface. E_j will be assumed to be given by

$$E_j = \frac{CA_j}{r_j} \quad , \tag{3.43}$$

where A_j is the amplitude of the wave illuminating the scattering surface at the jth scattering point and r_j is the distance from the scattering point to the point x', y' on the emulsion surface. C is a constant which we will assume to have approximately the same value for all scattering points and thus will cancel out in equation 3.42. The illumination wave is an unfocused laser beam passed through a 3/32 diameter aperature. The variation of A_j across the laser beam depends on the geometry of the laser cavity (mirror curvatures and spacing), the transverse mode structure, and the distance from the output mirror. The laser was operated in the lowest order transverse mode, and thus has an amplitude variation which is Gaussian (truncated by the aperture). Thus, the variation with r (defined normal to the laser beam axis) is

$$A(r) = E_o \, e^{-\alpha r^2} \quad . \tag{3.44}$$

The constant α was determined by fitting the above curve to measured data on the variation of the amplitude with r. A value of

$$\alpha = 1 \text{ mm}^{-2} = 645.16 \text{ (in)}^{-2} \tag{3.45}$$

was found to fit the measured data well. An examination of the geometry of the experiment shows that the amplitude of the illumination wave at the point (y_j, z_j) on the magnesium oxide surface is found by setting r^2 in equation 3.44 equal to

$$r^2 = y_j^2 - (z_j - 1)^2 \cos^2(26^\circ\ 30') \quad . \tag{3.46}$$

The distance r_j from the scattering point $(-1/2, y_j, z_j)$ to the point $(x', y', 0)$ on the emulsion surface is

$$r_j = [(x' + 1/2)^2 + (y' - y_j)^2 + z_j^2]^{1/2} \quad . \tag{3.47}$$

Equations 3.43 to 3.47 specify E_j as a function of the coordinates of the jth scattering point and the coordinates of the observation point on the hologram. The remaining quantity that needs to be specified in order to determine η is ℓ_j', the x' direction cosine of the propagation vector $\bar{k}_j$ of the wave scattered by the jth scattering point, evaluated at the observation point $(x', y', 0)$. We assume $\bar{k}_j$ points in the direction from $(-1/2, y_j, z_j)$ to $(x', y', 0)$, and thus

$$\ell_j' = -\frac{x' + 1/2}{r_j} \quad . \tag{3.48}$$

The application of equations 3.40 and 3.42 to 3.48 permit the computation of the normalized power diffracted into the virtual image as a function of the translation distance $\Delta x'$. This was done, with the calculations being done numerically with the aid of a digital computer. The results of this computation will be presented in Section

3.3.3, which follows the next section which deals with the details of the experiment.

3.3.2 Experimental Details

The basic idea of the experiment was to take a hologram of a simple wavefront, with the reference beam being of the form of a field produced by a large number of point sources. The hologram was then repositioned and the total power in the reconstructed signal beam wavefront was measured as a function of translation of the hologram plate from its initial position. The experimental apparatus which was used to perform the experiment is shown in Figure 3.2.

The apparatus on the iron surface plate is that which was used to expose and illuminate the hologram. The source used was a helium-neon laser which has a power output of about three milliwatts when operated in the lowest order transverse mode. A camera shutter was used to control the exposure time, which was of the order of ten seconds. A beam splitter provides two beams, one of which is passed through an optical system which performs a low pass spatial filtering operation and then produces a converging beam which comes to a focus about 27 inches behind the film plate. The other beam is directed by a series of mirrors (which are positioned so as to make the path lengths of the two beams approximately equal) positioned so that the beam falls on a diffusely reflecting surface oriented so that the scattered light reaches the region of the film plate which is illuminated by the converging beam. The diffusely reflecting surface is a layer of magnesium

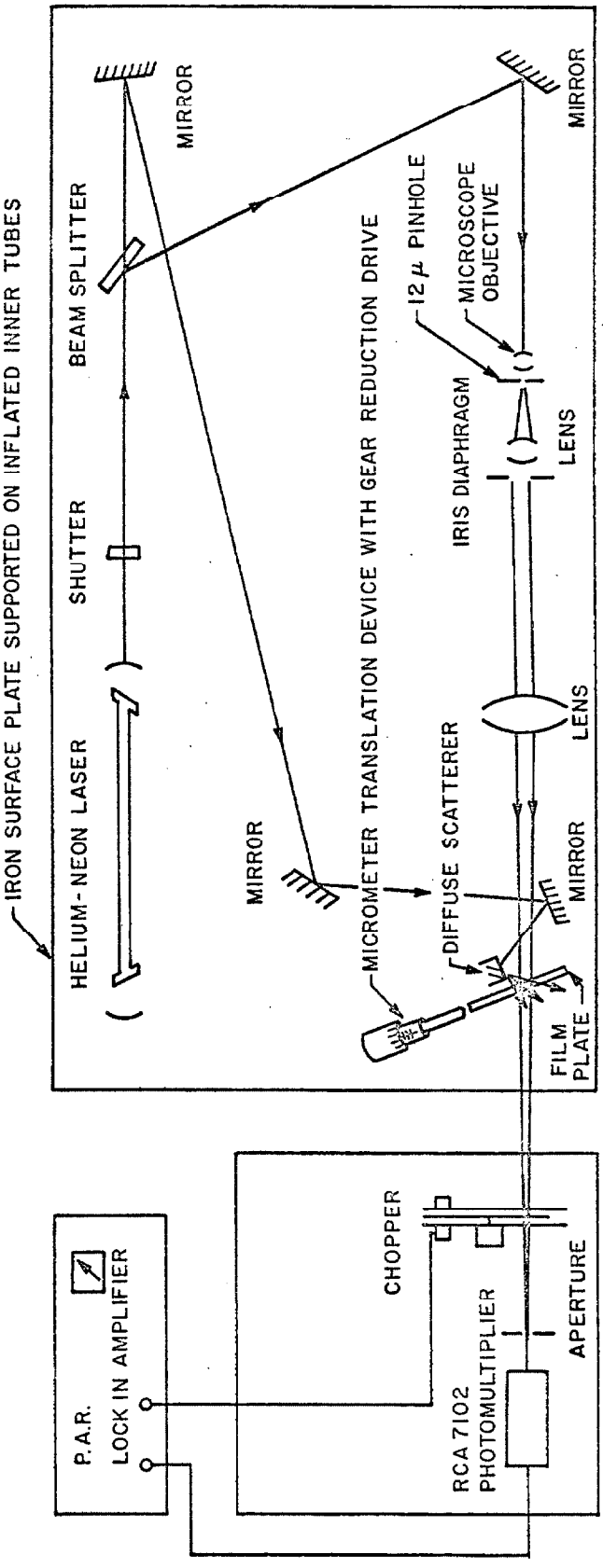


FIGURE 3.2 EXPERIMENTAL APPARATUS - TRANSLATION SENSITIVITY EXPERIMENT

oxide powder pressed between two microscope slides. The angle of incidence of the incident laser beam is adjusted so the reflected beams from the glass air interfaces do not strike the region of the film plate which constitutes the hologram area. The film plate (Kodak 649-f 4x5x.25 microflat plate) was held in a specially designed film plate holder which permitted accurate repositioning and translation of the film plate. Translations of the order of a micron or less were possible through the use of a micrometer drive mechanism with a 187:1 gear reduction attachment.

The apparatus on the smaller table (see Figure 3.2) is the system used to measure the power in the reconstruction of the converging wavefront. The converging beam is chopped at about 10^3 CPS, passed through a small aperture located at about the focal point, and then detected by a photomultiplier. The aperture has the function of discriminating against background light scattered from the film plate which is not part of the reconstructed signal beam. The output of the photomultiplier is measured using a lock-in amplifier, which makes use of a reference signal generated by the chopper.

During exposure of the holograms the iron surface plate was "floated " to uncouple the apparatus from building vibrations. In addition, acoustical shielding was used in order to reduce any acoustically excited vibrations in the critical elements of the hologram taking apparatus (beamsplitter, mirrors and film plate).

3.3.3 Discussion

The results of the numerical calculations referred to at the end of Section 3.3.1 and the experimental results are shown in Figure 3.3. The agreement between experiment and theory is excellent, considering the approximations made in the analysis and numerical computations as well as the difficulties involved with the experiment. Furthermore, we observe that the estimate of $\Delta x_v'$ (the translation distance required to make the reconstruction vanish) given by equation 3.20 of 26.7 microns (equation 3.37) is quite accurate.

There were two major areas of difficulty that were encountered in carrying out the experiment, namely low power levels in the reconstruction of the signal beam and difficulties in obtaining accurate translation motions of the order of a micron or less. The low power levels were due mainly to the fact that in this experiment we are forced to illuminate the hologram with the reference beam which was used to expose the hologram, which was quite weak due to the relatively low power of the laser, the inefficiency of the scattering step, and the $1/r^2$ loss between the scattering area and the film plate. This latter loss was minimized by placing the magnesium oxide scattering layer as close to the film plate as was mechanically possible. In addition, the relative power levels of the signal and reference beam as well as the exposure time were adjusted to obtain maximum efficiency from the hologram. This was accomplished by an "educated" trial and error procedure, making use of the results of an experimental study of the holographic diffraction grating, which is discussed in Chapter Four.

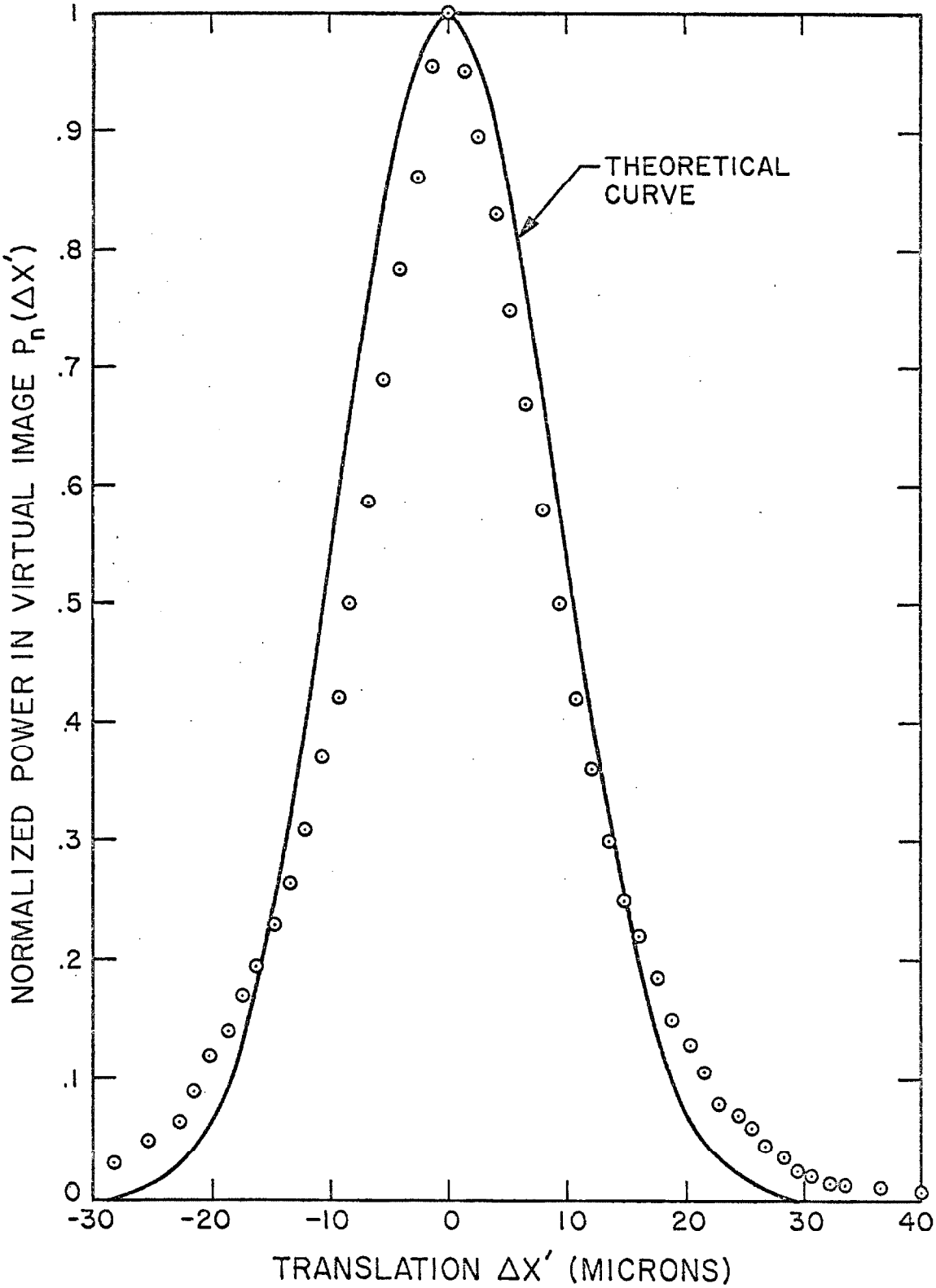


FIGURE 3.3 EXPERIMENTAL AND ANALYTICAL RESULTS

The most efficient hologram produced a reconstruction with a power level of the order of 2×10^{-9} watts, which allowed a measurement of $P_n(\Delta x')$ down to a value of about 0.05 before the sensitivity limit of our detection system was reached (about 10^{-10} watts). Lower values were measured but the noise in the detection system became objectionable.

The translation of the film plate was accomplished by pushing the film plate sideways in a special holder with a micrometer drive, which in turn was driven by a 187:1 gear reduction device. Once the backlash of the device was taken up there was no difficulty in moving the rod pushing on the film plate by very small increments. Trouble was experienced, however, with the movement of the film plates when thin .040" x 4" x 5" film plates were used, and satisfactory results were obtained only when thicker and heavier .250" x 4" x 5" microflat plates were used.

It was observed that a translation of the hologram plate resulted in a corresponding translation of the reconstructed beam. This required a corresponding movement of the small aperture used to block out the background noise. A similar motion of the focused beam was observed in earlier experiments (not reported in this thesis) dealing with holographic lenses, in which a hologram was taken of a converging wave, using a single plane wave for a reference beam. In that case the reconstructed converging wavefront remained in register with the hologram plate, as could be expected.

3.4 Fourier Transform Holograms with Multiple Wave Reference Beams

We recall that the fundamental process involved in holography is the interference of two fields in the exposure of the film plate. One of these fields is arbitrarily referred to as the reference beam and the other as the signal beam. By convention we refer to the field which is used to illuminate the hologram as the reference beam, and the field that is to be reconstructed (virtual image) as the signal beam. We recall that the most common situation is where the reference beam is a single wave and the signal beam is some complex field such as the light scattered by a diffusely reflecting object. It is now of interest to examine a particular case where the reference beam is the complicated field and the signal beam is the plane wave. To be more specific, we shall consider the case of the Fourier transform hologram, where we adopt the point of view that the field due to the transparency is the reference beam, and the plane wave is the signal beam.

This type of hologram is of particular interest, since it is in fact the complex part of the "matched filter" used by Vander Lugt (41,42) to perform signal detection by complex spatial filtering. Similarly, Gabor utilizes a variation of this type of hologram in a proposed character recognition system (43). We shall be interested in examining the translational sensitivity of Fourier transform holograms, both with respect to translations of the transparency and translations of the hologram itself. We will also be interested in considering the background noise, and will examine both the translational sensitivity and background noise when the transparency is diffusely illuminated.

Before considering Fourier transform holograms with multiple wave reference beams, however, we shall review the case where the reference beam consists of a single plane wave.

3.4.1 Review of Fourier Transform Holograms

By a Fourier Transform hologram we mean a hologram of a two-dimensional transparency, taken under such conditions that the exposing field at the film plate due to the transparency is of the form of a Fourier transform of the field transmitted by the transparency. This can be accomplished by the use of a converging lens, when the transparency and film plate are located in the front and back focal planes of the lens, respectively (the Fourier transform relationship between the amplitudes in the front and back focal planes of a lens is reviewed in Appendix V). Other experimental configurations are possible (34), and we shall briefly review Stroke's method of "lensless" Fourier transform holography at the end of this section.

The treatment which we shall give here will be a brief review of this well-known aspect of holography (2), (34), (38) and we shall limit our treatment to a demonstration of the formation of an image of the object transparency when a lens is used in the reconstruction process. A further discussion of certain aspects of Fourier transform holography will be given in the following sections.

Let us begin by considering the case where the object transparency is placed in the front focal plane of a converging lens and illuminated by a plane wave of unit amplitude at normal incidence. If

$\tau(x_1, y_1)$ is the amplitude transmittance of the transparency, then the amplitude distribution in the back focal plane of the lens is shown in Appendix V to be given by

$$E_S(x_2,y_2) = \frac{-i}{\lambda f} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau(x_1,y_1)\, e^{-i \frac{2\pi}{\lambda f} (x_1x_2 + y_1y_2)}\, dx_1 dy_1 \quad . \tag{3.49}$$

We shall take the reference beam to be an off-axis plane wave of unit amplitude which has the form (in the x_2,y_2 plane)

$$E_R(x_2,y_2) = e^{i \frac{2\pi}{\lambda} (\ell x_2 + m y_2)} \quad . \tag{3.50}$$

Such a reference beam could be produced by a point source suitably located in the plane of the object transparency, or by by-passing the lens entirely with a collimated beam. The hologram is formed by placing a film plate in the back focal plane of the lens, and illuminating it with the fields E_S and E_R. For simplicity we shall assume that the transmittance τ_H of the developed film plate, or hologram, is simply equal to the intensity of the exposing fields. Thus

$$\tau_H = E_R E_R^* + E_S E_S^* + E_R E_S^* + E_S E_R^* \quad . \tag{3.51}$$

In the reconstruction process, the hologram is placed in the front focal plane of a similar lens and the reconstructed images are formed in the back focal plane. Thus, the amplitude of the field in the back focal plane (x_3, y_3) is given by

$$E(x_3, y_3) = -\frac{i}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_R\, \tau_H\, e^{-i\frac{2\pi}{\lambda f}(x_2x_3 + y_2y_3)}\, dx_2dy_2 \quad . \tag{3.52}$$

The terms in the expression for τ_H as given by equation 3.51 which are of interest are $E_RE_S^*$ and $E_SE_R^*$, which correspond to the real and virtual images, respectively. Let us consider the virtual image. Substitution of $E_SE_R^*$ for τ_H in equation 3.52 yields, using equations 3.49 and 3.50,

$$E_v(x_3,y_3) = -(\frac{1}{\lambda f})^2 \iiiint \tau(x_1,y_1)\, e^{-i\frac{2\pi}{\lambda f}[x_2(x_1 + x_3) + y_2(y_1 + y_3)]} \times dx_1dy_1dx_2dy_2 \quad . \tag{3.53}$$

Integration over x_2,y_2 yields $(\lambda f)^2\, \delta(x_1 + x_3)\, \delta(y_1 + y_3)$ and hence

$$E_v(x_3, y_3) = -\tau(-x_3,-y_3) \quad . \tag{3.54}$$

Similarly, the amplitude distribution $E_r(x_3, y_3)$ due to the real image term $E_RE_S^*$ is given by

$$E_r(x_3, y_3) = (\frac{1}{\lambda f})^2 \iiiint \tau^*(x_1, y_1)\, e^{-i\frac{2\pi}{\lambda f}[x_2(-x_1 + x_3 - 2f\ell)]} \times e^{-i\frac{2\pi}{\lambda f}y_2(-y_1 + y_3 - 2fm)} dx_1dy_1dx_2dy_2 \tag{3.55}$$

and hence

$$E_r(x_3, y_3) = \tau^*(x_3 - 2f\ell, y_3 - 2fm) \quad . \tag{3.56}$$

Thus, both the real and virtual images yield a reconstruction of the transparency, and these are separated due to our choice of an off-axis reference beam. The zero order terms are likewise separated from the reconstruction being centered at $(f\ell, fm)$.

Let us now briefly review the technique of Stroke (38, 39, 40) whereby a Fourier transform hologram may be obtained without the use of a lens. The essential feature of this technique is the use of a point source located in the plane of the object for the reference beam. This results in the cancellation of certain quadratic phase factors with the result that the hologram obtained has a form similar to that of a Fourier transform hologram.

To put these statements on a more quantitative basis, let us assume that the transparency and film plate are located in parallel planes a distance z_o apart. The transparency with transmittance $\tau(\alpha, \beta)$ is illuminated with a plane wave of unit amplitude at normal incidence and this yields the field in the xy plane given by (Fresnel diffraction)

$$E_S(x,y) = -\frac{i}{\lambda z_o} e^{i\frac{2\pi z_o}{\lambda}} \iint \tau(\alpha, \beta)\, e^{\frac{i\pi}{\lambda z_o}[(x-\alpha)^2 + (y-\beta)^2]}\, d\alpha d\beta \quad . \tag{3.57}$$

Now suppose, as was suggested by Stroke, that the reference beam is provided by a point source located in the α,β plane. Taking the point source at the origin in this plane yields a field in the xy plane of

the form

$$E_R(x,y) = A_o \frac{e^{ikr}}{r} \tag{3.58}$$

or, taking $A_o = 1$,

$$E_R(x,y) \overset{\sim}{=} \frac{e}{z_o}^{i \frac{2\pi}{\lambda} (z_o + \frac{x^2 + y^2}{2z_o})} . \tag{3.59}$$

The resulting amplitude transmittance of the developed hologram plate is, using equations 3.51, 3.57 and 3.59

$$\tau_H(x,y) = E_R E_R^* + E_S E_S^* +$$

$$- \frac{i}{\lambda z_o^2} \iint \tau(\alpha,\beta) \, e^{\frac{i\pi}{\lambda z_o} (\alpha^2 + \beta^2)} \, e^{- \frac{2\pi i}{\lambda z_o} (\alpha x + \beta y)} \, d\alpha d\beta$$

$$+ \frac{i}{\lambda z_o^2} \iint \tau^*(\alpha,\beta) \, e^{- \frac{i\pi}{\lambda z_o} (\alpha^2 + \beta^2)} \, e^{\frac{2\pi i}{\lambda z_o} (\alpha x + \beta y)} \, d\alpha d\beta . \tag{3.60}$$

We observe that except for the factors

$$e^{\pm \frac{i\pi}{\lambda z_o} (\alpha^2 + \beta^2)} \tag{3.61}$$

equation 3.60 is of the same form as was obtained in the previous section, for the case of the Fourier transform hologram. Stroke (38) notes that the above phase factor merely makes the reconstructed object

appear as though it were recorded through a thin negative field lens, and that this can be compensated for in the reconstruction. Winthrop and Worthington (34), in their paper dealing with the Fresnel transform representation of holograms and hologram classification, refer to a hologram of the form of equation 3.60 as a quasi-Fourier transform hologram, and they discuss its imaging properties. Stroke (38) makes it clear that this arrangement is not limited to transparencies, but can also be used with diffusely reflecting objects, and presents experimental work in this area (40).

3.4.2 Translation Sensitivity - Displacement of the Transparency

In the previous section we reviewed Fourier transform holograms, and considered the case where the reference beam was taken to be a plane wave and the signal beam was the field due to the transparency. We now wish to turn this around, and consider the field due to the transparency as the reference beam and the plane wave as the signal beam. Thus

$$E_R(x_2,y_2) = -\frac{i}{\lambda f}\iint \tau(x_1,y_1)\, e^{-i\frac{2\pi}{\lambda f}(x_1x_2 + y_1y_2)}\, dx_1\, dy_1 \qquad (3.62)$$

and

$$E_S(x_2,y_2) = e^{i\frac{2\pi}{\lambda}(\ell x_2 + m y_2)} \qquad (3.63)$$

This re-labeling of the exposing fields, of course, doesn't change the hologram in any manner, but it is convenient because it allows us to

retain our designation of the reference beam as the field which is used to illuminate the hologram.

Thus, let us consider the case where a transparency having an amplitude transmittance $\tau(x_1, y_1)$ is placed in the front focal plane (x_1, y_1) of a converging lens of focal length f and illuminated with a plane wave at normal incidence. The film plate is located in the back focal plane of the lens (x_2, y_2) and illuminated by the field due to the transparency (the reference beam) and by an off-axis plane wave (the signal beam). As before, in the reconstruction process, the hologram is placed in the front focal plane (x_2, y_2) of a lens of focal length f and the reconstructed images are formed in the back focal plane of the lens (x_3, y_3). If the illumination is done with the reference beam, then the resulting amplitude distribution in the back focal plane is given by equation 3.52. That is,

$$E(x_3,y_3) = -\frac{i}{\lambda f}\iint E_R(E_R E_R^* + E_S E_S^*)\, e^{-i\frac{2\pi}{\lambda f}(x_2x_3 + y_2y_3)}\, dx_2dy_2 +$$

$$-\frac{i}{\lambda f}\iint E_R E_R^* E_S\, e^{-i\frac{2\pi}{\lambda f}(x_2x_3 + y_2y_3)}\, dx_2dy_2 +$$

$$-\frac{i}{\lambda f}\iint E_R E_R E_S^*\, e^{-i\frac{2\pi}{\lambda f}(x_2x_3 + y_2y_3)}\, dx_2dy_2 \quad . \qquad (3.64)$$

The last two terms are what is of interest, as they correspond to the virtual and real images, respectively. Let us consider the virtual image. Using equation 3.62 for the reference beam and equation 3.63

for the signal beam, the virtual image has an amplitude distribution in the (x_3, y_3) plane given by

$$E_v(x_3,y_3) = \left(\frac{i}{\lambda f}\right)^3 \int \cdots \int \tau(x_1,y_1)\, \tau^*(x_1',y_1')$$

$$\times e^{-i\frac{2\pi}{\lambda f}[x_2(x_1-x_1'+x_3-\ell f) + y_2(y_1-y_1'+y_3-mf)]}\, dx_1 dy_1 dx_1' dy_1' dx_2 dy_2 \ . \quad (3.65)$$

We see that integration over x_2 and y_2 yields $(\lambda f)^2\, \delta(x_1-x_1'+x_3-\ell f)$ $\delta(y_1-y_1'+y_3-mf)$ and hence, after integrating over x_1' and y_1', we obtain

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau(x_1,y_1)\, \tau^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)\, dx_1 dy_1 \ . \quad (3.66)$$

Similarly for the real image, we have

$$E_r(x_3,y_3) = -\left(\frac{i}{\lambda f}\right)^3 \int \cdots \int \tau(x_1,y_1)\, \tau(x_1',y_1')$$

$$\times e^{-i\frac{2\pi}{\lambda f}[x_2(x_1 + x_1' + x_3 + \ell f) + y_2(y_1 + y_1' + y_3 + mf)]}\, dx_1 dy_1 dx_1' dy_1' dx_2 dy_2 \quad (3.67)$$

which yields

$$E_r(x_3,y_3) = \frac{i}{\lambda f} \iint \tau(x_1,y_1)\, \tau(-x_1-x_3-\ell f,\, -y_1-y_3-mf)\, dx_1 dy_1 \ . \quad (3.68)$$

The above analysis is quite similar to that used by Vander Lugt (41), who has formulated these types of calculations in a convenient operational notation in a recent paper (44). An examination of

equations 3.66 and 3.68 shows the familiar result (38,41), that in a system of the type considered above, the formation of the virtual image involves a correlation operation, while the formation of the real image involves a convolution.

Let us now examine the effect of a translation of the transparency on the virtual image. Thus, we shall consider the case where the hologram has been exposed and developed, and then replaced in its original position, but where the transparency that was used to provide the reference beam during the exposure of the hologram is translated by an amount $\bar{r}_o$ in the x_1, y_1 plane from its original position, where

$$\bar{r}_o = \bar{e}_x \Delta x + \bar{e}_y \Delta y \quad . \tag{3.69}$$

The illumination field, provided by illuminating the displaced transparency with a plane wave of unit amplitude at normal incidence, is then given by equation 3.62 with $\tau(x_1,y_1)$ replaced by $\tau(x_1-\Delta x, y_1-\Delta y)$. Equation 3.65 then becomes

$$E_v(x_3,y_3) = \left(\frac{i}{\lambda f}\right)^3 \int \cdots \int \tau(x_1 - \Delta x,\, y_1 - \Delta y)\, \tau^*(x_1',\, y_1') \times e^{-i\frac{2\pi}{\lambda f}[x_2(x_1-x_1'+x_3-\ell f) + y_2(y_1-y_1'+y_3-mf)]}\, dx_1 dy_1 dx_1' dy_1' dx_2 dy_2 \tag{3.70}$$

which yields, after integrating over x_1, y_1, x_1' and y_1'

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau(x_1-\Delta x,\, y_1-\Delta y)\, \tau^*(x_1+x_3-\ell f,\, y_1+y_3-mf)\, dx_1 dy_1 \cdot \tag{3.71}$$

Making the change of variables $u = x_1 - \Delta x$ and $v = y_1 - \Delta y$, the above equation assumes the form

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau(u,v)\ \tau^*(u+x_3+\Delta x-\ell f,\ v+y_3+\Delta y-mf)\ du\ dv \qquad (3.72)$$

and it is seen that a translation of the transparency merely results in a similar translation of the virtual image in the (x_3,y_3) plane. An examination of equation 3.72 shows that this translation is equal to $-\bar{r}_o$, and thus the virtual image remains in register with the image of the displaced transparency in the x_3,y_3 plane, whose image is also displaced by $-\bar{r}_o$. This result is analogous to the prediction by van Heerden (12) that his "ghost image" remains in register with the image of the illuminating transparency, and is well known in the field of complex spatial filtering (41,44).

We have thus seen that except for a displacement in the (x_3,y_3) plane, the virtual image is not affected by a translation of the transparency. This can be explained in physical terms by noting that the field transmitted by the transparency can be expressed as a continuous distribution of plane waves. A translation of the transparency in the (x_1,y_1) plane doesn't alter the direction of any of these waves, and hence they are imaged at the same points in the (x_2,y_2) plane regardless of the translation. There is, however, a phase shift produced by the translation which accounts for the displacement of the virtual image. This phase shift, in the (x_2,y_2) plane, can be found by using equation V-14 (Appendix V), which gives the optical path

length between the points (x_1,y_1) and (x_2,y_2). What we are interested in is the change in optical path length as a function of (x_2,y_2) produced by a translation of the transparency or "source" in the x_1,y_1 plane. This will give us the phase shift $e^{i\phi}$ as a function of (x_2,y_2). We observe that

$$\phi = \frac{2\pi}{\lambda} [r(x_1 + \Delta x, y_1 + \Delta y, x_2,y_2) - r(x_1,y_1,x_2,y_2)] \quad (3.73)$$

where r is given by equation V-14 with $g = f$.

Thus, using equation V-14 we find

$$\phi = - \frac{2\pi}{\lambda f} (x_2 \Delta x + y_2 \Delta y) \quad (3.74)$$

and hence

$$e^{i\phi} = e^{-i \frac{2\pi}{\lambda f} (x_2 \Delta x + y_2 \Delta y)} , \quad (3.75)$$

which accounts for the displacement of the virtual image by $(\Delta x,\Delta y)$ in the (x_3,y_3) plane.

3.4.3 Translation Sensitivity - Displacement of the Hologram

Let us now consider the case where the transparency is left in its initial position, but the hologram plate is translated in the (x_2,y_2) plane by $\bar{r}_o$. As before, we shall examine the effect of the translation on the virtual image, in the (x_3,y_3) plane. In the absence of any translation of the hologram (i.e., when the hologram is replaced in exactly the position it was in during exposure) the

amplitude of the virtual image in the (x_3,y_3) plane is given by the second term in equation 3.64

$$E_v(x_3,y_3) = -\frac{i}{\lambda f}\iint E_R(x_2,y_2)\, E_R^*(x_2,y_2)\, E_S(x_2,y_2)\, e^{-i\frac{2\pi}{\lambda f}(x_2x_3+y_2y_3)}\, dx_2dy_2 \quad . \tag{3.76}$$

A translation of the hologram plate by $\bar{r}_o$ means that in the above equation

$$E_R^*(x_2,y_2)\, E_S(x_2,y_2) \rightarrow E_R^*(x_2 - \Delta x,\, y_2 - \Delta y)\, E_S(x_2 - \Delta x,\, y_2 - \Delta y)$$

and hence the expression for $E_v(x_3,y_3)$ becomes, using equation 3.62 for the reference beam and equation 3.63 for the signal beam,

$$E_v(x_3,y_3) = \left(\frac{i}{\lambda f}\right)^3 \int \cdots \int \tau(x_1,y_1)\, \tau^*(x_1',y_1')$$
$$\times e^{-i\frac{2\pi}{\lambda f}\left[x_2(x_1-x_1' + x_3 - \ell f) + y_2(y_1 - y_1' + y_3 - mf)\right]}$$
$$\times e^{-i\frac{2\pi}{\lambda f}\left[\Delta x(x_1' + \ell f) + \Delta y(y_1' + mf)\right]} dx_1dy_1dx_1'dy_1'dx_2dy_2 \quad . \tag{3.77}$$

Integrating first over x_2,y_2 and then over x_1',y_1' yields

$$E_v(x_3,y_3) = \frac{-i}{\lambda f}\iint \tau(x_1,y_1)\, \tau^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)$$
$$\times c^{-i\frac{2\pi}{\lambda f}\left[\Delta x(x_1 + x_3) + \Delta y(y_1 + y_3)\right]} dx_1dy_1 \quad . \tag{3.78}$$

We see from the above equation that the result of a translation of the hologram plate by an amount $\bar{r}_o$ is the appearance of the phase factor

$$e^{i\psi} = e^{-i\frac{2\pi}{\lambda f}[\Delta x(x_1 + x_3) + \Delta y(y_1 + y_3)]} \tag{3.79}$$

under the integral sign. It is clear that the effect of this phase factor depends on the nature of τ and upon the amount of translation. If $\bar{r}_o$ is sufficiently large such that the phase factor $e^{i\psi}$ oscillates rapidly in that portion of the range of integration which contributes to $E_v(x_3,y_3)$ then $E_v(x_3,y_3)$ will be essentially zero, and the translation will have resulted in the disappearance of the reconstruction of the signal beam. On the other hand, if $\bar{r}_o$ is sufficiently small such that $e^{i\psi}$ is very slowly varying, then the effect of the translation will also be small.

These results are analogous to those obtained in Section 3.2.3, where no lenses were used. In both cases a translation of the hologram plate can result in the disappearance of the reconstruction of the signal beam, with the amount of translation required to make the reconstruction vanish depending on the nature of the "source." There is, however, an important difference between the two cases as when there are no lenses used a translation of the source is equivalent to a translation of the hologram, and thus the reconstruction is equally sensitive to either a translation of the source or a translation of the hologram plate. On the other hand, the Fourier transform hologram is invariant to a translation of the source (i.e., the transparency) but

is quite sensitive to a translation of the hologram plate.

A more direct comparison between the two cases will be made in Section 3.4.5 when we consider diffuse illumination, but before pursuing the analogies between the analytical results of this section with those of Section 3.2, it should be noted that in Section 3.2 we were interested in computing the reconstructed signal field at an arbitrary point on the hologram plate, while in this section we have dealt with the entire reconstructed field imaged in the back focal plane of a lens. Furthermore, the virtual image that we have dealt with in this section includes what was referred to as noise in Section 3.2.

3.4.4 Background Noise

We saw in Section 3.2 that when a multiple wave reference beam was used the virtual image contained both a reconstruction of the original signal beam and a number of waves that were designated as background noise. In that section, because of the form of the fields, it was convenient to treat the reconstruction of the signal beam and the background noise separately. In this section, however, we have found it more convenient to treat the fields in their entirety. Thus the virtual image $E_v(x_3,y_3)$ contains not only a reconstruction of the signal beam, but also the "background noise." The form of $E_v(x_3,y_3)$ is found in a straightforward way by using equation 3.66 which expresses $E_v(x_3,y_3)$ in terms of the autocorrelation of the amplitude distribution of the reference beam in the (x_1,y_1) plane.

It is clear that since the signal beam has been taken to be a

plane wave, in the ideal case where the hologram and imaging lens are of infinite aperture, a noise-free reconstruction of the signal beam would require $E_v(x_3,y_3)$ to be a delta function. We can examine the effect of the nature of the reference beam on the form of $E_v(x_3,y_3)$ in physical terms by observing that the more closely $E_v(x_2,y_2)$ approaches that of $E_S(x_2,y_2)$ (uniform amplitude with a linear phase shift) the more closely $E_v(x_3,y_3)$ will approach being a delta function. Thus, since

$$E_v(x_2,y_2) = E_R(x_2,y_2)\, E_R^*(x_2,y_2)\, E_S(x_2,y_2) \tag{3.80}$$

we see that the more localized $E_R(x_2,y_2)$ is, the more degraded will be the reconstruction, as evidenced by the "spread" of $E_v(x_3,y_3)$. Thus the worst case would appear to be where $E_R(x_2,y_2)$ approaches a delta function, which corresponds to a rather uniform reference beam (plane wave) in the (x_1,y_1) plane. On the other hand, if the reference beam is produced by a single point source in the (x_1,y_1) plane, then

$$E_R(x_2,y_2)\, E_R^*(x_2,y_2) \cong \text{constant}$$

and hence the reconstruction will be essentially noise free. Thus, it is clear that the more "localized" the reference beam source in the x_1,y_1 plane, the more closely $E_v(x_3,y_3)$ will approach the noise-free case. Similar conclusions can be reached by an examination of equation 3.66.

3.4.5 Diffuse Illumination

We would now like to consider the case where the transparency is diffusely illuminated, such as would be the case if a piece of ground glass was placed behind the transparency and illuminated with a plane wave. It is clear that an exact specification of the field in the (x_1,y_1) plane is then quite impractical, because of the inherently random nature of the diffuser. Nevertheless, there is some definite field distribution in the (x_1,y_1) plane, and it can be represented by

$$E_R(x_1,y_1) = N(x_1,y_1)\ \tau(x_1,y_1) \tag{3.81}$$

where τ is the amplitude transmittance of the transparency and N accounts for the diffuser.

The question arises as to whether it is valid to apply the analysis used in the preceding four sections, where we merely replace $\tau(x,y)$ by $N(x,y)\ \tau(x,y)$. We recall that we did not place any explicit restrictions on $\tau(x,y)$, but there are some implied restrictions due to our use of the Fourier transform relationship between the amplitude distributions in the front and back focal planes of a lens, which is a good approximation to the actual case only for paraxial rays. Thus, while the Fourier transform relationship may yield quite accurate predictions for transparencies where the angular spread of the diffracted waves is relatively small, it may not be too accurate when a large angular spread exists, such as could be generated by a diffuser. A thorough investigation of these points will not be given here. Instead, we shall use the analysis of the preceding sections, keeping in mind

the limitations of the analysis.

Thus, it is clear that the same translational invariance of the transparency (plus the diffuser) exists as before, and likewise the same type of translation sensitivity exists with respect to a translation of the hologram plate.

We observe that with the diffuser, it is quite reasonable to consider the reference beam to arise from a very large number of point sources of varying amplitude and phase distributed throughout the portion of the (x_1,y_1) plane occupied by the transparency. Then, since each point source in the (x_1,y_1) plane generates a plane wave in the (x_2,y_2) plane, the exposing and illumination fields are of the form of a sum of plane waves, and the analysis of Section 3.2 is directly applicable. The invariance of the reconstructed signal beam to a translation of the transparency (plus diffuser) is immediately evident from equation 3.17 when we note that the lens has the property that all rays reaching a given point in the (x_2,y_2) plane must have left the (x_1,y_1) plane in the same direction. Thus all the $\bar{k}$'s in equation 3.17 are constant and hence the magnitude of η is a constant independent of $\bar{r}_o$, and thus the reconstruction is not affected by a translation of the transparency (plus diffuser).

Similarly, the analysis of Section 3.2 is directly applicable to the case where the hologram plate is translated, but we will not consider this further here, since this has been rather thoroughly covered in previous sections. What is of interest, however, is an examination of the effect of the diffuser on the "background noise." Thus, let us

consider equation 3.66, with τ replaced by $N\tau$:

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau(x_1,y_1)\, N(x_1,y_1)\, \tau^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)$$

$$N^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)dx_1 dy_1 \,. \quad (3.82)$$

It is apparent that in general the use of the diffuser will make $E_v(x_3,y_3)$ much more highly localized, since one could reasonably assume that the "random" nature of a diffuser would imply that the above integral will have negligible value unless the arguments of N and N^* are almost equal. To put this another way, we may think of the diffuser as generating "white noise," in the sense that the autocorrelation function of N may approach a delta function. This will obviously dominate the above integral, with the result that $E_v(x_3,y_3)$ will be quite localized.

These observations can be put in physical terms by recalling that the spread in $E_v(x_3,y_3)$ could be associated with a localization of $E_R(x_2,y_2)$. The effect of the diffuser is then to spread $E_R(x_2,y_2)$ which in turn results in a narrowing of $E_v(x_3,y_3)$.

3.4.6 Discussion

In the previous sections we have examined the translation sensitivity and background noise associated with Fourier transform holograms when the "reference beam" is the field generated by illuminating the transparency, and the signal beam is an off-axis plane wave.

This type of hologram is of particular interest because of its use in various character recognition systems, such as those discussed by Vander Lugt (41,42) and by Gabor (43). In systems of this nature a Fourier transform hologram is taken of a transparency containing the character or characters to be recognized, and the recognition operation is accomplished by illuminating the hologram with the field from a transparency (using the same optical system as was used to generate the hologram) which may or may not contain the original character or characters. If the transparency is essentially the same as that used to produce the hologram, then the field that illuminates the hologram is essentially the "reference" beam, and the signal beam (the plane wave) is reconstructed. This is brought to a focus in the output plane (x_3,y_3) where its presence signifies that the "test transparency" is in fact the one used to make the hologram.

If the test transparency is different from the original then the signal in the output plane, $E_v(x_3,y_3)$, will be of the form of a correlation of the amplitude transmittance of test transparency τ_T with the amplitude transmittance of the original transparency τ (this type of operation is often called correlation filtering or correlation detection). It is quite straightforward to apply the analysis of Sections 3.4.2, 3.4.3, 3.4.4 and 3.4.5 to this case, as all that is involved is to use

$$E(x_2,y_2) = \frac{-i}{\lambda f} \iint \tau_T(x_1,y_1)\, e^{-i \frac{2\pi}{\lambda f}(x_1x_2 + y_1y_2)}\, dx_1 dy_1 \qquad (3.83)$$

for the illumination field, instead of $E_R(x_2,y_2)$ as given by equation 3.62. We observe that the effects of this substitution on $E_v(x_3,y_3)$ is merely to change τ to τ_T in equations 3.65, 3.66, 3.70, 3.71, 3.72, 3.77, 3.78 and 3.82 (the factor τ^* in these equations remains unchanged). We shall not go through the analysis for this case here, as it would add little if anything new.

It might, however, be useful to make a few comments within the context of character recognition about the case where we have diffuse illumination of the transparency. Suppose we keep the same diffuser but use a different transparency, with amplitude transmittance τ_T, to illuminate the hologram. Then $E_v(x_3,y_3)$ is given by equation 3.82 with τ replaced by τ_T

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau_T(x_1,y_1)\ \tau^*(x_1 + x_3 - \ell f,\ y_1 + y_3 - mf)$$

$$\times N(x_1,y_1)\ N^*(x_1 + x_3 - \ell f,\ y_1 + y_3 - mf)dx_1dy_1 \quad . \qquad (3.84)$$

It is reasonable to assume that N will be a rapidly varying irregular function and hence the integral will have an appreciable value only when the arguments of N and N^* are very close to being equal. This means that the $E_v(x_3,y_3)$ will be quite localized, as discussed in Section 3.4.5. We also observe, however, that $E_v(x_3,y_3)$ will depend on how closely τ_T and τ are correlated, and it thus appears that the system performance may be improved by the use of a diffuser since the output is still dependent on how closely τ_T and τ are correlated

but is now much more highly localized in the output plane. The possibility of improvement of the operation of the system when diffuse illumination is used seems reasonable from a physical point of view when we note that the diffuser results in the field at the hologram plate due to the transparency being much more spread out than before, and hence a much greater area of the hologram plate is utilized.

Another interesting case is where a different diffuser is used with each test transparency, as this situation is equivalent to the situation which would exist if one were to use characters printed on a diffusely reflecting surface for input to the character recognition system. In this case $E_v(x_3,y_3)$ would be given by

$$E_v(x_3,y_3) = \frac{-i}{\lambda f} \iint \tau_T(x_1,y_1)\, \tau^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)$$

$$\times N_T(x_1,y_1)\, N^*(x_1 + x_3 - \ell f,\, y_1 + y_3 - mf)\, dx_1 dy_1 \quad . \qquad (3.85)$$

In general, it would appear reasonable to assume that N_T and N^* are uncorrelated, and hence $E_v(x_3,y_3)$ will be essentially zero except when $N_T = N$. The system thus performs exceptionally well with respect to recognizing the diffuser, but this is not in general what is desired, as what one wishes to recognize is τ , independent of the diffuser. Thus, when there is a different diffuser associated with each test transparency, the system will not function satisfactorily as a character recognition device. We observe that the situation is not improved by setting $N^* = 1$ in equation 3.85 (i.e., by illuminating

the master transparency τ with a plane wave when we produce the hologram), since the factor N_T will still be present and will result in $E_v(x_3,y_3)$ being essentially zero. It is thus clear that based on the assumption that in general N_T and N will be uncorrelated, that in any practical character recognition system the input data must be in the form of a transparency, and if one wishes to construct a character recognition system where the input data is of the form of characters on an opaque diffusely reflecting surface (i.e., printed page) then an auxiliary step is required to put the input into the form of a transparency.

CHAPTER FOUR

AMPLITUDES OF THE DIFFRACTED FIELDS

4.1 Introduction

In this chapter we shall extend the scattering theory of wavefront reconstruction, as developed in Chapter One, to include the computation of the amplitudes of the diffracted waves. We will deal only with transmission holograms, and will make use of both analytical and experimental techniques to determine the quantities which are of interest.

We shall be interested in dealing with transmission holograms involving high spatial frequencies and thick emulsion layers, and hence we must take into account the entire emulsion layer. We recall that this was done in Chapter One, as the grain density was specified at every point within the emulsion layer. We should note, however, that although the grain density is specified throughout the emulsion layer, the grain density alone does not provide a complete description of the emulsion layer. The problem of specifying the characteristics of the individual grains will not be considered, but instead, the effects which are related to the detailed nature of the film grains will be investigated experimentally.

In Chapter One we computed the directions and phases of the diffracted plane waves produced when the grains associated with a particular periodicity term in equation 1.9 are illuminated by an arbitrary plane wave. We saw that they were completely specified by the

wavelength and direction cosines of the incident plane wave and by the periodicity of the recorded interference pattern in the plane of the emulsion surface (or in any plane z = constant). The problem which we wish to consider now is that of determining the amplitudes of the diffracted waves.

It is clear that there are a number of factors which are of significance with regard to the problem of computing the amplitudes of the diffracted waves. They are as follows:

(a) The amplitude of a particular diffracted wave will be strongly dependent on how well the waves scattered by the grains in the various planes z = constant add in phase. This will be a function of the particular diffracted order considered, the wavelength and direction of the illumination wave, the spatial variation of the grain density throughout the emulsion layer, as well as the thickness of the emulsion layer.

(b) The amplitude of a particular diffracted wave will depend on the angular dependence of the amplitude of the wave scattered by each individual grain. This will, of course, depend on the characteristics of the particular grain under consideration.

(c) The amplitudes of the diffracted waves will be affected by losses due to attenuation within the film emulsion layer and reflection losses at the various interfaces. The reflection losses are quite straightforward to compute but the losses due to attenuation within the emulsion layer will depend on the film characteristics, processing procedures, exposure times, etc.

(d) The amplitudes of the diffracted waves will depend on the amount of light scattered by each grain as well as the total number of grains present.

The problem of computing the fields produced when a "volume" diffraction grating is illuminated by a plane wave is not unique to the particular approach which has been used here in connection with holography. Indeed, it arises in connection with information storage in three-dimensional media (8,13) as well as in the diffraction of light by ultrasonic waves (52). We can designate the methods used to obtain a solution of a problem of this type as being either "rigorous" or "scalar" in nature. By rigorous we mean where the solution is obtained by a direct application of Maxwell's equations. We realize, of course, that certain idealizations may be necessary in describing the diffracting volume and certain approximations may be required to obtain solutions of the equations, and thus certain "rigorous" methods may be "more rigorous" than others. Born and Wolf, in Chapter Twelve of their book, "Principles of Optics" (52) treat the problem of diffraction of light by ultrasonic waves using rigorous methods. More recently, and with direct reference to holograms employing "thick" emulsions, Burckhardt (53) solved the problem of computing the diffraction of a plane wave at a sinusoidally stratified (lossless) dielectric grating using a rigorous approach, where the stratifications were perpendicular to the surface of the grating, and where the wave vector of the illumination wave has no component in the direction of the grating lines.

A "scalar" approach, on the other hand, refers to the case where we regard each element of volume of the grating as a source of secondary waves (when the grating is illuminated, of course), and where the diffracted waves are computed by coherently summing the waves produced by each element of volume of the grating. This type of approach was used by van Heerden (13) in connection with his investigation of information storage in solids. Van Heerden restricted his analysis to the case where the illumination plane wave is identical to one of the two original exposing plane waves, and demonstrated that for a very thick grating, the amplitude of the diffracted field is negligible except in the direction of the other original exposing plane wave. More recently, Leith et al (8), in a paper dealing with holographic data storage in three-dimensional media, made a careful investigation of the problem of determining the amplitudes of the diffracted waves produced when a holographic diffraction grating is illuminated with a plane wave. In particular, they considered the case where the propagation vectors of the two waves which generated the grating and the propagation vector of the illumination wave all lie in the same plane, and made a rather comprehensive study of the effect of varying various parameters on the amplitudes of the diffracted waves. They considered both transmission and reflection holograms.

The analysis which we shall employ to compute the amplitudes of the diffracted waves will be based on the scalar approach, and hence will be similar in many respects to that employed by van Heerden (13) and Leith et al (8). There will, however, be a number of significant

differences, both in the formulation and development of the solutions and in the scope of the analysis. For example, we will consider the general case where the wave vectors of the two original exposing plane waves and the wave vector of the illumination plane wave may have arbitrary directions, being restricted only to being incident on the hologram plate from the same side (since we are considering transmission holograms). Furthermore, we shall account for attenuation within the emulsion layer, as well as reflection losses at the different interfaces.

Our basic analytic treatment of the problem will consist of deriving general expressions for the power in the first order diffracted waves, for the case where the grain density varies sinusoidally with position, which corresponds to considering the grains associated with any one of the basic periodicity terms in equation 1.9. This is done in the following section (Section 4.2). Supporting and extending this analytical work is an experimental study dealing with holographic diffraction gratings, formed by recording the interference pattern generated by two plane waves. Due to the nonlinear response of the film, the variation of the grain density in the experimental gratings is not sinusoidal. This, however, presents no difficulties as it is easily handled within the framework of the analysis of Chapter One. The result is simply by the appearance of additional periodicities which can be treated independently of the basic periodicity under consideration.

4.2 Sinusoidally Varying Grain Density - Amplitudes of the Diffracted Waves

In this section we shall be interested in summing the waves scattered by all of the grains within an emulsion layer having a grain density of the form

$$D = D_o + D_o \cos[(\bar{k}_1 - \bar{k}_2)\cdot\bar{r} + \phi_1 - \phi_2] \qquad (4.1)$$

for $0 \leq Z \leq T$, where T is the thickness of the emulsion layer. We shall assume that the emulsion layer is illuminated by a plane wave, and that $\bar{k}_1$, $\bar{k}_2$, ϕ_1 and ϕ_2 are constants.

The assumption that each grain acts as an independent scatterer allows us to group the grains in any convenient manner when summing the scattered waves. A particularly convenient way of doing this is to first sum over x and y and then sum over z. We saw in Chapter One that the field produced by the waves scattered by the grains lying between z and $z + dz$ consists of a series of plane waves, whose directions are specified by equations 1.26 and 1.28. The number of plane waves produced corresponds to the number of allowable values of N in equation 1.28, with one additional constraint, which is that when the variation of the grain density in the transverse direction is sinusoidal, the second and higher order waves will be absent. This is discussed in Appendix VI. It can be seen from equation 4.1 that this is the case here, and hence we need consider only the two first-order waves produced by the grains in each infinitesimal layer of the emulsion.

We saw in Section 1.3.4 that the directions of these two first-order waves depend only on the direction and wavelength of the illumination wave and the spacing and orientation of the loci of constant grain density, none of which are a function of z. Thus there will be only two plane waves produced when the emulsion layer is illuminated by a plane wave. Each of these two plane waves can be thought of as being a superposition of a number of plane waves having the same propagation vectors $\bar{k}_d$ but different amplitudes and phases. Each of these component plane waves is associated with a different "layer" within the emulsion volume (i.e., with a different value of z). Thus, if $E(z)dz$ is the complex amplitude of the wave resulting from the scattering by the grains within the region between z and $z + dz$, then the total amplitude E of the wave is

$$E = \int_0^T E(z)dz \quad . \tag{4.2}$$

It will be convenient to write E in the form

$$E = g\int_0^T A(z)\, e^{i\phi(z)}dz \tag{4.3}$$

where the function g accounts for all factors which are not a function of z, and A and ϕ are real functions.

We shall assume that g is proportional to the amplitude E_i of the illumination wave and to the number of grains contributing to a unit area of the diffracted wave under consideration. Thus g will be

proportional to D_o and to $1/n_d$, where n_d is the z direction cosine of the diffracted wave. The factor $1/n_d$ accounts for the increased number of grains contributing to a unit area of the diffracted wave by virtue of the inclination of the diffracted wave with respect to each elemental "scattering layer" within the emulsion volume.

It is clear that since we are adding the amplitudes of the waves scattered by the individual grains in certain specific directions, the angular dependence of these waves should be accounted for. We know from the solution of the problem of determining the field scattered by such simple objects as spheres, ellipsoids, discs, etc. (50) that the amplitude of the scattered fields often exhibit a variety of lobes and nulls which are a function of the size, shape and orientation of the scattering particle as well as of the polarization of illumination wave. Thus, if the grains are all identical, then the function g will be proportional to some function F which accounts for the angular dependence of the field scattered by a typical grain. If, however, the grains have different sizes, shapes or orientations, then the angular dependence of the field scattered by each grain will be different. In our analysis we treat each grain as being identical, and for the case where the actual silver grains within the emulsion layer are quite different from one another (as is usually the case (51)), then we take as a model an emulsion layer containing identical grains whose characteristics represent the average characteristics of the various different grains.

In the analysis which will be presented here, we will assume that the averaging of the radiation patterns of the different grains is such that we can regard the function F to be a constant. Thus, we can write the function g in the form

$$g = \frac{CE_i D_o}{n_d} \tag{4.4}$$

where C is a constant.

4.2.1 Attenuation

In this section we wish to determine the function $A(z)$ appearing in equation 4.3. It is clear that $A(z)$ will depend on the attenuation of the emulsion layer, which will be a function of position. Thus the attenuation suffered by a wave passing through the emulsion layer will be a function of the particular path along which the attenuation is computed. However, any measurements which we are likely to make will involve a collimated beam whose diameter is much larger than the fringe spacing d, and thus the variations in attenuation will effectively "average out," and the attenuation can be accounted for by some "average attenuation constant" α. We shall thus assume that the dependence of $A(z)$ on the attenuation can be expressed in the form

$$A(z) = e^{-\alpha L'(z)} \tag{4.5}$$

where $L'(z)$ is the total path length within the emulsion. The path length $L'(z)$ is composed of two parts, namely the distance traveled

by the illumination wave in reaching the scattering layer at the depth z and the distance traveled by the diffracted wave in going from the scattering layer to the edge of the emulsion at $z = T$. It is quite straightforward to show that

$$L'(z) = \frac{z}{n_i} + \frac{T-z}{n_d} \quad , \tag{4.6}$$

where n_i and n_z are the z direction cosines of $\bar{e}_i$ and $\bar{e}_d$, the unit vectors pointing in the direction of the illumination and diffracted waves, respectively (within the emulsion layer).

4.2.2 Phase Factor

The phase factor $\phi(z)$, it is recalled, represents the phase difference between the wave generated by the grains in the layer between z and $z + dz$ and the wave generated by the grains in the layer adjacent to the emulsion surface plane, $z = 0$. There are two factors which contribute to the phase difference. First, there will, in general, be a path length difference between the two cases by virtue of the fact that the diffracted wave is "generated" at a different depth within the emulsion surface, and second, there may be an inclination of the "fringe planes" which will result in a phase shift due to the resulting "displacement" of the loci of maximum grain density, which will increase linearly with z. The phase shift due to a "displacement" of the fringe planes was computed in Appendix IV in connection with the problem of translation sensitivity with multiple wave reference beams as discussed in Chapter Three. The results

of this analysis will not be used here, but instead, an approach similar to that used in Section 1.3.4 to derive the generalized grating equations will be employed.

Let us begin by considering the grains lying in the layer between z and $z + dz$. If we use the xyz coordinate system defined in Section 1.3.3 then the loci of maximum grain density will be lines parallel to the x axis, spaced a distance d apart. Since the directions of the two first-order diffracted waves are defined by requiring that the phase difference between the waves scattered by any two grains whose y coordinates differ by d be 2π, the phase of the individual scattered waves in directions of the two first-order diffracted waves will vary linearly with y, being of the form $\frac{2\pi y}{d}$. It is clear that since the grain density also varies with y with the period d, when we add up the contributions to the two first-order diffracted waves by all the grains in the layer, the phase of these diffracted waves will be equal to the phase of the waves scattered by the grains located at positions of maximum grain density.

Thus, since we are interested only in a phase difference, we can neglect such factors as phase shifts of π on scattering and compute the phase factor $\phi(z)$ strictly in terms of path length differences. We can write $\phi(z)$ in the form

$$\phi(z) = \frac{2\pi}{\lambda_i} (L(z) - L(0)) \tag{4.7}$$

where $L(z)$ is the path length from an arbitrary wavefront of the illumination plane wave to an arbitrary wavefront of the diffracted plane wave under consideration, where the scattering takes place from any grain in a plane of maximum grain density which is located at a depth z. Although phase differences of $2N\pi$ (N an integer) are not of significance, it is most convenient to compute $L(z)$ using the same plane of maximum grain density, so that $\phi(z)$ will be a continuous function.

In computing $L(z)$, we shall make use of the analysis of Section 1.3.4, and define the path length of interest in terms of the planes A and B defined therein. We recall that plane A is normal to the wave vector of the illumination wave, and is located on the incident side of the emulsion layer, while plane B is normal to the wave vector of the diffracted wave under consideration and is located on the opposite side of the emulsion layer. In general the region surrounding the film plate will have an index of refraction differing from that of the emulsion layer and its supporting substrate, and hence refraction will occur at the various interfaces. We find, however, that this does not affect $\phi(z)$, as the path length differences which yield $\phi(z)$ occur solely within the emulsion layer. Thus we can assume, without loss of generality, that the surrounding medium has the same index of refraction as the emulsion layer.

The path length $L(z)$ is given by equation 1.24, which we shall rewrite here with $\bar{r}_o$ replaced by $\bar{r}(z)$, the position vector of a point lying in a plane of maximum grain density at a depth z :

$$L(z) = C_2 - C_1 + \bar{r}(z)\cdot\bar{e}_i - \bar{r}(z)\cdot\bar{e}_d \quad . \tag{4.8}$$

Using equations 4.7 and 4.8, the expression for $\phi(z)$ becomes

$$\phi(z) = \frac{2\pi}{\lambda_i}\left[(\bar{r}(z) - \bar{r}(0))\cdot\bar{e}_i - (\bar{r}(z) - \bar{r}(0))\cdot\bar{e}_d\right] \quad , \tag{4.9}$$

where λ_i is the wavelength of the illumination wave within the emulsion layer.

The vector $\bar{r}(z) - \bar{r}(0)$ is a vector pointing from a point in a plane of maximum grain density with z coordinate equal to zero to another point in this same plane with z coordinate equal to z. In actuality, the value of $\phi(z)$ is not changed if we simply use a plane of constant grain density, rather than a plane of maximum grain density. The planes of constant grain density are defined by

$$(\bar{e}_1 - \bar{e}_2)\cdot\bar{r} = C \tag{4.10}$$

where $\bar{r}$ is the position vector, C is a constant and $\bar{e}_1$ and $\bar{e}_2$ are unit vectors pointing in the directions of $\bar{k}_1$ and $\bar{k}_2$ in equation 4.1.

It is clear that since both $\bar{r}(z)$ and $\bar{r}(0)$ satisfy equation 4.10 for the same value of C, we can write

$$(\bar{e}_1 - \bar{e}_2)\cdot[\bar{r}(z) - \bar{r}(0)] = 0 \quad . \tag{4.11}$$

It is convenient to express both the vectors $(\bar{e}_1 - \bar{e}_2)$ and $(\bar{r}(z) - \bar{r}(0))$ in the xyz coordinate system, which is defined so that

the vector $(\bar{e}_1 - \bar{e}_2)$ has no x component (see Section 1.3.3). In this coordinate system we can write

$$\bar{e}_1 - \bar{e}_2 = \bar{e}_y|\bar{e}_1 - \bar{e}_2| \cos\gamma - \bar{e}_z|\bar{e}_1 - \bar{e}_2| \sin\gamma \qquad (4.12)$$

where γ is the inclination angle of the fringe planes, defined as the angle the normal to the planes of constant grain density makes with the plane of the emulsion surface. Similarly, we can write

$$\bar{r}(z) - \bar{r}(0) = \Delta x\bar{e}_x + \Delta y\bar{e}_y + z\bar{e}_z \quad , \qquad (4.13)$$

where Δx is arbitrary and Δy is a function of γ and z. We can evaluate Δy by using equations 4.12 and 4.13 in equation 4.11. That is

$$|\bar{e}_1 - \bar{e}_2|\Delta y \cos\gamma - |\bar{e}_1 - \bar{e}_2|z \sin\gamma = 0 \quad , \qquad (4.14)$$

and hence

$$\Delta y = z \tan\gamma \quad . \qquad (4.15)$$

Thus, using equations 4.13 and 4.15 in equation 4.9, the expression for $\phi(z)$ becomes

$$\phi(z) = \frac{2\pi}{\lambda_i}[(\Delta x\bar{e}_x + z \tan\gamma\, \bar{e}_y + z\bar{e}_z)\cdot(\bar{e}_i - \bar{e}_d)] \qquad (4.16)$$

or

$$\phi(z) = \frac{2\pi}{\lambda_i}[\Delta x(\ell_i - \ell_d) + z(m_i - m_d)\tan\gamma + z(n_i - n_d)] \qquad (4.17)$$

where ℓ, m and n are the x, y and z direction cosines in the xyz coordinate system. In this coordinate system, the x direction cosines of the illumination and diffracted waves are equal (equation 1.26) and hence the coefficient of Δx is zero, and $\phi(z)$ becomes

$$\phi(z) = \frac{2\pi z}{\lambda_i} [(m_i - m_d) \tan \gamma + (n_i - n_d)] \quad . \tag{4.18}$$

4.2.3 Integration Over z - Summing the Fields Generated at Different Depths Within the Emulsion Layer

Making use of equations 4.5, 4.6 and 4.18, the integrand in equation 4.3 can be written in the form

$$A(z)\, e^{i\phi(z)} = e^{-\alpha T/n_d}\, e^{wz} \tag{4.19}$$

where w is a complex valued function which is independent of z :

$$w = a + ib \tag{4.20}$$

where

$$a = -\alpha\left(\frac{1}{n_i} - \frac{1}{n_d}\right) \tag{4.21}$$

and

$$b = \frac{2\pi}{\lambda_i} [(m_i - m_d) \tan \gamma + (n_i - n_d)] \quad . \tag{4.22}$$

The integral in equation 4.3 is thus easily evaluated to yield the following expression for the amplitudes of the first order diffracted waves:

$$E = ge^{-\alpha T/n_d} \frac{(e^{wT}-1)}{w} \quad . \tag{4.23}$$

The power per unit area in the diffracted waves is given by $P = EE^*$, which can be expressed in the form

$$P = \frac{CD_o^2 E_i E_i^* e^{-2\alpha T/n_d}}{n_d^2(a^2 + b^2)} [1 + e^{2aT} - 2e^{aT} \cos bT] \tag{4.24}$$

where we have used the value of g given by equation 4.4 and set CC^* equal to a new constant C .

We shall not attempt to compute either C or α , but we can obtain an estimate of the value of α for a particular grating by making a direct measurement of the attenuation of the emulsion layer. The other quantities appearing in the above equation can be computed from a knowledge of the characteristics of the original exposing fields, the illumination wave and the physical characteristics of the emulsion layer.

4.2.4 Computing the Diffracted Waves

In this section we shall describe in detail the computational steps involved in computing the amplitudes and directions of the two first-order diffracted waves, when refraction at the various interfaces and reflection losses are taken into account. We shall assume that the initial exposing plane waves and the illumination wave are given, being specified in the region adjacent to the film emulsion layer.

We begin by defining a general reference coordinate system, (x', y', z') defined so that the emulsion layer occupies the region $0 \leq z' \leq T$, where T is the thickness of the emulsion layer. The original exposing field is taken to be two plane waves, whose propagation vectors in the region $z' \leq 0$ are given by

$$\bar{k}_{10} = \frac{2\pi}{\lambda_o} \bar{e}_{10} \tag{4.25}$$

and

$$\bar{k}_{20} = \frac{2\pi}{\lambda_o} \bar{e}_{20} \cdot \tag{4.26}$$

The unit vectors $\bar{e}_{10}$ and $\bar{e}_{20}$ are specified by their direction cosines in the $x'y'z'$ coordinate system, namely ℓ'_{10}, m'_{10}, n'_{10}, and ℓ'_{20}, m'_{20}, n'_{20}. These two plane waves undergo refraction at the interface between the emulsion layer and the adjacent medium, the $z' = 0$ plane. If η is the index of refraction of the emulsion layer and η_o is the index of refraction of the region $z' \leq 0$, then the wavelength in the emulsion layer is given by

$$\lambda = \frac{\eta_o}{\eta} \lambda_o \cdot \tag{4.27}$$

The wave vectors of the two plane waves in the emulsion layer are then of the form

$$\bar{k}_1 = \frac{2\pi}{\lambda_o} \frac{\eta}{\eta_o} \bar{e}_1 \tag{4.28}$$

and

$$\bar{k}_2 = \frac{2\pi}{\lambda_o} \frac{\eta}{\eta_o} \bar{e}_2 \cdot \tag{4.29}$$

The directions of $\bar{e}_1$ and $\bar{e}_2$ are found from $\bar{e}_{10}$ and $\bar{e}_{20}$ by applying Snell's Law. We find that we can express the direction cosines of $\bar{e}_1$ and $\bar{e}_2$ in terms of those of $\bar{e}_{10}$ and $\bar{e}_{20}$ as follows:

$$\ell'_j = \frac{\eta_o}{\eta} \ell'_{jo} \tag{4.30}$$

$$m'_j = \frac{\eta_o}{\eta} m'_{jo} \tag{4.31}$$

$$n'_j = \sqrt{1 - (\frac{\eta_o}{\eta})^2 (1 - n'^2_{jo})} \tag{4.32}$$

where $j = 1,2$. Next, we define the xyz coordinate system by a rotation of Φ about the z' axis (equation 1.16), where

$$\Phi = \tan^{-1} (\frac{\ell'_2 - \ell'_1}{m'_1 - m'_2}) \quad . \tag{4.33}$$

We observe that the same value of Φ is obtained if we use the fields in the region external to the emulsion layer, since the factor η_o/η will cancel in equation 4.33.

The transformation equations for the direction cosines are given by (Appendix III)

$$\ell_j = \ell'_j \cos \Phi + m'_j \sin \Phi \tag{4.34}$$

$$m_j = -\ell'_j \sin \Phi + m'_j \cos \Phi \tag{4.35}$$

$$n_j = n'_j \tag{4.36}$$

where $j = 1,2$. By virtue of the way the xyz coordinate system is defined, the intersection of the fringe planes with the $z = 0$ plane are lines parallel to the x axis. The periodicity d in the $z = 0$ plane is the distance in the y direction between planes of maximum grain density, and is given by (equation 1.19)

$$d = \frac{\lambda}{|m_1 - m_2|} \quad . \tag{4.37}$$

The fringe plane inclination angle γ is the angle that the normal to the fringe planes $\bar{n}$ makes with the y axis. Thus

$$\gamma = \cos^{-1} (\bar{n} \cdot \bar{e}_y) \tag{4.38}$$

where the sign of γ is fixed by equation 4.15 and

$$\bar{n} = \frac{\bar{e}_1 - \bar{e}_2}{|\bar{e}_1 - \bar{e}_2|} \quad . \tag{4.39}$$

The xyz coordinate system, d , and γ along with the physical properties of the emulsion layer such as its index of refraction, attenuation, and thickness are adequate to specify the nature of the grating and hence allow us to compute the amplitudes and directions of the first-order diffracted waves produced when the grating is illuminated with a plane wave. It should be kept in mind, however, that we are still considering only the case where the spatial variation of the grain density is sinusoidal.

Having specified the nature of the grating, let us now apply the results of Section 1.3.4 to compute the directions of the two first-order diffracted waves produced when the grating is illuminated with a

plane wave (there may, of course, be no diffracted waves produced for certain illumination waves). If the propagation vector of the illumination wave is given by

$$\bar{k}_{io} = \frac{2\pi}{\lambda_{io}} \bar{e}_{io} \tag{4.40}$$

in the region $z \leq 0$, then it will be given by (after refraction at the interface at $z = o$)

$$\bar{k}_{i} = \frac{2\pi}{\lambda_{i}} \bar{e}_{i} \tag{4.41}$$

in the region $z > 0$, where

$$\lambda_{i} = \frac{\eta_{o}}{\eta} \lambda_{io} \tag{4.42}$$

and

$$\ell_{i} = \frac{\eta_{o}}{\eta} \ell_{io} \tag{4.43}$$

$$m_{i} = \frac{\eta_{o}}{\eta} m_{io} \tag{4.44}$$

$$n_{i} = \sqrt{1 - (\frac{\eta_{o}}{\eta})^2 (1 - n_{io}^2)} \quad . \tag{4.45}$$

The direction cosines of the two first-order diffracted waves are given by (equations 1.26 and 1.27)

$$\ell_{d} = \ell_{i} \tag{4.46}$$

$$m_{d} = \frac{N\lambda_{i}}{d} + m_{i} \quad , \tag{4.47}$$

where N has the values $+1$ and -1, corresponding to the two different first-order diffracted waves. The above two equations specify n_d by virtue of the relationship

$$\ell_d^2 + m_d^2 + n_d^2 = 1 \quad . \tag{4.48}$$

That is

$$n_d = \sqrt{1 - \ell_i^2 - \left(\frac{N\lambda_i}{d} + m_i\right)^2} \quad , \tag{4.49}$$

where we have taken the positive sign for the square root by virtue of the fact that we are interested in transmitted, rather than reflected waves. This completes the specification of the quantities m_d, n_d, d and γ in terms of the parameters of the original exposing waves and of the illumination wave, which are assumed to be given. Thus, once we assign values to C, D_o^2, $E_i E_i^*$, α and T the diffracted power can be computed, using equation 4.24. This will yield the diffracted power at the boundary plane of the emulsion layer $z = T$ in terms of the illumination power $E_i E_i^*$ at the other boundary surface, $z = 0$. In practice, however, the illumination power is specified in the region $z < 0$ and the power in the diffracted waves is measured after the waves leave the film plate, and hence it may be necessary to take reflection losses into account.

We shall be interested in the case where the film plate consists of an emulsion layer supported by a glass substrate (such as Kodak 649-f glass film plates), and thus there are three interfaces at which

reflection losses can occur. In general, however, the index of refraction of the substrate is close enough to that of the emulsion layer so that reflection losses at the emulsion-substrate interface can be neglected. This leaves two interfaces to be considered, the front and back surfaces of the film plate.

The quantity which is of interest to us is the transmissivity τ of the interface under consideration, defined as the ratio of the transmitted power to the incident power. It is shown by Born and Wolf (54) that the transmissivity depends only on the polarization and angle of incidence of the incident wave, and on the index of refraction on either side of the interface. Furthermore, they utilize Snell's Law to obtain

$$\tau_{\parallel} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t)\cos^2(\theta_i - \theta_t)} \tag{4.50}$$

and

$$\tau_{\perp} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t)} \quad , \tag{4.51}$$

where $\tau_{\parallel}$ applies to the case where the electric field vector lies in the plane of incidence and $\tau_{\perp}$ applies to the case where it is perpendicular to the plane of incidence. In the above expressions θ_i is the angle of incidence of the wave incident on the interface and θ_t is the corresponding angle of the transmitted wave after refraction at the interface.

The transmissivity of the front surface of the film plate can be found using whichever of the above expressions is appropriate (or perhaps both if the illumination wave has an electric field vector with components in both directions), where the angles θ_i and θ_t are given by

$$\theta_i = \cos^{-1}(n_{io}) \tag{4.52}$$

and

$$\theta_t = \cos^{-1}(n_i) \tag{4.53}$$

where n_{io} is assumed to be given and n_i is computed using equation 4.45.

Similarly, the transmissivity of the back surface of the film plate is found using values of θ_i and θ_t given by

$$\theta_i = \cos^{-1}(n_d) \tag{4.54}$$

$$\theta_t = \cos^{-1}(n_{dl}) \tag{4.55}$$

where n_d is given by equation 4.49 and n_{dl} is given by

$$n_{dl} = \sqrt{1 - \left(\frac{n}{n_1}\right)^2 (1 - n_d^2)} \quad , \tag{4.56}$$

where n and n_1 are the indices of refraction of the glass substrate and adjacent medium respectively.

The diffracted power per unit area in the two first-order

waves, taking into account reflection losses, is thus given by

$$P = \frac{CD_o^2 E_i E_i^* \tau_i \tau_d}{n_d^2(a^2 + b^2)} e^{-2\alpha T/n_d} [1 + e^{2aT} - 2e^{aT} \cos bT] \qquad (4.57)$$

where τ_i and τ_d are the transmissivities of the front and back surfaces, respectively.

4.2.5 Special Cases

In this section we shall examine the case where the wave vector of the illumination wave has no component in the direction of the "grating lines" (the x direction), and hence where

$$\ell_{io} = 0 \quad . \qquad (4.58)$$

Then, it follows from equations 4.43 and 4.46 that

$$\ell_i = \ell_d = 0 \qquad (4.59)$$

and hence the wave vectors of the illumination and diffracted waves all lie in the yz plane. If we define the angle θ by

$$\theta = \sin^{-1}(m) \qquad (4.60)$$

then

$$m = \sin\theta \qquad (4.61)$$

and, from equation 4.48 with $\ell_d = 0$,

$$n = \cos \theta \quad . \tag{4.62}$$

Using equations 4.61 and 4.62 in equations 4.21 and 4.22, the expressions for a and b become

$$a = -\alpha\left(\frac{1}{\cos \theta_i} - \frac{1}{\cos \theta_d}\right) \tag{4.63}$$

and

$$b = \frac{2\pi}{\lambda_i}\left[(\sin \theta_i - \sin \theta_d) \tan \gamma + (\cos \theta_i - \cos \theta_d)\right] \tag{4.64}$$

or, using simple trigonometric identities, b can also be shown to be equal to

$$b = \frac{2\pi}{\lambda_i \cos \gamma}\left[\cos (\theta_i - \gamma) - \cos (\theta_d - \gamma)\right] \quad . \tag{4.65}$$

Next, let us consider the case where we neglect attenuation losses ($\alpha = 0$), neglect reflection losses ($\tau_i \tau_d = 1$), and let $\gamma = 0$ (fringe planes normal to the emulsion surface). Then, equation 4.57 assumes the form

$$P = \frac{2CD_o^2 E_i E_i^* \{1 - \cos[\frac{2\pi T}{\lambda_i}(\cos \theta_i - \cos \theta_d)]\}}{\cos^2 \theta_d \left[\frac{2\pi}{\lambda_i}(\cos \theta_i - \cos \theta_d)\right]^2} \tag{4.66}$$

or, using the identity $1 - \cos A = 2 \sin^2 (\frac{A}{2})$,

$$P = \frac{CD_o^2 E_i E_i^* \sin^2 \left[\frac{\pi T}{\lambda_i}(\cos \theta_i - \cos \theta_d)\right]}{\left[\frac{\pi}{\lambda_i} \cos \theta_d (\cos \theta_i - \cos \theta_d)\right]^2} \quad . \tag{4.67}$$

The above expression is equal, to within a constant, to the results obtained by Leith et al (8) for the same case (equation 31 in reference 8).

4.3 Producing Holographic Diffraction Gratings - Experimental Apparatus and Techniques

In this section we shall describe the experimental apparatus which was used to produce the holographic diffraction gratings referred to in this and the next chapter. Variations of the basic experimental setup described here were used in the experimental studies described in Chapters Three and Five.

Figure 4.1 shows a diagram of the experimental apparatus. The source is a helium-neon laser and the optical system consists of a beam splitter, mirrors and lenses positioned so as to illuminate the film plate with two plane waves. The film plate is held in a rotatable holder, which allows the orientation of the film plate to be varied with respect to the illumination plane waves, which are fixed in direction (the angle between them being 30°). The various components are fastened to the surface of a 3 x 6 foot surface plate, which is "floated" to uncouple the apparatus from building vibrations.

4.3.1 Source

The source is a helium-neon laser having a cavity length of 57 cm, with a 60 cm radius output mirror (Spectra Physics #7259 HT, with collimation correction) and a high reflectance flat for the other mirror

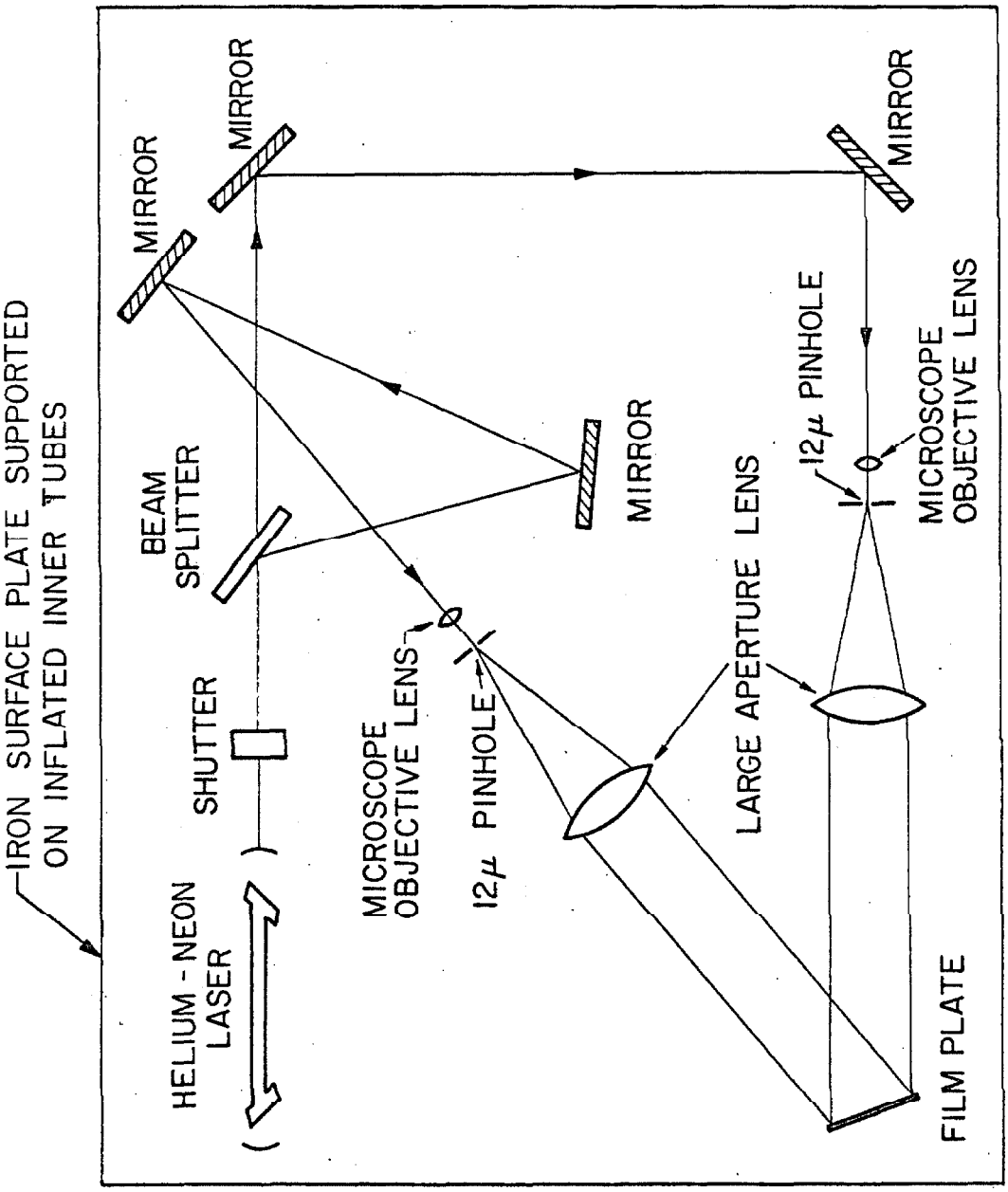


FIGURE 4.1 EXPERIMENTAL APPARATUS FOR THE PRODUCTION OF HOLOGRAPHIC DIFFRACTION GRATINGS

(Spectra Physics #8283). The CW power output at 6328 A° ranged from 1 to 4 milliwatts at a beam current of approximately 15 milliamps, depending on the mirror alignment and on the cleanliness of the Brewester angle windows.

The laser normally oscillates in more than one axial mode, and no attempts were made to achieve single mode operation. Efforts were made, however, to limit the oscillations to the lowest order transverse mode (by mirror adjustments) and were in general reasonably successful. The fact that the laser is oscillating in more than one mode implies that the coherence length of the laser output is correspondingly reduced. Questions regarding the coherence length were simply avoided by making the path lengths approximately equal by suitably positioning the mirrors.

The laser tube was oriented so that the output beam (which is linearly polarized) had its E vector perpendicular to the surface of the surface plate. This choice of polarization is advantageous as it results in higher values of reflectivity for the beam splitter and avoids depolarization upon reflection from the mirrors. A more important reason for choosing this polarization arises from the fact that the film is sensitive only to the total electric field, and with this polarization the electric field vectors of the two illumination waves are colinear. With the other polarization, however, this is not the case and when the two waves are propagating at right angles to each other there will be no interference pattern formed (with respect to the electric field). This is discussed in detail in Appendix I.

4.3.2 Mechanical Stability

Before discussing the steps we have taken to achieve mechanical stability, let us briefly review the reasons why it may be necessary to insist on a high degree of mechanical stability in a hologram-taking apparatus. In recording a hologram, we are in essence recording an interference pattern which is quite sensitive to path length changes. To be assured of an adequate recording of the interference pattern we must require that the film plate remain in register with the interference pattern to within at least one quarter of a fringe during the duration of the exposure. It is clear that the stability requirements are directly related to the exposure time, which for the apparatus described here ranged from several seconds to more than one hour.

A change in register of the film plate with respect to the interference pattern can be caused by either a movement of the film plate or by a shift of the interference pattern. While the effect of the former is proportional to the spatial frequency of the interference pattern, the latter is not. Indeed, a path length change of λ will cause a shift of one fringe regardless of the fringe spacing. Such path length changes can be caused by similar changes in the position of the reflecting elements in the optical system that encounter the beam after it has been divided (beamsplitter and mirrors). Changes in the optical path length that occur prior to the beamsplitting operation are not important.

In the apparatus described here, all components are rigidly mounted and securely fastened to the surface plate. The surface plate,

which weighs approximately one thousand pounds, provides structural rigidity as well as inertial damping for the system. The surface plate is isolated from mechanical vibrations in the floor of the building by supporting it on a layered structure of felt, neoprene sponge rubber, plywood and low pressure rubber inner tubes. Measurements made by J. Azmuth (55) on this system in the 20-20,000 CPS range have shown that the peak surface plate acceleration is approximately 1/28 that of the nearby floor. Isolation from acoustical vibrations is accomplished by covering the components with felt covered boxes. This also reduces the effects of air turbulence.

The apparatus, as described above, was found to be very stable. Holograms were obtained using it in which the exposure times exceeded one hour. Holograms were also obtained with the table "unfloated," with exposure times of the order of five minutes.

4.3.3 Optical Components

The optical components consist of the beam splitter, mirrors and the two collimating lens systems. The beam splitter is simply a 4 1/2 inch diameter quartz plate, .242 inches thick, with a wedge angle of 47 seconds. It yields two primary beams of equal intensity when the angle of incidence is 75 degrees, with 40% of the power being lost due to multiple reflections. The mirrors are high quality front surface mirrors (Davidson Optronics Model D615). It was found that ordinary front surface mirrors (coated select plate glass) often exhibit what is referred to as the orange peel effect, which is evidenced by a

mottled appearance of the reflected beam.

The two collimating lens systems were identical, each consisting of a 16 mm microscope objective lens and a 6" aperture 19" focal length lens, placed so that their focal planes coincided. It was found that invariably dust or other small particles would be present on the surface of the microscope objective lens or on the mirrors, and that these particles generated diffraction patterns that caused rapid amplitude variations across the beam. These effects were removed by placing a small aperture, or "pinhole," at the focal point of the microscope objective lens. The operation of the pinhole is readily explained in terms of a low pass spatial filtering operation, using the analysis contained in Appendix V. It was found that pinholes with a diameter in the range from 10 to 30 microns were quite satisfactory (the pinholes were obtained from Buckbee Meers Inc.). Accurate positioning of the pinholes was required (to within at least .001 inch in all three directions) and this was achieved by using three-dimensional micropositioners (Kulicke and Soffa Model 200).

Although the low pass spatial filtering operation eliminated the rapid amplitude fluctuations across the beam, there was still a slow variation due to the decrease of amplitude with radius which is characteristic of the lowest order transverse mode of the laser. This produced a decrease in the power of an order of magnitude at a radius of 3 cm when a 16 mm microscope objective was used.

4.4 Measurement of the Power in the Diffracted Waves - Experimental Apparatus and Techniques

In this section we shall describe an experimental apparatus whose function is to measure the direction and power of the various diffracted waves which are produced when a holographic diffraction grating is illuminated by a collimated beam. A diagram of the apparatus is shown in Figure 4.2, from which we see that the apparatus consists essentially of a source to illuminate the grating and a detection system to measure the power in the diffracted waves. The photomultiplier can be rotated about the vertical axis only, and thus we are restricted to measurements in which the illumination wave and the diffracted waves all have their propagation vectors in the same (horizontal) plane.

The device used to hold the grating and photomultiplier is a converted spectrometer. The grating is mounted on a rotatable table which in turn is mounted on the spectrometer table. The use of this additional rotatable table, whose rotation with respect to the illumination beam can be measured to within 5 minutes of arc, allows the angle of incidence of the illumination beam to be read directly, without being dependent on the angular position of the viewing telescope.

The photomultiplier, an RCA 7102, is mounted on the body of the viewing telescope of the spectrometer, which can be rotated to pick up the various diffracted orders. The spectrometer is of quite high quality, and the angular position of the viewing telescope, and hence the photomultiplier, can be read to within one minute of arc. In mounting the photomultiplier on the viewing telescope, the objective lens

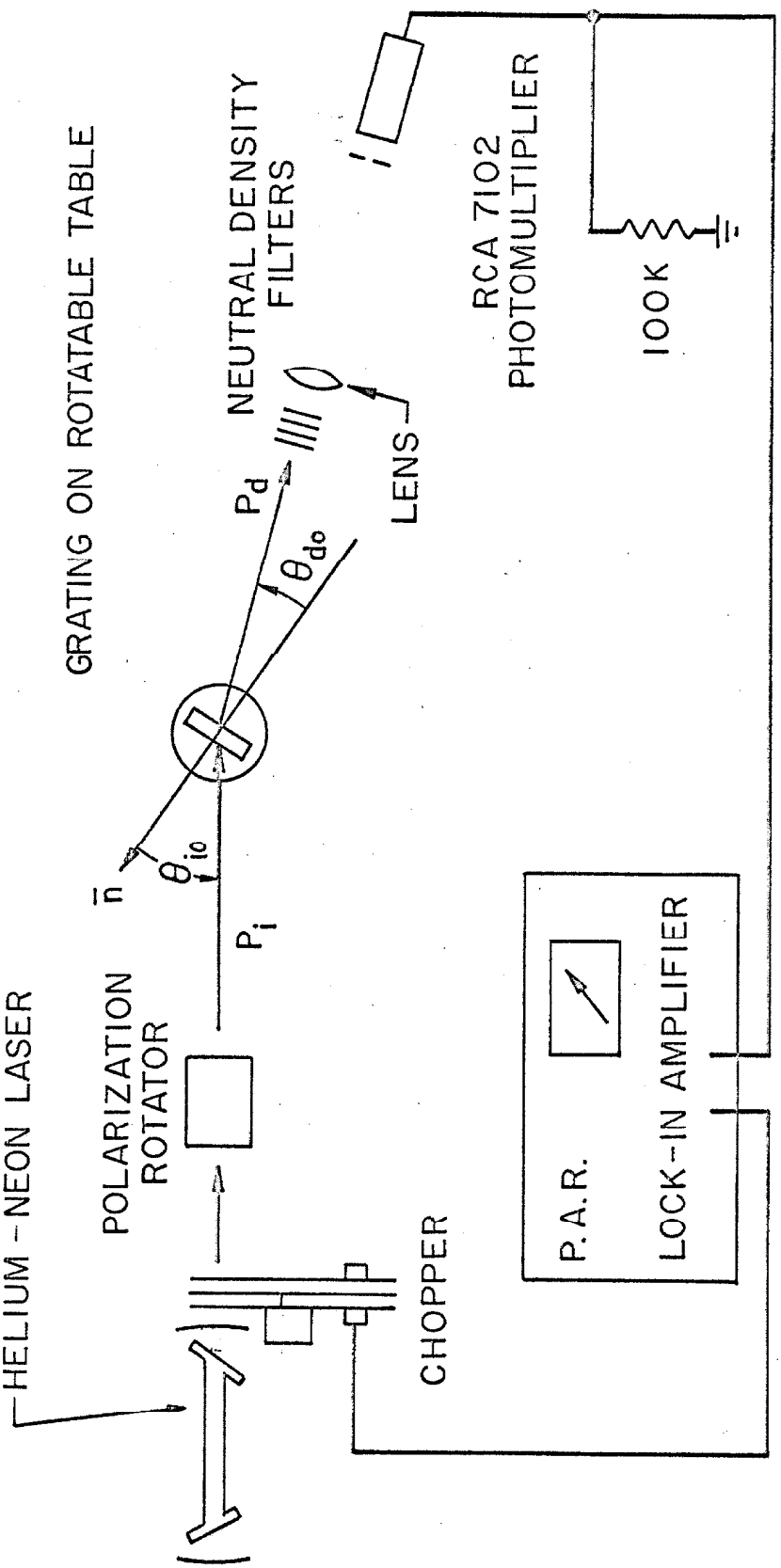


FIGURE 4.2 EXPERIMENTAL APPARATUS FOR MEASURING THE DIRECTION AND POWER OF THE DIFFRACTED WAVES PRODUCED WHEN A HOLOGRAPHIC DIFFRACTION GRATING IS ILLUMINATED WITH A COLLIMATED BEAM.

was retained but the eyepiece assembly was replaced by a small aperture. By making this aperture sufficiently small, the angular sensitivity of the detection system could be made equal to the accuracy of the graduated circle, which was one minute of arc. In practice, however, we found it convenient to use a larger (.078" diameter) aperture, which yielded an angular sensitivity on the order of 20 minutes of arc.

The source was a helium-neon laser which produced a linearly polarized output beam at .6328μ . A Spectra Physics polarization rotator was used to enable us to illuminate the grating with any desired linear polarization.

4.4.1 Photomultiplier Detection System

A phase sensitive detection system was employed which allowed the detection of very low signal levels. The output beam of the laser was chopped at about 10^3 CPS, with the chopper providing a reference signal at the same frequency, which remains in phase with the chopped laser beam. The output signal from the photomultiplier is detected by a lock-in amplifier (Princeton Applied Research Model JB-5) which utilizes the reference signal generated by the chopper to discriminate against that portion of the signal from the photomultiplier not in a narrow frequency band centered at 10^3 CPS.

What was desired was the ability to detect signals having a wide range of power levels, from 1 milliwatt (the power level of the direct laser beam) down to 10^{-6} or 10^{-7} mw. Rather than have the input to the detection system vary over such a large range, this range was

achieved by the use of neutral density filters, with the input to the photomultiplier varying only over one order of magnitude. What was done was to position the photomultiplier so as to pick up the direct laser beam (with no grating in place) and then set the photomultiplier voltage and lock-in amplifier gain so that the meter on the lock-in amplifier read 100 with a neutral density of 5 in front of the photomultiplier (five N.D. filters of N.D. = 1.0). Weaker signals could then be read by appropriate removal of neutral density filters. This gave a reliable range of readings over 60 db, with an additional range of 10 db corresponding to meter readings from 1 to 10 with no neutral density filters in the beam. This method has the additional advantage that the data is automatically normalized with respect to the power in the illumination beam.

The sensitivity of the system could be increased in two ways, either by increasing the photomultiplier voltage and working at lower signal levels or by increasing the laser power. This latter method, of course, is most desirable but one is limited by the power output of the lasers that one has available.

A number of different helium-neon lasers were used in the setup, and there was a fairly wide range in the stability of the output power between the different lasers. The instabilities were observed to be of two types, slow long term drifting of the power level and rapid noise-like fluctuations. The long term drifts could be corrected by periodically resetting the photomultiplier voltage and the rapid fluctuations could be allowed for in reading the meter, provided they

were not too large. It is estimated that together these effects limited the accuracy of the readings to about ±5% for the worst cases.

4.4.2 Relating the Measured Power Ratio to the Amplitudes of the Diffracted Waves

In the analysis of Section 4.2 it was assumed that the grating was essentially infinite in extent and was illuminated by a plane wave which was likewise infinite in extent. The amplitude of the diffracted plane waves were then computed and the power per unit area was computed by taking the square of the amplitude of the diffracted wave. In the experimental apparatus described here only a portion of the grating is illuminated, and this with a collimated beam whose amplitude varies across the beam. The ratio of the total power in the diffracted beam to that in the illumination beam is what is measured, and from this we wish to determine the power per unit area in the diffracted beam, normalized with respect to the square of the amplitude of the illumination field.

Let us define (x', y'), (x,y), and (x'', y'') to be the transverse coordinates of the illumination beam, the grating and the diffracted beam, respectively. The amplitude of the illumination beam will be some function of x', y', which we shall denote by $E_i(x', y')$, and the total power in the illumination beam will be of the form

$$p_i = c \int_{x'} \int_{y'} E_i E_i^* \, dx' \, dy' \tag{4.68}$$

where c is the appropriate constant and the integration extends over the cross sectional area of the illumination beam. Likewise, if $E_d(x'', y'')$ is the amplitude of the diffracted field under consideration, then the total power in the diffracted beam is of the form

$$p_d = c \int_{x''} \int_{y''} E_d E_d^* \, dx'' \, dy'' \quad , \tag{4.69}$$

where the integration extends over the cross sectional area of the diffracted beam and the constant c is the same constant appearing in equation 4.68 . We assume that the characteristics of the grating are constant over the illuminated area and hence that the variation of E_d with x'', y'' is due solely to the variation of E_i across the illuminated portion of the grating. We can thus write p_d in the form

$$p_d = cK \int_{x''} \int_{y''} E_i E_i^* \, dx'' \, dy'' \tag{4.70}$$

where

$$K = \frac{E_d E_d^*}{E_i E_i^*} \quad . \tag{4.71}$$

We observe that the illumination beam and the diffracted beam share the same area of the grating, and that we can write

$$p_i = cn_i \int_{x} \int_{y} E_i E_i^* \, dx \, dy \tag{4.72}$$

and

$$p_d = cKn_d \int_x \int_y E_i E_i^* \, dx \, dy \quad , \tag{4.73}$$

where n_i and n_d are the z direction cosines of the illumination and diffracted beams, respectively. The quantity which is measured, p_d/p_i, is thus of the form

$$\frac{p_d}{p_i} = \frac{n_d}{n_i} \frac{E_d E_d^*}{E_i E_i^*} \quad . \tag{4.74}$$

We thus have the choice, in comparing the measured power ratio p_d/p_i with the computed power as given by equation 4.57, of converting the measured data by multiplying by the factor n_i/n_d or of computing $\frac{n_d P}{n_i E_i E_i^*}$ rather than $P/E_i E_i^*$. In general we shall do the latter, since it merely requires a simple change in the computer program. Thus, what we will wish to compute is

$$p = \frac{C \tau_i \tau_d e^{-2\alpha T/n_d}}{n_i n_d (a^2 + b^2)} [1 + e^{2aT} - 2e^{aT} \cos bT] \quad , \tag{4.75}$$

where we have absorbed the factor D_o^2 in the constant c.

4.5 Experimental Study of Holographic Diffraction Gratings

In this section we shall consider the dependence of the power diffracted into the first-order waves on the direction, wavelength and polarization of the illumination wave, as well as on the characteristics of the particular grating being considered. In particular, we shall discuss a number of measurements that were made and compare the experimental data with the theoretical values computed using the analysis developed in Section 4.2. A brief description of the results of some of these measurements was reported in a previous publication (9).

The application of the analysis of Section 4.2, which deals with the case where the variation of grain density is sinusoidal, to the case of a holographic diffraction grating, is quite straightforward. According to the analysis developed in Chapter One the grain density of a hologram formed by recording the interference pattern generated by two plane waves can be written as a sum of sinusoidal terms, the lowest order of which generates the two first-order diffracted waves. By virtue of the assumption of the neglect of multiple scattering, the waves generated by the grains associated with the different sinusoidal terms can be dealt with separately, and hence the problem reduces to the problem considered in Section 4.2. We observe, however, that we have not as yet specified the constant D_0 appearing in equation 4.1 and thus we can only compute the power in the diffracted waves to within a multiplicative constant. In comparing theoretical and experimental results, this constant will be chosen so that the two cases agree at some convenient point.

The holographic diffraction gratings that are described in this section were made using Kodak 649-f 3 1/4 x 4 1/4 x .040" glass film plates, and were processed in the same manner. The processing procedure was:

5 minutes in d-19 developer at 70 f

30 seconds in a 1.65 percent acetic acid stop bath

5 minutes in a fixing solution (757 cc paper fixer, 355 cc ammonium thiosulfate, and 3,030 cc distilled water)

30 minute rinse in distilled water.

Constant agitation was maintained throughout and the plates were air dried.

4.5.1 Orientation Sensitivity

We have seen in Section 4.2 that the amplitudes of the diffracted waves depends on how well the waves generated by the grains at various depths with the emulsion layer add in phase, and that this in turn is a function of the direction of the illumination wave. In this section we shall consider a specific grating, and we shall measure the power diffracted into the first-order waves as a function of the direction of the illumination wave. We will then use the analysis developed in Section 4.2 to compute the diffracted power, and then we shall compare the experimental and theoretical results.

The grating which will be considered was made using the apparatus discussed in Section 4.3, with the two exposing plane waves (with $\lambda = .6328\mu$) being incident on the film plate at $\theta_{io} = +15^{o}$

and $\theta_{io} = -15^{\circ}$, respectively. This produced a grating whose fringe planes are perpendicular to the emulsion surface ($\gamma = 0$) and whose fringe spacing d in the plane of the emulsion surface is equal to 1.223μ (equations 4.37 and 4.61).

The grating was then illuminated by a collimated laser beam ($\lambda = .6328\mu$) and the ratio of the power diffracted into each of the two first-order diffracted waves to the illumination power was measured as a function of the direction of the illumination beam. The propagation vector of the illumination wave was restricted to lie in the plane of the two original exposing waves (i.e., $\ell_{io} = 0$) in order that the wave vectors of the illumination wave and diffracted waves would all lie in the same plane.

The experimental results are shown in figure 4.3, where the power ratio is plotted as a function of θ_{io}, the angle of incidence (defined by equation 4.61) of the illumination wave prior to refraction at the emulsion air interface. Three experimental curves are shown, one for the zeroth order (the direct transmitted beam) and the other two for the $N = -1$ first-order diffracted wave - one with the emulsion side facing the illumination beam and the other with the emulsion side away. The $N = +1$ curves are not shown as they were simply the mirror images of the $N = -1$ curves with respect to the origin. In all three cases the illumination beam was linearly polarized with its electric field vector perpendicular to the plane of incidence.

Also shown in figure 4.3 is the theoretical curve, computed

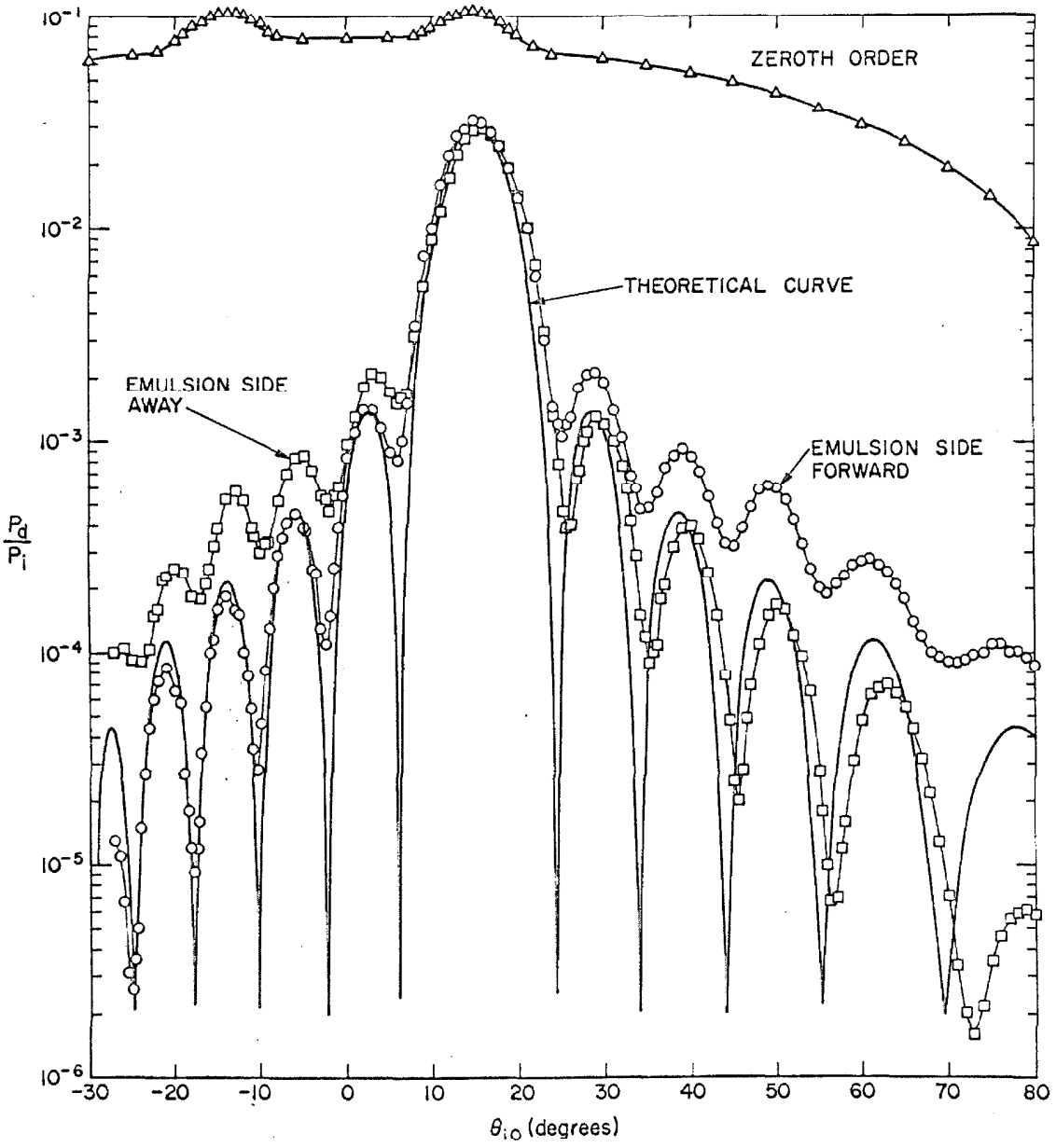


FIGURE 4.3 RATIO OF THE POWER DIFFRACTED INTO THE $N = -1$ FIRST-ORDER WAVE AND THE ZEROTH-ORDER WAVE TO THE POWER IN THE ILLUMINATION WAVE VS θ_{io}, THE ANGLE OF INCIDENCE OF THE ILLUMINATION WAVE PRIOR TO REFRACTION AT THE EMULSION-AIR INTERFACE. $\lambda = .6328\mu$, $d = 1.223\mu$, $\gamma = 0$.

using the analysis of Section 4.2. The quantity which was computed was p_d/p_i, as given by equation 4.75, which is the ratio of the total power in the diffracted wave under consideration to the total power in the illumination beam. The calculations were done numerically, using a digital computer and plotter, with points being computed every 10th of a degree over the range of θ_{io} of interest.

In order to carry out the calculations, it was necessary to assign a value to the constant C appearing in equation 4.75, as well as to specify α(the attenuation constant), T(the emulsion thickness), and η(the index of refraction of the emulsion layer). The constant C was chosen so the maximum value of the computed value of p_d/p_i was equal to the maximum measured value. The attenuation constant α was estimated from the attenuation suffered by the transmitted portion of the illumination beam, at $\theta_{io} = 0$, and found to be of the order of $.1\mu^{-1}$ for the grating under consideration. The emulsion thickness of the 649-f plates before processing is 15μ, but shrinkage occurs during the processing procedure, with the result that the emulsion thickness T is somewhat less then 15μ. The index of refraction of the emulsion layer was taken to be equal to that of the gelatin matrix of the emulsion layer, 1.52 (66).

The value of T used to generate the theoretical curve shown in figure 4.3 was 12μ. This value was determined by comparing a number of similar computed curves (with $T = 5,6, \ldots, 15\mu$) with the experimental data shown in figure 4.3. It was observed that the

effect of varying T was to change the number and width of the various maxima, with the width decreasing and the number of secondary maxima increasing with increasing T . The location of the central maxima did not vary with T to any significant extent.

The effect of α on the shape of the theoretical curve was also investigated, and it was found that the difference between $\alpha = 0$ and $\alpha = .1\mu^{-1}$ was slight, amounting to a slight filling in of the nulls and a slight suppression (less than 2db) of the secondary maxima. The difference in the region of the central maximum was negligible. As α is increased the filling in of the nulls and the suppression of the secondary increases, and becomes rather pronounced at $\alpha = 1\mu^{-1}$. The shape of the central maxima remains essentially unaffected as do the secondary maxima on either side of the central maximum.

An examination of figure 4.3 shows that the agreement of the theoretical curve and the experimental curves is quite good in the region of the central maximum. Outside this region the agreement is not quite as good but the general nature of the curves is the same. The difference between the case where the emulsion side faces the beam and the case where the emulsion side is away from the beam is considerable, and appears, perhaps, to be due to a decrease of grain density with depth, due perhaps to attenuation within the emulsion layer during exposure of the film plate (the amplitude attenuation constant for an undeveloped 649-f emulsion layer was measured and found to be $.022\mu^{-1}$) and /or a decrease of development activity with depth during the processing of the film plate. These factors are outside the scope of the

analysis presented here and will not be considered further.

The shape of the zeroth-order curve can easily be explained by attenuation within the emulsion layer and reflection losses at the interfaces, except for the two humps at $\theta_{io} = \pm 15^{o}$. These anomalies, which have also been observed by Leith et al (8), do not appear to be explained by the analysis presented here.

We observe that the maximum diffracted power occurs when $\theta_{io} = \pm 15^{o}$, depending on which diffracted order is being considered. These are the two angles of incidence of the original exposing plane waves. We shall consider this point in some detail shortly (Section 4.5.3), but before doing so we shall examine the polarization dependence of the ratio of the power diffracted into the first-order waves to the power in the illumination beam.

4.5.2 Polarization Dependence

In order to examine the polarization dependence of the power diffracted into the first-order waves, measurements were made with the electric field vector perpendicular to the plane of incidence and parallel to the plane of incidence. The results, for the case where the emulsion side is facing the illumination beam, are shown in figure 4.4. In making these measurements, the polarization was changed at each value of θ_{io} with a polarization rotator (figure 4.2).

The curves shown in figure 4.4 contain the effects of reflection losses at the different interfaces, which differ for the two polarizations. Figure 4.5 shows these same two curves with the re-

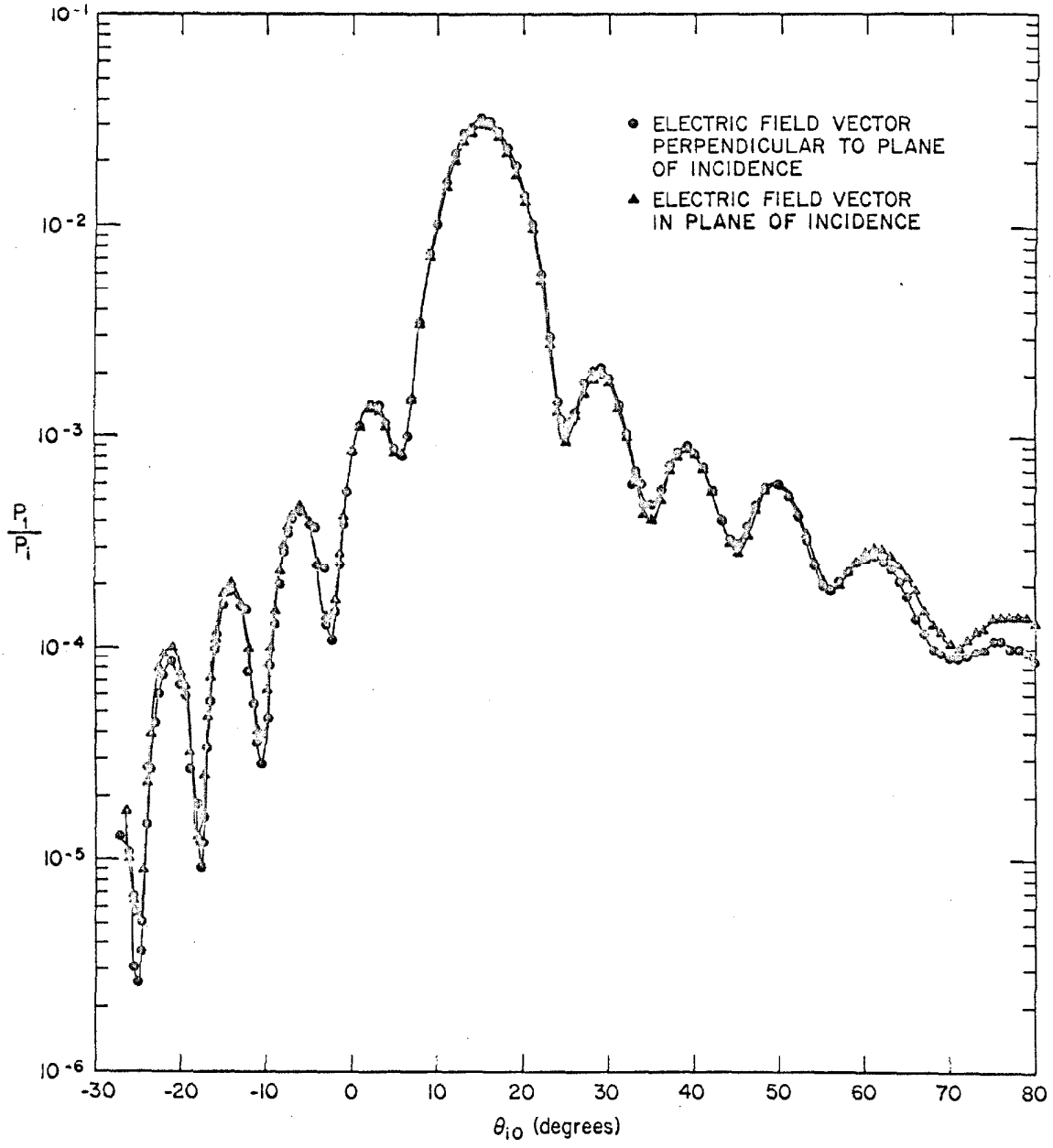


FIGURE 4.4 RATIO OF THE POWER DIFFRACTED INTO THE N = - 1 FIRST-ORDER WAVE TO THE POWER IN THE ILLUMINATION WAVE VS θ_{io}, THE ANGLE OF INCIDENCE OF THE ILLUMINATION WAVE PRIOR TO REFRACTION AT THE EMULSION-AIR INTERFACE. EMULSION SIDE TOWARDS BEAM, λ = .6328μ, d = 1.223μ, γ = 0.

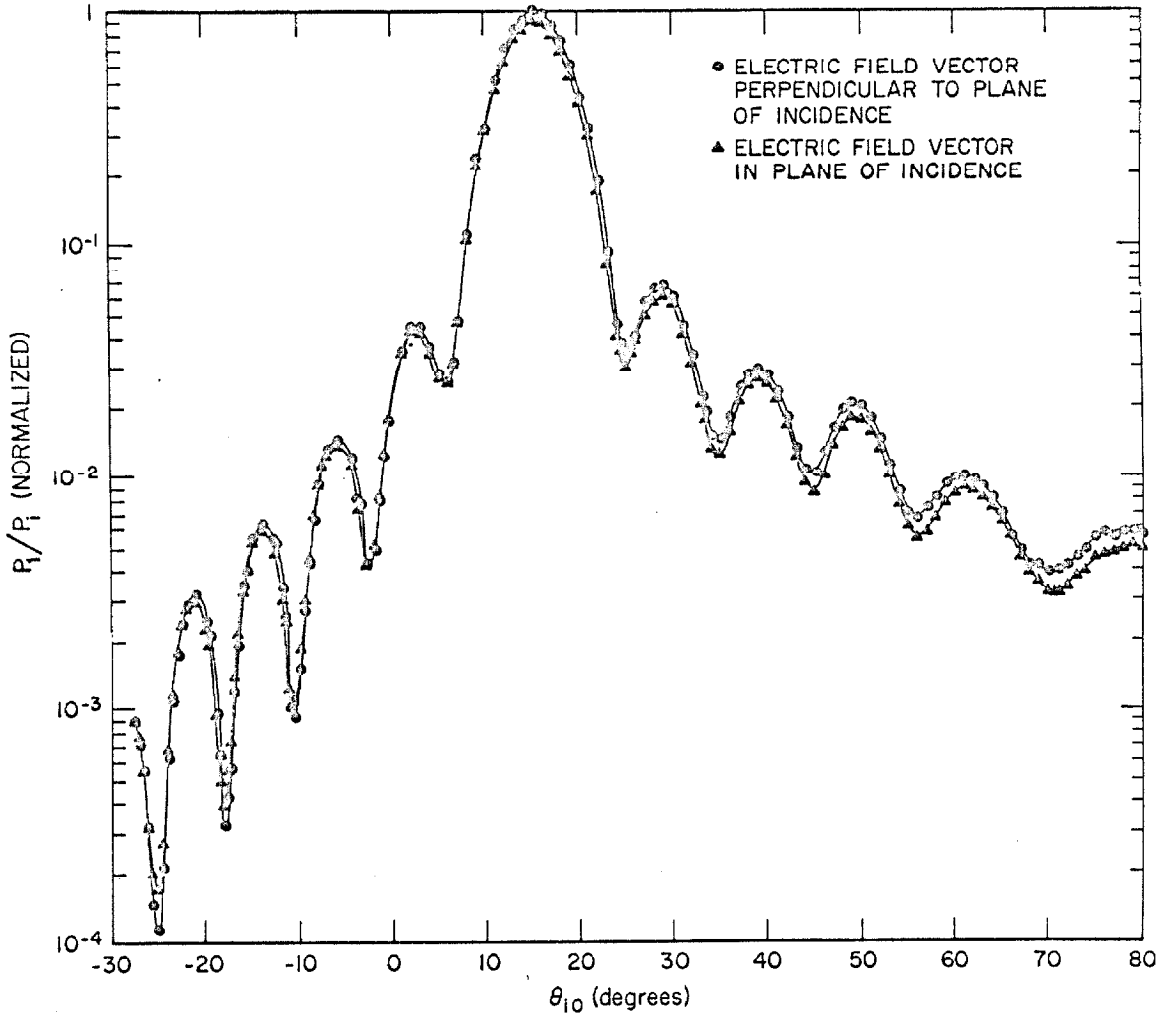


FIGURE 4.5 NORMALIZED RATIO OF THE POWER DIFFRACTED INTO THE N = - 1 FIRST-ORDER WAVE, WITH REFLECTION LOSSES DIVIDED OUT, TO THE POWER IN THE ILLUMINATION WAVE VS θ_{io} . EMULSION SIDE TOWARDS BEAM, $\lambda = .6328\mu$, $d = 1.223\mu$, $\gamma = 0$.

flection losses divided out, with the power ratio normalized to one at the maximum value. The conversion of the data was done with the aid of a digital computer, and amounted to dividing each data value by the transmissivity of the front and back interfaces, which were computed for each value of θ_{io} using the analysis outlined in Section 4.2.4. In making these calculations the index of refraction of both the emulsion layer and the supporting glass substrate was taken to be equal to 1.52, and multiple reflections were neglected.

We observe from figures 4.4 and 4.5 that the difference between the two polarizations is rather small for the particular case which we have considered. There is, however, an observable difference and it may well be that for a grating having a different fringe spacing d or a different thickness T that this difference may be greater. Burckhardt (53), in his paper dealing with the diffraction of a plane wave at a sinusoidally stratified lossless dielectric grating, presents a rigorous solution of the problem of computing the amplitudes of the diffracted waves when the fringe planes are perpendicular to the emulsion surface and when the propagation vector of the illumination wave has no component in the direction of the grating lines. His results indicate that the difference in the maximum diffracted power for the two polarizations increases with increasing T and decreasing d. Burckhardt also shows that in all cases the diffracted power is greatest when the electric field vector is perpendicular to the plane of incidence. We see from figures 4.3 and 4.4 that this is the case for the grating under consideration here, with the maximum

diffracted power being about 1/2 db greater for the case where the electric field vector is perpendicular to the plane of incidence. A similar result was obtained with the emulsion side away from the illumination beam.

We recall that in the analysis developed in Section 4.2, the polarization of the illumination wave entered in the analysis through its effect on the individual scattered waves produced by the various grains within the emulsion layer. We made the approximation of neglecting the angular dependence of these individual scattered waves, as well as the effect of different polarizations of the illumination wave. It appears from the experimental results shown in figures 4.3 and 4.4 that the neglect of the polarization dependence is quite a reasonable approximation, at least for gratings whose thicknesses and fringe spacings are of the order of those considered here.

4.5.3 Gratings with Inclined Fringes - Bragg Condition

We recall from Section 1.3.5 that when we illuminate a grating with one of the two original exposing plane waves, one of the two-first order diffracted plane waves has the same direction and phase (except for a phase factor of $-\pi$) as the other original exposing wave. We now wish to show that the analysis developed in Section 4.2 predicts that the power diffracted into that particular first-order wave is a maximum for this case.

The power per unit area in the first-order diffracted waves is given by equation 4.24, and is of the form

$$P = C' \frac{e^{-2\alpha T/n_d}}{n_d^2} \frac{(1 + e^{2aT} - 2e^{aT} \cos bT)}{a^2 + b^2} \quad , \tag{4.76}$$

where we have absorbed the various multiplicative constants in the new constant C' . It is clear that if $\alpha = 0$ (so that $a = 0$) and the factor $1/n_d^2$ is presumed to be slowly varying, then P will be a maximum where $b = 0$. The quantity b is given by equation 4.22 for the general case (no restrictions on the direction of the illumination wave other than $n_i > 0$) and by (equation 4.65)

$$b = \frac{2\pi}{\lambda_i \cos \gamma} [\cos (\theta_i - \gamma) - \cos (\theta_d - \gamma)] \tag{4.77}$$

for the case when the propagation vector of the illumination wave lies in the plane of the wave vectors of the two original exposing waves (i.e., when $\ell_{i_o} = 0$). It sufficies to consider the expression for b given above, as we are interested in the case where the illumination wave is the same as one of the two original exposing waves.

The fringe plane inclination angle γ is given by

$$\gamma = \frac{\theta_1 + \theta_2}{2} \tag{4.78}$$

where θ_1 and θ_2 are the angles of incidence (θ defined by equation 4.61) of the two original exposing plane waves after refraction at the emulsion air interface. Using the expression for γ given above, equation 4.77 becomes

$$b = \frac{2\pi}{\lambda_i \cos\left(\frac{\theta_1 + \theta_2}{2}\right)} \left[\cos\left(\theta_i - \frac{\theta_1}{2} - \frac{\theta_2}{2}\right) - \cos\left(\theta_d - \frac{\theta_i}{2} - \frac{\theta_2}{2}\right)\right]. \quad (4.79)$$

Let us now consider the case where we let $\theta_i = \theta_1$. We are interested in the first-order wave for which $\theta_d = \theta_2$ (equation 4.37, 4.47 and 4.61). We observe that substitution of $\theta_i = \theta_1$ and $\theta_d = \theta_2$ in equation 4.79 yields $b = 0$. The same result is obtained for $\theta_i = \theta_2$ and $\theta_d = \theta_1$.

Thus, apart from questions regarding the effect of α, a, and n_d in equation 4.76, we observe that the maximum diffracted power is obtained when we illuminate the grating with one or the other of the two original exposing waves, depending on which of the two first-order waves is being considered. The effect of α, a and n_d on the location of the maximum does not appear to be too significant for the gratings which we have considered, judging from the numerical calculations we have made, and will not be considered further here.

The above results are equivalent to the results obtained by Leith et al (8), and are in agreement with the results of van Heerden (13) and Burckhardt (53). As has been noted by a number of authors (8,9,53), the above results can be stated in terms of Bragg reflection from the planes of constant grain density. Bragg's law, or the Bragg condition, is simply a statement of the conditions on the angle of incidence, wavelength and distance between the parallel planes for which the waves reflected from the different planes add in phase. We

can express Bragg's Law in the form (67)

$$\sin \Psi_i = \frac{N\lambda}{2d'} \tag{4.80}$$

where Ψ_i is the angle the illumination wave makes with the parallel planes, λ is the wavelength of the illumination wave (within the structure), d' is the distance between the planes, and N is an integer.

It is straightforward to show that when θ_i equals either θ_1 or θ_2, the Bragg Condition is satisfied. Suppose we let $\theta_i = \theta_1$, then Ψ_i is given by

$$\Psi_i = \theta_1 - \gamma \tag{4.81}$$

or, using equation 4.78 for γ,

$$\Psi_i = \frac{\theta_1}{2} - \frac{\theta_2}{2} \quad . \tag{4.82}$$

In order to have Bragg reflection, the angle of reflection of the diffracted wave must equal the angle of incidence of the illumination wave. This requires that $\Psi_d = -\Psi_i$, where

$$\Psi_d = \theta_d - \gamma \quad . \tag{4.83}$$

We recall that when $\theta_i = \theta_1$, $\theta_d = \theta_2$ for the first-order diffracted wave of interest. Thus

$$\Psi_d = \theta_2 - \gamma \tag{4.84}$$

or, using equation 4.78

$$\Psi_d = \frac{\theta_2}{2} - \frac{\theta_1}{2} \quad , \tag{4.85}$$

and hence $\Psi_d = -\Psi_i$ as required. The distance d' between the planes of maximum grain density is given by (equation 1.62)

$$d' = \frac{\lambda}{2 \sin\left(\frac{\theta_1 - \theta_2}{2}\right)} \tag{4.86}$$

and hence, substituting the expression for d' given above in equation 4.80, we see

$$\sin \Psi_i = N \sin\left(\frac{\theta_1 - \theta_2}{2}\right) \quad . \tag{4.87}$$

An examination of the above equation and equation 4.82 shows that the Bragg Condition is satisfied, which is what we wished to show. A similar result is obtained when we let $\theta_i = \theta_2$.

A number of different holographic diffraction gratings were made having different fringe plane inclination angles γ . These gratings were made using the experimental apparatus described in Section 4.3. The beam spread angle $\theta_{10} - \theta_{20}$ (prior to refraction) was held constant at 30 degrees, and the film plate holder was rotated in steps of 10 degrees to provide a range of values of γ , as shown in Table 4.1.

The variation of the power diffracted into the $N = -1$ first-order was measured as a function of θ_{io} for each of these gratings, and the results are shown in Figure 4.6, with the power ratio P_1/P_i

ϵ_{10} (degrees)	θ_{20} (degrecs)	θ_1 (degrees)	θ_2 (degrees)	γ (degrees)	d (microns)	COMPUTED $\theta_{io\ max}$ (N = - 1)	MEASURED $\theta_{io\ max}$ (N = - 1)	SHIFT: MEASURED MINUS COMPUTED
15	-15	9.804	-9.804	0	1.223	15	15	0
25	- 5	16.143	-3.287	6.428	1.241	25	26.9	1.9
35	5	22.170	3.287	12.728	1.301	35	38.2	3.2
45	15	27.723	9.804	18.764	1.412	45	51	6
55	25	32.610	16.143	24.377	1.596	55	58	3
65	35	36.602	22.170	29.386	1.902	65	68°	3

TABLE 4.1 HOLOGRAPHIC DIFFRACTION GRATINGS WITH INCLINED FRINGES. PARAMETER VALUES COMPUTED USING AN INDEX OF REFRACTION OF 1.52.

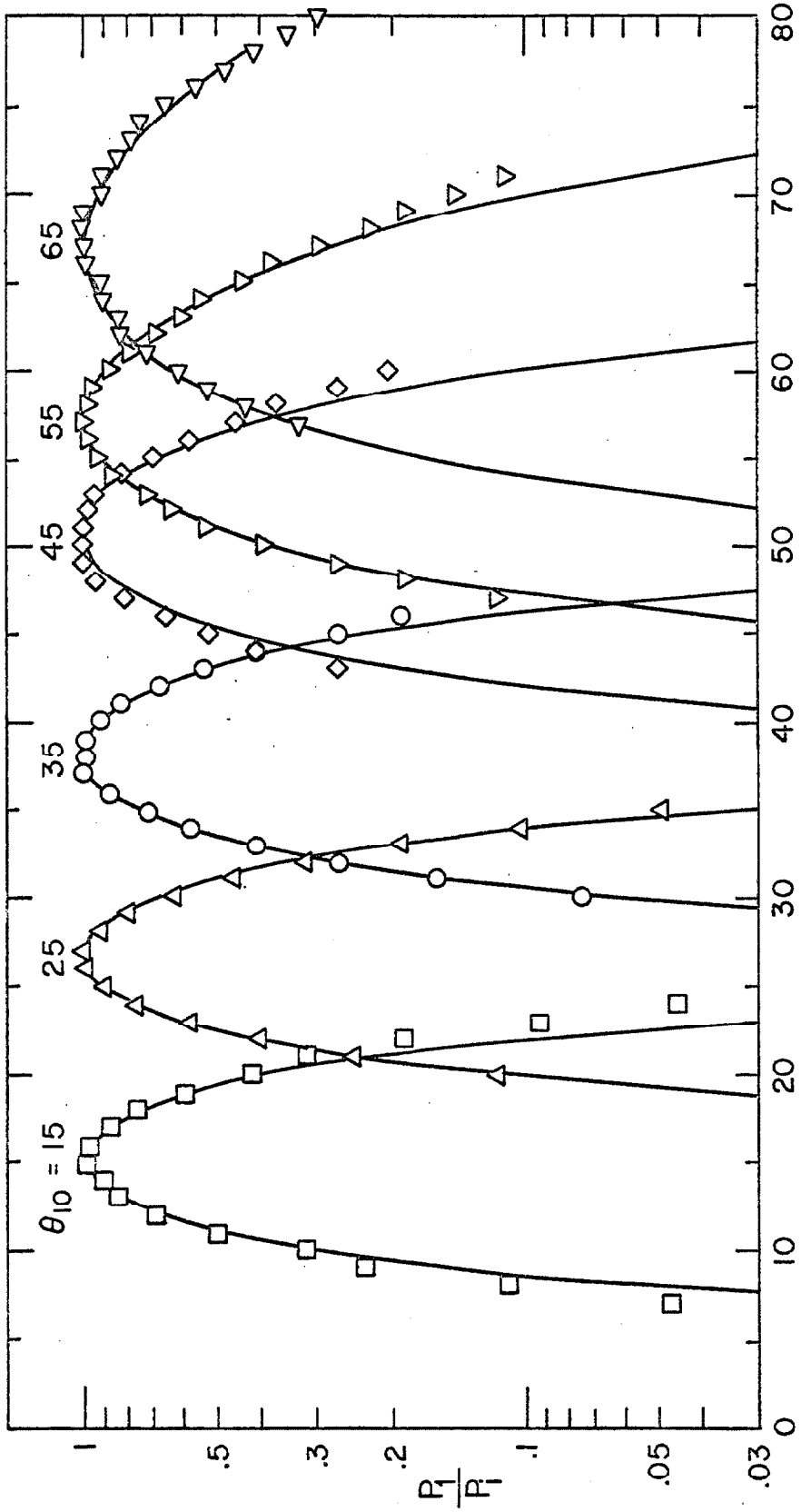


FIGURE 4.6 NORMALIZED RATIO OF THE POWER DIFFRACTED INTO THE N = - 1 FIRST-ORDER WAVE TO THE ILLUM. POWER VS θ_{io}, FOR THE SIX GRATINGS WHOSE PARAMETERS ARE GIVEN IN TABLE 4.1. EMULSION SIDE FOREWARD, λ = .6328, ELECTRIC FIELD VECTOR PERPENDICULAR TO PLANE OF INCIDENCE. THEORETICAL CURVES COMPUTED USING α = .1μ^{-1} AND T = 12μ.

being normalized to 1 at the maximum value for each grating. The solid curves are the corresponding theoretical curves (computed using equation 4.75 and the analysis outlined in Section 4.2.4), which have been shifted horizontally to yield the best fit, and where the constant C appearing in equation 4.75 has been chosen so that the maximum computed value of p_1/p_i is equal to 1. The amount of horizontal shift for each curve is shown in the last column of Table 4.1.

We observe that, apart from the observed shift of the location of the peaks, the theoretical curves are in excellent agreement with the experimental data. The broadening of the curves with increasing γ is due to the increasing value of d. The cause of the shift in the location of the peaks is not clear, but may be due to emulsion shrinkage effects. Another possibility is that the actual index of refraction of the emulsion layer might be different from 1.52, which was the value used in the computations. Our numerical computations show that changing the index of refraction changes the location of the peaks. Neither of these two effects was investigated in detail and we shall not consider these points further here.

4.5.4 Varying the Wavelength of the Illumination Beam

In the measurements discussed in the previous sections, the source was a helium-neon laser, and hence the value of λ was fixed at .6328μ (in air). In this section we shall describe measurements made with other values of λ. Experimentally, this was accomplished by replacing the laser source shown in Figure 4.2 with a low pressure

mercury arc lamp and using monopass spectral filters to isolate the various spectral lines. A small aperture and collimator were used to provide a collimated beam with which to illuminate the holographic diffraction grating. Using these techniques the ratio of the power diffracted into the $N = -1$ first-order wave to the illumination power was measured with λ (air) equal to .5790μ, .5460μ, .4358μ and .4046μ, for the grating used in Sections 4.5.1 and 4.5.2 ($d = 1.223\mu$, $\gamma = 0$).

The experimental results are shown in Figure 4.7a, where we have also plotted the curve for $\lambda = .6328\mu$. We observe that the peaks are shifted to smaller values of θ_{io} for shorter wavelengths, as predicted by the analysis of Section 4.2. Computer generated curves showed that the peaks occurred at $\theta_{io} = 15°$, 13.9°, 13°, 10.4° and 9.7° for $\lambda = .6328\mu$, .5790μ, .5460μ, .4358μ and .4046μ, respectively. Observation of the experimental curves shown in Figure 4.7a shows that the location of the peaks are within 1/2 degree of the predicted values.

We also observe that there is a difference in efficiency for the different wavelengths, the efficiency being greatest at $\lambda = .6328\mu$. Furthermore, there is a broadening of the curves with shorter wavelengths. This broadening is not predicted by the theory if one assumes that the index of refraction is the same for all wavelengths. However, it can be accounted for by using larger values of the index of refraction in the computations. This is illustrated in Figure 4.7b, where we have plotted the experimental curve for $\lambda = .4046\mu$ and theoretical curves computed using $\eta = 1.5$, 2.0 and 2.5. No shift of the peak

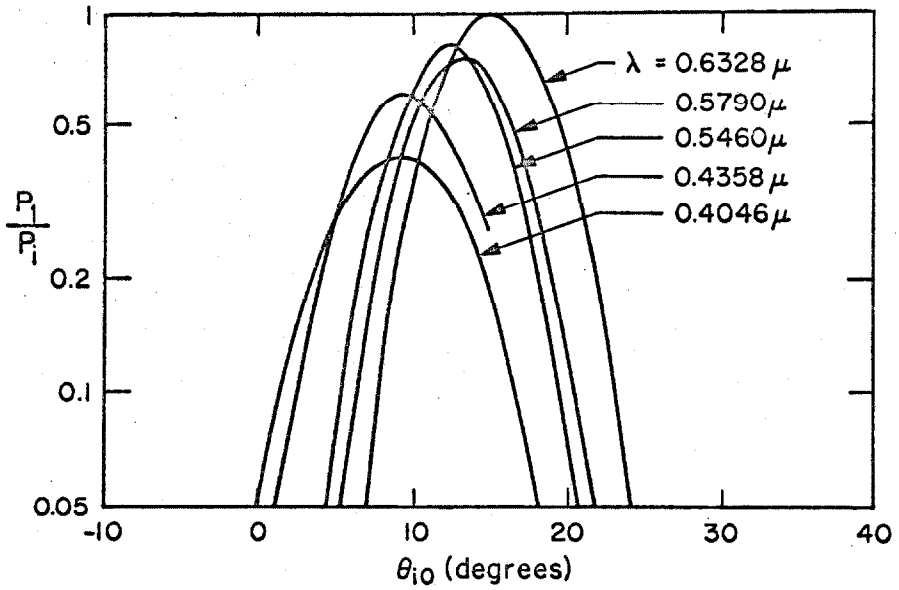


FIGURE 4.7a NORMALIZED POWER RATIO (N = - 1) VS θ_{io} ; d = 1.223μ, γ = 0, EMULSION SIDE FOREWARD.

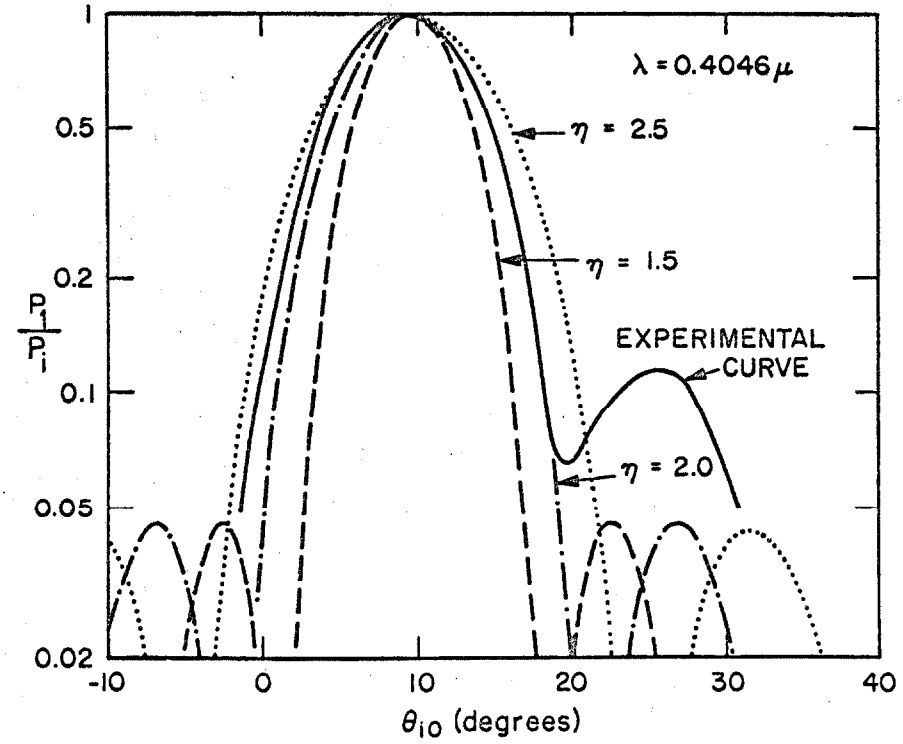


FIGURE 4.7b NORMALIZED POWER RATIO (N = - 1) VS θ_{io} ; d = 1.223 μ, γ = 0, EMULSION SIDE FOREWARD. THEORETICAL CURVES COMPUTED USING T = 12μ, $\alpha = .1\mu^{-1}$, AND THREE DIFFERENT VALUES OF η , THE INDEX OF REFRACTION OF THE EMULSION LAYER.

occurs when we vary n when the fringe planes are perpendicular to the emulsion surface ($\gamma = 0$). We see from Figure 4.7b that an index of refraction of the order of 2.0 provides a reasonable fit, which suggests the possibility that the film emulsion layer may be quite dispersive. We shall merely note this possibility here and shall not consider this point further, as it is outside the scope of our treatment of the problem.

4.5.5 Orientation Sensitivity - Grating Lines in the Plane of Incidence

In the previous sections, we have restricted our experiments and numerical computations to the case where the propagation vector of the illumination beam had no component in the direction of the grating lines (i.e., $\ell_{io} = 0$), in order that all the propagation vectors would lie in the same plane. The analysis developed in Section 4.2, however, is not limited to this case, and in order to verify the general validity of the analysis an experiment was performed where the wave vector had components in the direction of the grating lines.

What was done was to rotate the grating by 90° about the z axis from its normal position on the rotatable table shown in Figure 4.2, which results in the grating lines being parallel to the horizontal plane. Rotating the rotatable table then results in ℓ_{io} being varied, with m_{io} being equal to zero. With this change, the wave vectors of the illumination wave and diffracted waves no longer lie in the same plane, and it was necessary to modify the experimental apparatus so that the photomultiplier could be positioned to pick up the diffracted waves.

What was done was to add another degree of rotational freedom to the photomultiplier holder, so that in addition to a rotation about the vertical axis, rotation could be achieved about a horizontal axis which passes through the illuminated portion of the grating.

The experimental results are shown in figure 4.8 for the same grating that was used in Section 4.1 ($\gamma = 0$, $d = 1.223\mu$). The ratio of the power diffracted into the $N = +1$ first-order wave to the power in the illumination beam is plotted against ϕ, which is defined by

$$\ell_{io} = \sin \phi \quad . \tag{4.88}$$

Also plotted is the theoretical curve, for three different emulsion thicknesses. The agreement is best for $T = 11\mu$. The point $\phi = 0$ corresponds to the point $\theta_{io} = 0$ in figure 4.3 (emulsion side forward), and the observed difference in p_1/p_i is due to the fact that different portions of the grating were illuminated in the two cases. This was due to the design of the holder, and could not be avoided without cutting the film plate, which was not done.

We observe from figure 4.8 that the variation of p_1/p_i with ϕ is much smoother then with θ_{io}, and that the central peak and secondary maxima are absent. Our numerical calculations show that this same general behavior is to be expected for other gratings having different parameters, but that the general shape of the curve will vary considerably, and in some cases will be a minimum at $\phi = 0$ (this has been observed).

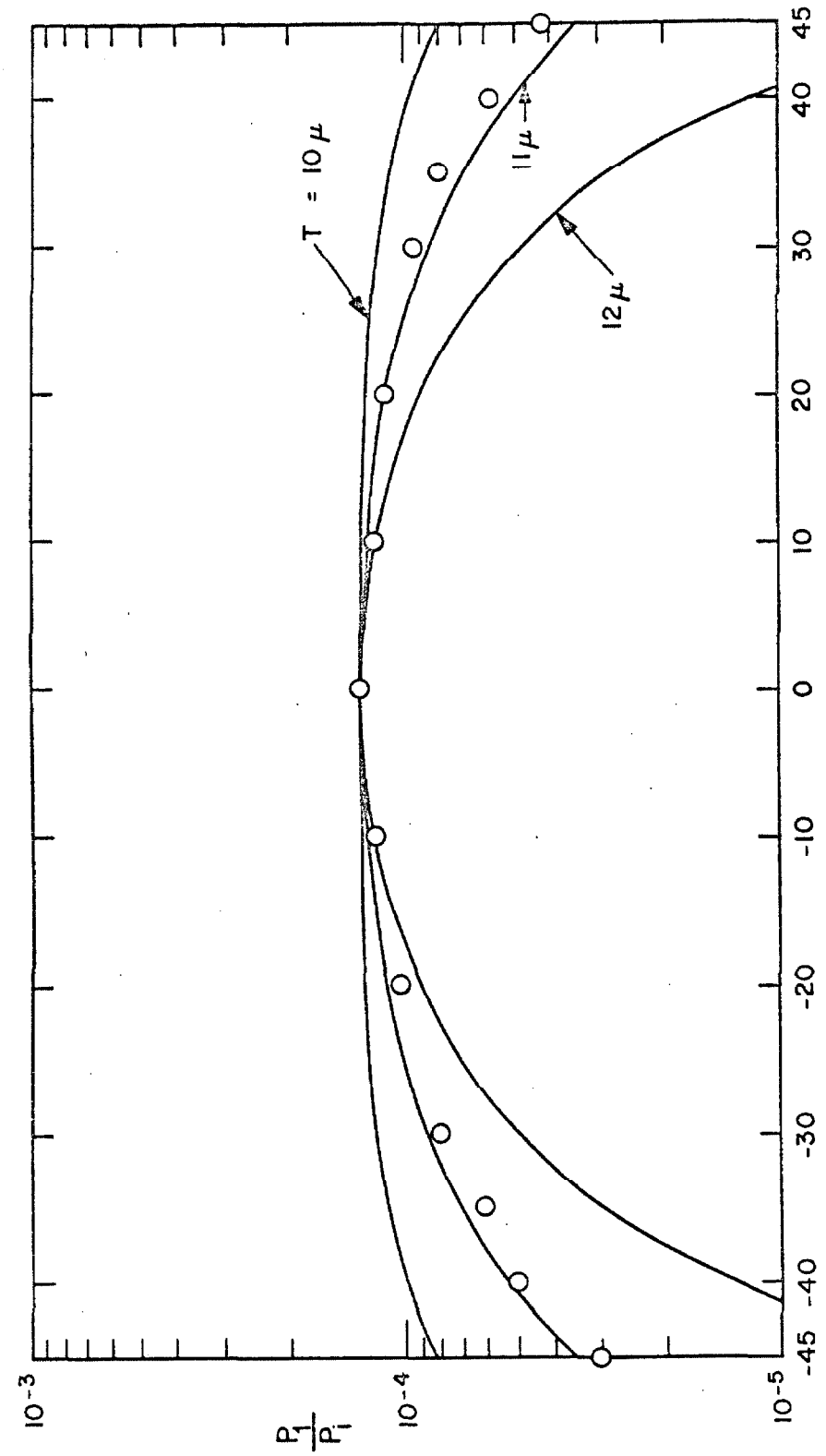


FIGURE 4.8 RATIO OF THE POWER DIFFRACTED INTO THE N = + 1 FIRST-ORDER WAVE TO THE POWER IN THE ILLUMINATION BEAM VS ϕ. EMULSION SIDE FOREWARD, ELECTRIC FIELD VECTOR PERPENDICULAR TO PLANE OF INCIDENCE, d = 1.223μ, γ = 0. THEORETICAL CURVES COMPUTED USING α=0.1μ^{-1} AND VALUES OF T AS SHOWN. λ = 0.6328μ.

4.5.6 Efficiency

We have seen in the previous sections that for a given grating, source, and diffracted order, there is a particular direction of the illumination wave for which the diffracted power is a maximum. This "optimum" direction of illumination, or "optimum illumination condition," is the same for similar gratings (same d, γ, T), but one finds that the diffracted power that is obtained under such conditions is not necessarily a constant, but may vary from one grating to the next, depending on the exposure and processing procedures used in producing the various gratings.

It is of interest to examine the effect of varying the exposure and processing procedures on the efficiencies of the different gratings. This was done experimentally, using gratings similar to the one discussed in Sections 4.5.1, 4.5.2, 4.5.4, and 4.5.5 ($d = 1.223$, $\gamma = 0$). What was done was to make a number of such gratings, where the exposure time was varied to provide a range of exposure. The efficiency of each grating (i.e., p_1/p_i for optimum illumination conditions) was then measured and plotted against the attenuation suffered by the zeroth-order beam (p_o/p_i at $\theta_{io} = 0$) and the results are shown in figure 4.9. Recalling that we estimate the attenuation constant α by

$$\alpha = \frac{1}{2T} \operatorname{Ln} \left(\frac{p_i}{p_o}\right) \tag{4.89}$$

We can view figure 4.9 as being equivalent to a plot of efficiency

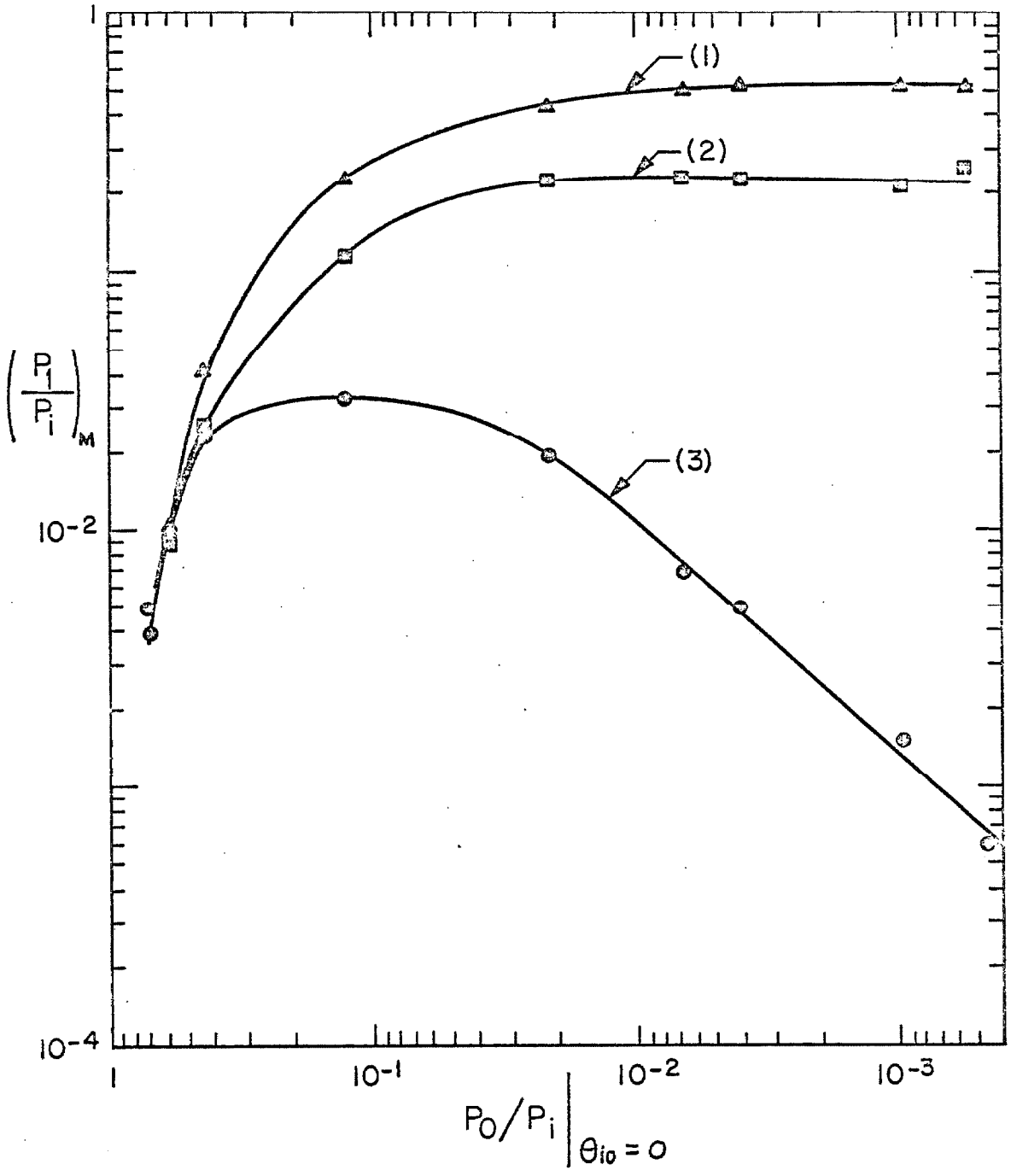


FIGURE 4.9 EFFICIENCY VS TRANSMITTANCE OF FILM PLATE AT $\lambda = 0.6328\mu$
(1) BLEACHED, STILL WET,
(2) BLEACHED, DRY,
(3) NOT BLEACHED.

vs α (where p_i is the power in the illumination beam, p_o is the power in the transmitted portion of the illumination beam, at normal incidence, and T is the emulsion thickness).

The lower curve in figure 4.9 corresponds to the case where the film plates are processed in the normal manner, as outlined in the beginning part of Section 4.2. This curve can be interpreted in a qualitative way with the aid of equation 1.9, the grain density equation. There appears to be two competing processes that occur as we increase the exposure. First of all, we increase the number of grains which contribute to the first-order diffracted waves which increasesthe power diffracted into the first-order waves. At the same time, however, we increase the total number of grains present which increases the attenuation of the emulsion layer. Initially, the first factor dominates and the efficiency increases rapidly with increasing exposure. The film response slowly begins to saturate and the coefficients of the other terms in equation 1.9 begin increasing faster then the term contributing to the first order diffracted wave. The efficiency then begins to level off and finally begins to decrease, and when the film response has become completely saturated, a further increase in exposure results only in a corresponding increase in attenuation, as evidenced by the straight line portion of the curve.

This description is supported by the upper curves shown in figure 4.9, which corresponds to the case where the gratings are bleached (using Kodak Chromium Intensifier). Here the attenuation has been removed and the saturation of the film is clearly evident.

It is interesting to note that the efficiencies of the bleached gratings are about a factor of two greater when the emulsion layer is still wet, following the bleaching process. This appears to be due to the chemistry of the bleaching process, rather than to the increased thickness of the emulsion layer due to swelling, as resoaking of the emulsion layer after it had dried did not result in an increase in efficiency. In addition, it was observed that bleaching of the very highly exposed plates resulted in rather severe light scattering, but that scattering was not a problem for the more lightly exposed plates.

The gratings used to generate the curves shown in Figure 4.9 were made using approximately equal power in the two original exposing waves. Gratings were made using unequal power in the two waves and they were found to be less efficient, as is expected.

4.6 Discussion

In this chapter we have extended the analysis of Chapter One to include the computation of the amplitudes of the first-order diffracted waves that are produced by the grains associated with a particular periodicity term in equation 1.9. The analysis was then applied to the case where the hologram was a holographic diffraction grating, for which a series of experiments were carried out, and the experimental and analytical results were found to be in good agreement.

We found that there are two optimum directions of illumination that maximize the power diffracted into the first-order waves, one direction being associated with each of these waves. These directions were found to be (apart from relatively small shifts that were observed experimentally) the directions of the two original exposing plane waves that were used to generate the grating. That is, if one illuminates the grating with one of the two original exposing plane waves, then the amplitude of the first-order diffracted wave, which corresponds to a reconstruction of the other original exposing wave, is a maximum.

It was seen that the sensitivity of the diffracted power to the direction of the illumination wave increases with increasing emulsion thickness and decreasing periodicity d, with the result that for a "thick" grating ($d/T \lesssim .2$) the amplitude of the first-order diffracted wave corresponding to the reconstruction of the original exposing wave will be much greater than that of the other first-order wave.

It is thus clear that for the case of a more general hologram, the best reconstruction of the original signal beam will be obtained when the illumination beam is the same as the reference beam. If the hologram is "thick," then the power diffracted into the reconstruction of the signal beam (i.e., the virtual image) will be much greater than that diffracted into the real image beam, when the illumination is done with the reference beam.

CHAPTER FIVE

DUPLICATION OF HOLOGRAMS

5.1 Introduction

The idea of duplicating holograms has been a subject of considerable interest for a number of reasons. First of all, if one has a "master hologram," then oftentimes one can produce copies of comparable quality without the need for the more elaborate apparatus required to make the original hologram. Holograms have been successfully duplicated in a number of laboratories (45), and indeed, it appears that at least part of the interest in the duplication of holograms arises from the fact that a certain amount of experimental research can be done in this area without the need for the somewhat specialized apparatus required for making holograms. This was the case here, where the initial experimental work in the field of holography (March 1965) consisted of duplicating a borrowed hologram using both a helium-neon laser and conventional light sources of different spectral width.

More recently, during the summer of 1966, a more comprehensive study of the duplication process was made, with particular emphasis being placed on the case where the hologram must be regarded as "thick," in the sense that the periods of the fringe patterns are small compared to the emulsion thickness. It was determined, from a careful study of the duplication of a very simple type of hologram, the holographic diffraction grating, whose properties were reported in a previous publication (9), that the duplication process should be

viewed as that of making a hologram of a hologram. Described in these terms, the various aspects of copying holograms can be treated in a straightforward and consistent manner.

Various aspects of the idea of making a hologram of a hologram have been considered by other researchers. F. B. Rotz and A. A. Friesem (47) demonstrated the interesting result that if one takes a hologram of the real image of a hologram, then the real image of the new hologram doesn't exhibit any of the pseudoscopic effects normally associated with the real image. In the experimental arrangement used by Rotz and Friesem, the film plate for the second hologram was located sufficiently far from the original hologram so that it was illuminated only by the real image field. The reference beam was provided in the normal manner. D. B. Brumm, in a recent publication (48), developed this idea further, and pointed out that one could effectively duplicate holograms in this manner, and that it was not necessary to separate the two film plates, as the zeroth-order beam can provide the reference beam, and either the real or virtual image beams can provide the signal beam.

5.2 Duplication of "Thick" Transmission Holograms

In this section we shall consider the duplication of transmission holograms where the spatial frequencies of the recorded interference patterns are sufficiently high such that periods of these patterns are small compared to the emulsion thickness. In such a case, as was discussed in Section 2.2.2, the variation of the grain density

with the depth must be accounted for, and the concept of average amplitude transmission becomes of questionable use. We shall consider the case of duplication of low spatial frequency holograms briefly in Section 5.6, where we shall treat them as a limiting case of the case considered here.

5.2.1 Duplication Process

The basic duplication process can be described in somewhat general terms as that of exposing a film plate to the field produced by illuminating a hologram in some manner. We would normally think of the film plate as being in close proximity to the hologram, but this need not always be the case. It is clear that the nature of the duplicate hologram will depend primarily on the nature of the field that exposes it. This, in turn, depends on the nature of the field used to illuminate the master hologram, the characteristics of the master hologram, and the location of the duplicate hologram film plate with respect to the master hologram. It is thus apparent that what is involved in a detailed description of the duplication process is the solution of the general problem of specifying the diffracted fields produced when a hologram is illuminated by some arbitrary field. Certain aspects of this problem were considered in the previous four chapters, and we shall apply the analyses, results, and conclusions contained therein to the solution of the problem at hand. In particular, we shall make frequent use of the material contained in Chapter Four, and will develop our treatment along the lines of the analysis con-

tained in Chapter One.

5.2.2 Production of an Exact Copy

It is clear that the duplicate hologram will be an exact copy of the original only if the exposing fields are identical to those used to produce the original hologram. This can, to a large extent, be achieved provided that the original reference beam was a single wave, preferably a plane wave (this requirement must be satisfied if the virtual image is to be an accurate reconstruction of the signal beam, as discussed in Chapter Three). Then, as we recall from Chapters One and Two, illumination of the master hologram plate with the reference beam will yield a reconstruction of the signal beam. This reconstruction of the signal beam, plus the transmitted portion of the illumination or reference beam, are then essentially the same as the original exposing fields.

There are, however, additional fields produced which correspond to the real image, second order images, and other fields as discussed in Chapter One. There is always a certain amount of background scattering also, as well as fields due to multiple reflections. All these additional fields can be neglected provided that their amplitudes are sufficiently small with respect to the two fields of interest. We saw in Chapter Four that this is usually the case for "thick" transmission holograms, provided that we illuminate with the reference beam. As was discussed there, this arises by virtue of the fact that when we illuminate with the reference beam, the waves scattered by the grains

at various depths within the emulsion layer that contribute to the virtual image add in phase, while those contributing to the other images do not. The result is that the amplitudes of the real and higher order images may be sufficiently small such that they can be neglected.

5.2.3 Effects of Varying the Geometrical Characteristics of the Illumination Wave

In the previous section we considered the case where the master hologram plate was illuminated with the reference beam, assumed to be a laser generated plane or spherical wave, of the same wavelength λ as that used to produce the master hologram plate. In this section we shall deal with the case where the illumination wave is still a laser generated plane or spherical wave of wavelength λ, but incident at a different angle of incidence or having a different radius of curvature, or both.

We shall find it convenient to deal with the problem using the description of the holographic process developed in Chapter One. This allows us to determine the effect of varying the geometrical characteristics of the illumination wave on the total field by a careful examination of the effects of changing the angle of incidence of a plane wave on a holographic diffraction grating, since we can consider the hologram to be composed of a "linear" sum of such gratings, as discussed in Chapter One. Looking at the problem from this point of view, we need consider only the case where the illumination wave is a plane wave, as at any given point we consider a spherical wave to be

approximated by a plane wave. We should bear in mind, however, that while a change in the direction of a plane wave illumination beam brings about the same change in angle of incidence of the illumination wave at all points on the film plate, this will not be the case for a spherical wave illumination beam.

It is clear that changing the angle of incidence of a plane wave used to illuminate a holographic diffraction grating will change the amplitudes and directions of the various diffracted waves. It is shown in Section 5.3, where the duplication of holographic diffraction gratings is considered in detail, that in general only the two first-order diffracted waves and the transmitted portion of the illumination wave need be considered. Furthermore, it is shown in Section 5.3.1 that although the directions of the first-order diffracted waves are changed by varying the angle of incidence of the illumination wave, the periodicity of the interference pattern generated by either of the two first-order diffracted waves and the transmitted portion of the illumination wave is a constant in any plane parallel to the plane of the master film plate. This constant is independent of the angle of incidence of the illumination wave and its wavelength, and is equal to the periodicity of the master holographic diffraction grating in the plane of the emulsion surface. This means that the periodicity in the plane of the emulsion surface of the duplicate hologram will be the same as that of the original (provided the duplicate film plate is placed in a plane parallel to that of the original during the duplication process). Since this is the periodicity that determines the

directions of the diffracted waves which are produced when the duplicate hologram is illuminated, this means that with respect to the directions of the diffracted waves produced, the duplicate gratings will be an accurate duplicate of the original.

These results, while derived for the special case of a holographic diffraction grating, can be extended to more general types of holograms. There is, however, one modification which must be made. We recall that with the holographic diffraction grating, illuminated by a plane wave, there was no need to specify the distance between the duplicate film plate and the master film plate. This was the case because the diffracted waves were all plane waves, and hence the interference pattern is the same in any plane parallel to the emulsion surface. With the more general type of hologram, however, we may have spherical waves, and while they may be considered as "locally" plane, the interference pattern may vary considerably with the distance from the master hologram plate. Thus we must add the additional constraint that the duplicate film plate be in close proximity to the master film plate, if the periodicity of the duplicate hologram is to be equal to that of the original.

Thus, if the above mentioned condition is satisfied, then we would expect, for example, that if we duplicate a pictorial hologram with a laser-generated plane or spherical wave, the reconstruction of the object produced by the duplicate hologram should look essentially the same as that produced by the original hologram, regardless of the geometry of the illumination wave used in the duplication process (with-

in reasonable limitations, of course). This was observed to be the case by Landry (46), who reported that either parallel or diverging laser light produces similar results in the copying of holograms.

Let us now go back and consider the effect of varying the geometrical characteristics of the illumination wave on the amplitudes of the various diffracted waves. As mentioned earlier, we are interested in the case where the emulsion layer of the master hologram plate must be regarded as "thick," and hence, as was seen in Chapter Four, the amplitudes of the various diffracted waves are highly dependent on the angle of incidence of the illumination wave. In addition, we recall from Chapter Four that the efficiency of a hologram depends on the ratio of the amplitudes of the signal and reference beams. For the case of the simple two-beam holographic diffraction grating, we saw that the optimum ratio was unity.

In the case where we are duplicating a hologram, we have essentially three beams to consider, the two-first order beams and the transmitted portion of the illumination beam (the zeroth order). We can think of each of these first-order beams as interfering with the zeroth order to yield a separate hologram encoded on the duplicate hologram film plate. The efficiency of each of these "holograms" depends on the ratio of the amplitude of the corresponding first-order beam to that of the zeroth order. In general, the amplitudes of these first-order beams will be considerably below the value required to give maximum efficiency, so that the higher the amplitude of either of

the two first-order beams, the higher will be the efficiency of its corresponding "hologram."

It may appear at first glance that the separation of the duplicate hologram into two "holograms" is somewhat artificial, since we have shown that both of these holograms yield identical diffracted images. A closer examination of the situation, however, will show that this separation is quite meaningful when the emulsion layer of the duplicate hologram film plate must be considered as "thick." In such a case, the complete spatial dependence of the interference patterns generated by the two first-order beams and the zeroth-order beam is of importance, rather than just the periodicity in the front surface plane of the emulsion layer.

It is clear that since the directions of the two first-order beams are quite different, the two corresponding interference patterns will also be quite different, even though they have the same periodicity in the plane of the emulsion surface. The net result is that when the emulsion layer of the duplicate hologram is "thick," the amplitudes of the waves diffracted by each of the two holograms encoded on the duplicate hologram film plate will be highly dependent on the angle of incidence of the wave used to illuminate the duplicate hologram, and that this dependence will be different for the two holograms. There will thus be two angles of incidence at which the "Bragg condition" is satisfied for any given diffracted order produced by the duplicate hologram, one for each of the two recorded interference patterns. This is shown experimentally to be the case in Section 5.3.1,

where a careful study was made of the duplication of a holographic diffraction grating.

We recall that the relative amplitudes of the two first-order beams are highly dependent on the angle of incidence of the beam used to illuminate the master hologram plate, and hence the efficiencies of the two holograms which form the duplicate hologram may be quite different. In fact, if the amplitude of one of the two first-order diffracted waves is much greater then that of the other, then the efficiency of the corresponding hologram will be sufficiently high compared with that of the other hologram such that only it need be considered. This is usually the case, as was mentioned in the previous section, when the master hologram plate is illuminated with its reference beam.

5.3 Duplication Experiments with Holographic Diffraction Gratings

We have seen that the nature of the duplicate hologram depends, as does any hologram, on the characteristics of the exposing field. The nature of this exposing field, of course, depends on the nature of the master film plate and the nature of the field used to illuminate it. In this section we will consider the case where the master hologram plate is a holographic diffraction grating, and where the duplication apparatus is as shown in figure 5.1. With this apparatus, we are able to illuminate the master hologram with a laser-generated plane wave at various angles of incidence. For reasons of experimental and computational convenience, we shall deal only with the case where the propagation vectors of the original exposing plane waves (which

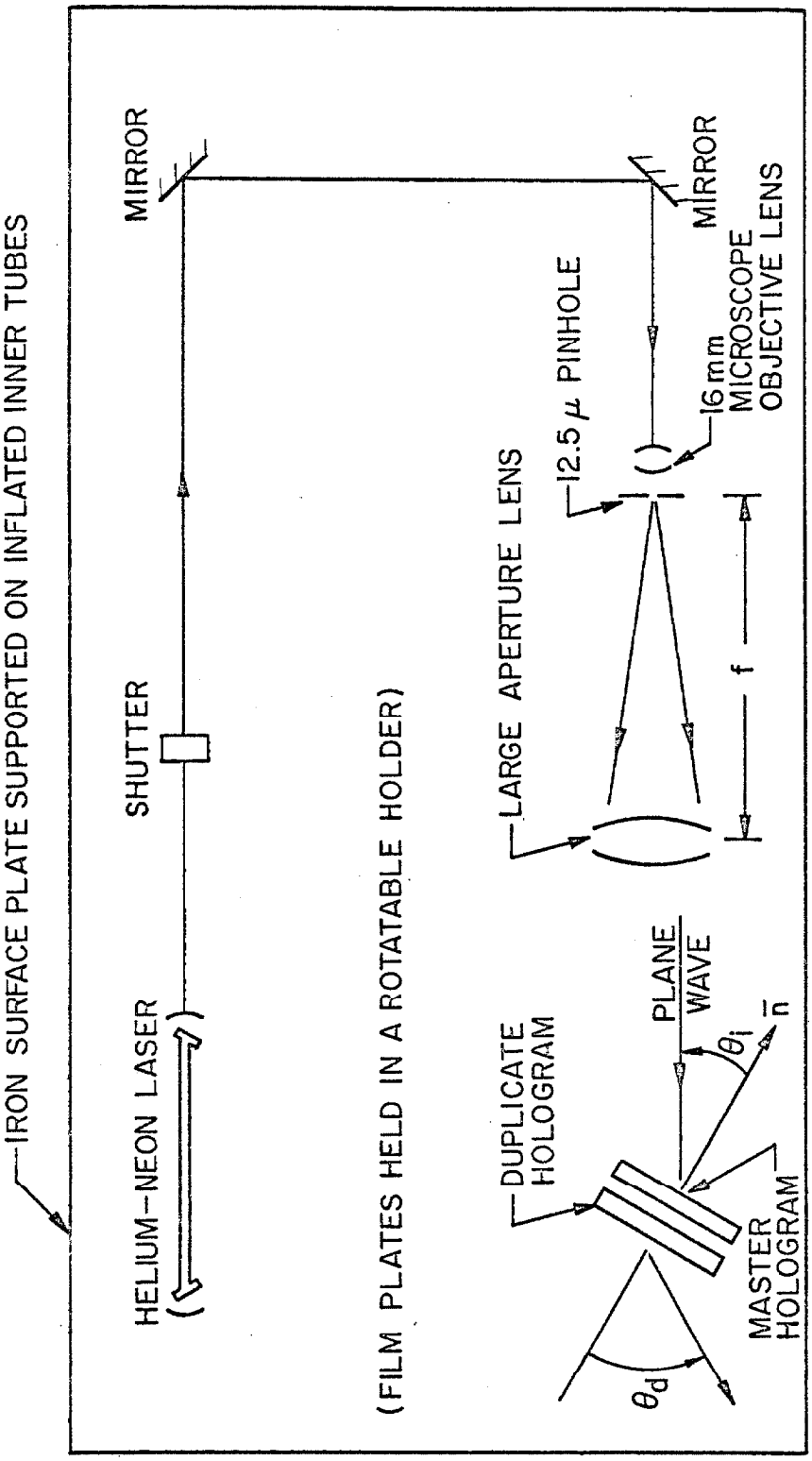


FIGURE 5.1 EXPERIMENTAL APPARATUS FOR DUPLICATING HOLOGRAMS

generated the master hologram), the propagation vector of the illuminating wave, and the normal to the film plate, all lie in the same plane (the horizontal plane). The grating lines are thus normal to the horizontal plane, and the y axis of the xyz coordinate system defined in Section 1.3.3 lies in the horizontal plane. The generalized grating equations given by equations 1.26 and 1.27 reduce to the ordinary grating equation, which is given by

$$\sin \theta_d = \sin \theta_i + \frac{N\lambda}{d} \tag{5.1}$$

where (equation 4.60)

$$\theta = \sin^{-1} (m) , \tag{5.2}$$

m being the y direction cosine of the wave under consideration. Equation 5.1 can be applied either to the fields inside or outside the film plate, provided the appropriate value of λ is used (this follows directly from equations 4.42 and 4.44). In this chapter we shall deal primarily with the fields outside the film plate, and we shall drop the subscript o on θ_i which was used in Chapter Four to designate θ_i prior to refraction at the emulsion air interface.

In the experiments that will be described in this section, the holographic diffraction gratings that are duplicated all have the same fringe spacing, $d = 1.223$ microns, and all have their fringe planes perpendicular to the emulsion surface ($\gamma = 0$). These gratings were produced with the apparatus shown in figure 4.1, with the two plane waves being symmetrically incident at $\theta_i = \pm 15^\circ$. The wavelength of the two plane waves was .6328 μ.

5.3.1 Varying the Direction of the Illumination Wave

The effect of varying the angle of incidence θ_i of the plane wave that illuminates the holographic diffraction grating (i.e., the master hologram) is twofold. First of all, the directions of the diffracted waves are a function of θ_i, as specified by equation 5.1, and second, the amplitudes of these diffracted waves are strongly dependent on θ_i, as discussed in Chapter Four. The power in each of the two first-order diffracted waves, for an illumination wavelength of $0.6238\ \mu$, is shown in figure 5.2 as a function of θ_i for the holographic diffraction gratings under consideration. In figure 5.2 we have used the values for the case where the emulsion side is away from the beam and the polarization is perpendicular to the plane of incidence, as this is the configuration used in the duplication of the gratings. The second-order diffracted waves, whose powers are of the order of two orders of magnitude smaller than those of the first-order waves in the range of θ_i of interest, are neglected. We will also neglect waves arising from reflections at the various interfaces.

The field which exposes the duplicate film plate thus consists of three plane waves, corresponding to the transmitted portion of the illuminating plane wave and the two first-order diffracted waves produced by the holographic diffraction grating. Recalling that the holographic diffraction grating was originally produced by two plane waves incident at $\theta_i = \pm 15^0$, it follows that letting $\theta_i = \pm 15^0$ in the duplication process will result in a duplicate hologram that is

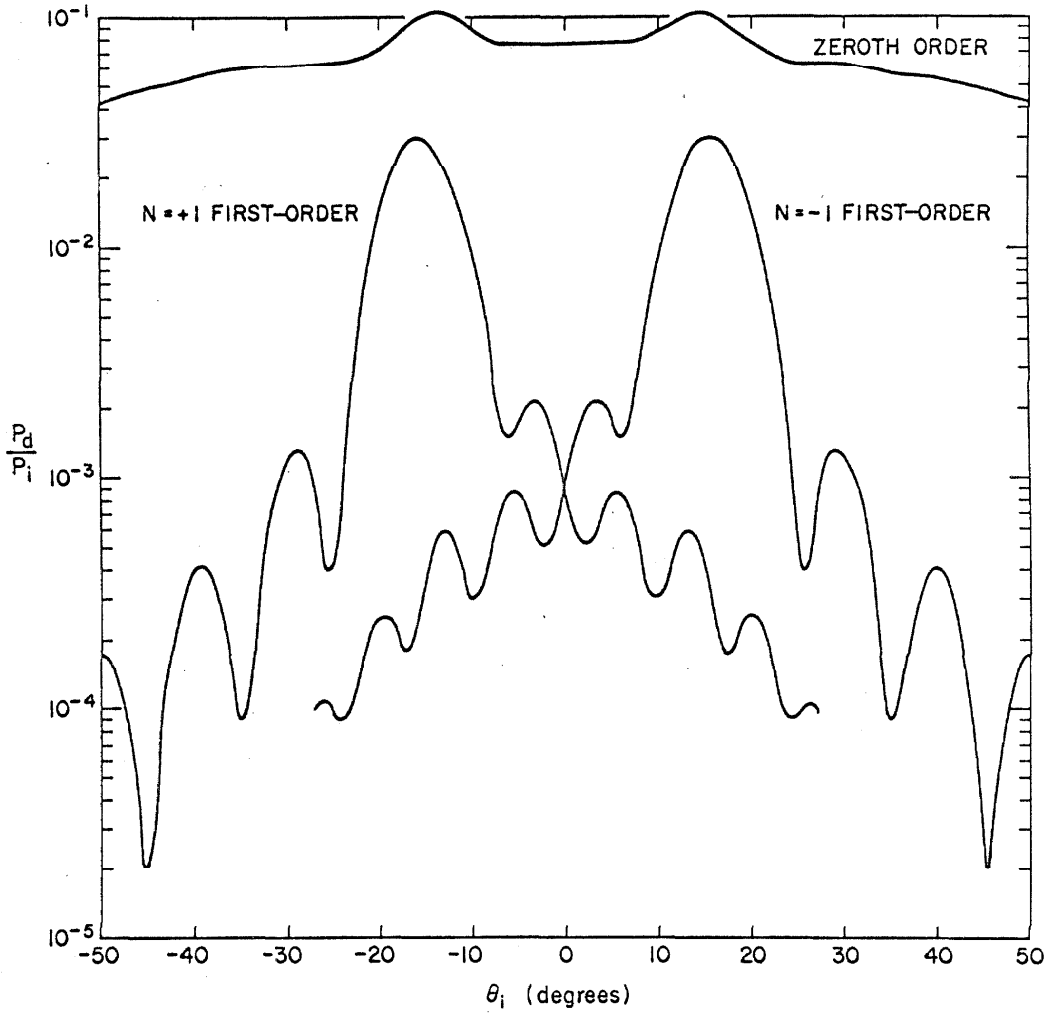


FIGURE 5.2 RATIO OF DIFFRACTED POWER TO INCIDENT POWER VS θ_i, THE ANGLE OF INCIDENCE OF THE ILLUMINATION BEAM PRIOR TO REFRACTION (EXPERIMENTAL CURVES - DATA POINTS NOT SHOWN). ELECTRIC FIELD VECTOR PERPENDICULAR TO PLANE OF INCIDENCE, EMULSION SIDE AWAY FROM BEAM, $d = 1.223$, AND $\gamma = 0$.

essentially the same as the master hologram, the holographic diffraction grating. This can be seen to be the case by letting $\theta_i = 15^o$ in equation 5.1. Then the three plane waves that illuminate the duplicate hologram film plate will consist of the transmitted portion of the illumination plane wave, which is incident on the duplicate film plate at 15^o, and the two first-order diffracted waves which are incident at -15^o and 50.7^o. The relative amplitudes of the two first-order diffracted waves are found from figure 5.2 with $\theta_i = 15^o$, and it is seen that the $N = -1$ first-order diffracted wave, which is incident on the duplicate hologram film plate at -15^o, is considerably stronger than the other first order. Thus, except for the considerably weaker beam at 50.7^o, the field that exposes the duplicate film plate is essentially the same as that which was used to produce the original holographic diffraction grating, and hence the duplicate hologram will be an accurate reproduction of the master hologram. A number of duplicate gratings were made using the apparatus shown in figure 5.1 (with $\theta_i = 15^o$) and they were observed to be very similar to the original, as expected. Both the original and duplicate gratings were made with Kodak 649f film plates, 3 1/4 x 4 1/4 x .040 size, and were processed in the same manner. The basis of comparison between the original and duplicate gratings was taken to be d, the fringe spacing, as determined by applying equation 1.19, and the fringe plane orientation, as determined by measuring the amplitudes of the two first-order diffracted waves as a function of θ_i. The various duplicate gratings exhibited different efficiencies, which depended on the rela-

tive amplitudes of the two principal exposing waves, as well as the total exposure. These factors will be discussed in more detail in the next section.

Let us now consider the case where the illumination wave has some angle of incidence other than $\theta_i = \pm 15^\circ$. This will mean that the field that exposes the duplicate film plate will consist of a different set of plane waves than in the previous case. Their directions and magnitudes are found from equation 5.1 and figure 5.2, respectively.

Let us consider a specific case, for example, $\theta_i = 0$. This case is of special interest, as it would be the configuration most likely to be used by someone who might view the duplication process as that of making a "contact print." For $\theta_i = 0$ the second-order diffracted waves are quite negligible, and we need only consider the three waves corresponding to the zeroth order (i.e., the transmitted portion of the illumination wave) and the two first-order waves. It is seen from figure 5.2 that these two first-order waves will have equal amplitudes when $\theta_i = 0$, but that this amplitude is considerably smaller than that of the primary first-order wave when $\theta_i = \pm 15^\circ$ (it should be kept in mind that power, rather than amplitudes, are plotted in figure 5.2).

The directions of the two diffracted first-order waves are found from equation 5.1, and are equal to $\pm 31.174^\circ$. The field that exposes the duplicate hologram film plate thus consists of a relatively large amplitude plane wave incident at $\theta_i = 0$ (the transmitted illumination wave) and two relatively small amplitude plane

waves (of equal amplitude) incident at $\theta_i = \pm 31.174^\circ$. The resulting interference pattern, which is what is recorded to form the duplicate hologram, thus consists of the interference patterns of each of the two first-order waves with the zeroth-order wave and the pattern corresponding to the interference of the two first-order waves with each other. Because of the relatively large amplitude of the zeroth-order wave, the first two of the above mentioned interference patterns will be the most important, and if the amplitudes of the first-order waves are sufficiently small compared with that of the zeroth-order wave, their mutual interference pattern can be neglected. In such a case the duplicate grating can be considered as the superposition of two gratings. It is clear that in the special case under consideration these two gratings have the same periodicity d' in the plane of the emulsion surface. This periodicity, or fringe spacing, can be computed using equation 1.19. We find

$$d' = 1.223 \text{ microns} \tag{5.3}$$

which is the same as that of the "master" grating.

The above result is not merely a coincidence for the special case considered, but is a consequence of a general rule which can be stated as follows:

The periodicity of the interference pattern which is generated by the zeroth-order wave and either of the two first-order waves, in any plane parallel to the plane of the master grating, is a constant independent of the illumination angle θ_i and the illumination wave-

length λ and this constant is equal to the periodicity d of the master grating.

This can be demonstrated with the aid of equations 1.19 and 5.1. Writing equation 1.19 in terms of the angles of incidence θ_{i1} and θ_{i2} of the two plane waves which generate the duplicate grating, the periodicity d' of the duplicate grating can be expressed in the form

$$d' = \frac{\lambda}{\left|\sin \theta_{i1} - \sin \theta_{i2}\right|} \tag{5.4}$$

where λ is the wavelength of the illumination plane wave. In the case under consideration θ_{i1} and θ_{i2} are the angles of incidence of the transmitted portion of the illumination wave and either one of the two first-order waves, respectively. The angle θ_{i1} is arbitrary and the angle θ_{i2} is specified by equation 5.1. That is

$$\sin \theta_{i2} - \sin \theta_{i1} = \pm \frac{\lambda}{d} \tag{5.5}$$

where d is the periodicity of the master grating. It is clear that substitution of equation 5.5 in equation 5.4 yields

$$d' = d \tag{5.6}$$

which is what we wished to demonstrate.

Returning to the duplicate grating formed with $\theta_i = 0$, it is clear that its periodicity in the plane of the emulsion surface is the same as that of the master grating, and hence as far as the

directions of the diffracted waves it produces when illuminated, it is equivalent to the master grating. The basic structure of the duplicate grating, however, is considerably different from that of the master grating. In the case of the master grating, the fringe planes are normal to the emulsion surface, and there is only a single set of them. The duplicate grating, on the other hand, has two sets of oppositely inclined fringes, corresponding to the interference patterns of the two first-order diffracted waves with the zeroth-order wave.

These two sets of fringe planes have the same periodicity in the plane of the emulsion surface, and hence the fields scattered by the grains associated with either set of fringe planes add in phase in the same directions. Thus, although there are two distinct sets of fringe planes, there will only be two first-order diffracted waves produced by the duplicate grating, and as was mentioned earlier, the directions of these diffracted waves are the same as for the original master grating.

The amplitude of either of the first-order diffracted waves produced by the duplicate grating is clearly equal to the sum of the amplitudes of the waves contributed by the two sets of fringe planes. Because these two sets of fringe planes are inclined in opposite directions, the relative contributions to the total amplitude of either of the two first-order diffracted waves varies greatly with the angle of incidence of the illumination wave. In fact, there are essentially two distinct ranges of θ_i where either one or the other of the two sets of fringe planes dominates and the other can be neglected. This

is verified by the experimental results shown in Figure 5.3, where the normalized power diffracted into the $N = -1$ first-order beam is plotted against θ_i. The apparatus used to make these measurements is shown in Figure 4.2, and a discussion of the experimental details for this case would be essentially the same as that given in Chapter Four, and hence will be omitted.

The curve plotted in Figure 5.3 is seen to consist of two similar curves, one centered at $\theta_i = 1°$ and the other centered at $\theta_i = 34.5°$, with a transition region in between. The curve centered at $\theta_i = 1°$ is essentially due only to the fields scattered by the grains in the fringe planes associated with the original exposing waves at $\theta_i = 0°$ and $\theta_i = -31.174°$. Similarly, the curve centered at $\theta_i = 34.5°$ is essentially due only to the fields scattered by the grains in the other set of fringe planes, which are associated with the original exposing waves at $\theta_i = 0$ and $\theta_i = 31.174°$. We observe that the maxima are shifted slightly from the values ($0°$ and $31.2°$) of θ_i that we would expect on the basis of Bragg reflection from the inclined fringe planes. A similar shift was observed and discussed in Chapter Four, in the section dealing with holographic diffraction gratings with inclined fringes (Section 4.5.3).

5.3.2 Efficiency

It is often of interest to compare different holograms on the basis of how "bright" a reconstruction can be obtained, with a given

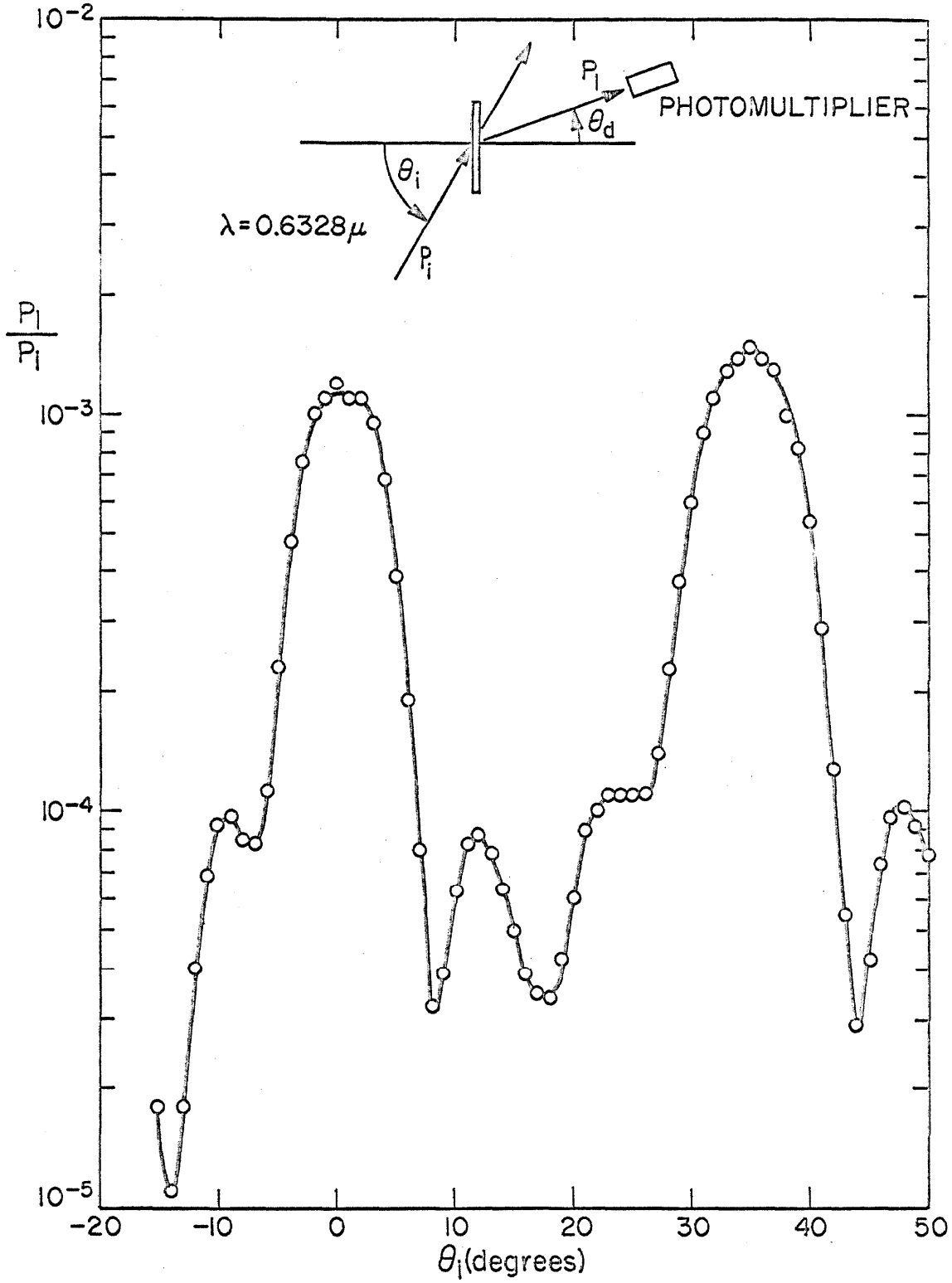


FIGURE 5.3 POWER RATIO FOR THE N = -1 FIRST-ORDER WAVE. DUPLICATE HOLOGRAPHIC DIFFRACTION GRATING MADE WITH $\theta_i = 0$.

illumination power. While a meaningful comparison may be difficult between two entirely different holograms, such is not the case when we have duplicates of the same master hologram. In the special case of the holographic diffraction grating, comparisons of this nature are particularly straightforward, as was seen in Chapter Four. A convenient basis of comparison is the efficiency, which we define as the ratio of the power diffracted into the primary first-order beam to the power in the illumination beam, when the "optimum" illumination angle θ_i is used. For example, consider the grating corresponding to figure 5.3. It has an efficiency given by

$$\varepsilon_d = 1.5 \times 10^{-3}$$

while the efficiency of the master hologram plate from which this grating was duplicated is seen from figure 5.2 to be

$$\varepsilon_m = 3 \times 10^{-2} \quad .$$

We observe that in this case the efficiency of the duplicate grating is much lower than that of the original. This is not always the case, however, as it is quite possible to have the efficiency of a duplicate grating exceed that of the original. Furthermore, the original grating which yields duplicate gratings of the highest efficiencies is not necessarily the one with the highest efficiency itself.

These statements follow directly from the results of Section 4.5.6 of the previous chapter, where we saw that the efficiency of a two-beam holographic diffraction grating depends primarily on the ratio

of the amplitudes of the two beams and on the total exposure. For a given amplitude ratio, there is some optimum exposure that will yield maximum efficiency. The maximum efficiency is greatest for an amplitude ratio of unity, and becomes less and less as the ratio departs farther and farther from unity.

Let us return to the case where we wish to duplicate a holographic diffraction grating of the type described in the previous section (i.e. d = 1.223μ, fringe planes normal to the emulsion surface). The variation of the power in the zeroth-order beam and in the two first-order beams with θ_i as shown in figure 5.2 for a particular grating of this type is typical, with differences between different gratings amounting to displacements of the zeroth-and first-order curves in the vertical direction.

In all cases (except possibly for bleached gratings) the amplitude ratio is closest to unity when $\theta_i = \pm 15^\circ$, and thus the illumination angle which gives the most accurate duplicate gratings also gives the most efficient duplicate gratings.

5.4 Duplication with a Non-laser Source

The first duplication of holograms was done by Gabor (5), using "conventional" or "non-laser" light sources. Indeed, the formation of a "positive" (or duplicate hologram from our point of view) was an important part of the holographic process as described by Gabor. Since the holograms which Gabor was dealing with involved fairly low spatial frequencies, the duplication process consisted of essentially making a "contact print" of the original hologram. Later, with the

invention of the laser, it became practical to make holograms having much higher spatial frequencies, and these, too, have been duplicated using conventional light sources in what appears to be a contact print type of process (45, 46). It was observed (45, 46), however, that it was quite important to have close contact between the master hologram and the duplicate film plate, otherwise no reconstruction can be obtained from the duplicate hologram which is produced. This and other effects are easily explained in terms of the description of the duplication process as that of producing a hologram of a hologram, rather than as the formation of a contact print.

5.4.1 Coherence Length and Path Length Differences

In the previous sections we assumed that the illumination of the master hologram plate was done with a laser generated plane or spherical wave. Thus, although the duplication process has been shown to involve the recording of the interference patterns generated by the two first-order beams and the zeroth-order beam, it was not necessary to take into account path length differences, as the coherence length of the illumination field could be considered as quite long. We shall now consider the case where we have a point source which has some finite, perhaps large, spectral width, and hence may have a very short coherence length. In such a case, if the path length differences exceed the coherence length of the source, then there will be no interference pattern, and hence no duplicate hologram produced.

There is, of course, no specific path length difference at

which point the interference pattern abruptly disappears, but rather, as discussed by Born and Wolf (65), the disappearance is gradual. The coherence length L_c is defined in such a way as to give a measure of the path length difference for which the interference pattern is essentially gone, as evidenced by a very low value of the visibility of the interference fringes, and is given by

$$L_c = \frac{\bar{\lambda}_o}{\Delta\lambda_o} \tag{5.7}$$

where $\Delta\lambda_o$ and $\bar{\lambda}_o$ are the spectral width and the mean wavelength of the source, in a vacuum. In comparing path length differences with L_c, we must use the optical path length rather than the geometrical path length.

In general, there will be a range of path length differences associated with a hologram of a fairly complex nature, and these path length differences will depend on a number of factors, which, to a certain extent, are under our control. In examining this problem, it is convenient to use the approach developed in Chapter One, as the path length differences associated with each periodicity of the master hologram plate can be computed separately, in a straightforward manner. This allows us to consider the effect of varying certain parameters in the duplication process independently of the details of any particular hologram.

Thus, let us consider the case where we wish to duplicate a holographic diffraction grating, of periodicity d, using a point source whose mean wavelength is $\bar{\lambda}_o$. We shall assume, for computa-

tional convenience, that the grating is illuminated with a collimated beam, and that the projection of the propagation vector of the illumination beam on the emulsion surface plane is perpendicular to the interference fringes. In this case, the propagation vectors of the diffracted waves lie in the same plane as that of the illumination wave, and their directions are specified by the simple grating equation

$$\sin \theta_d = \sin \theta_i + \frac{N\bar{\lambda}_o}{d} \quad . \tag{5.8}$$

In writing the above equation, we have assumed that $\Delta\lambda o$ is sufficiently small such that we can neglect the angular dispersion which it produces. If this is not the case, then we must take into account the range of λ and apply the grating equation separately for each wavelength.

The interference patterns which are of interest are the two which are generated by the interference of the first-order waves with the zeroth order. If the coherence length of the source is relatively short, then these two interference patterns will be localized in the immediate vicinity of the emulsion layer of the master hologram plate, the holographic diffraction grating. Thus, if one is to be able to record these interference patterns, and thus obtain a duplicate hologram, then the emulsion layer of the duplicate film plate must be placed within the region where the interference patterns exist. In practice, this is usually accomplished by placing the two emulsion layers in contact, and illuminating the master hologram plate from the

back side. It is not possible, however, to reduce the path length differences to zero, as the emulsions themselves have a finite thickness. Furthermore, it may not be possible to reduce the separation distance to zero, especially if an index matching fluid is placed between the two emulsion layers.

The optical path length difference arising from a separation of the two emulsion layers by an amount δ, neglecting emulsion thicknesses, is shown in Appendix VII to be given by

$$\Delta L = \eta\delta \left[(\tan \theta_i' - \tan \theta_d') \sin \theta_i' + \frac{1}{\cos \theta_d'} - \frac{1}{\cos \theta_i'}\right], \tag{5.9}$$

where η is the index of refraction of the medium between the emulsion layers and θ_i' and θ_d' are related to θ_i and θ_d by Snell's law:

$$\sin \theta_i' = \frac{1}{\eta} \sin \theta_i \tag{5.10}$$

and

$$\sin \theta_d' = \frac{1}{\eta} \sin \theta_d \quad . \tag{5.11}$$

The values of θ_d, which are of interest, correspond to $N = \pm 1$ in equation 5.8, and are specified once the angle of incidence θ_i of the illumination wave is specified.

We observe that the path length difference, while being proportional to δ, is also a function of η_e, θ_i, η, $\bar{\lambda}o$, and d. The dependence of ΔL on these quantities was investigated numerically and some typical results are shown in figure 5.4 (for $N = -1$, $\eta = 1.0$).

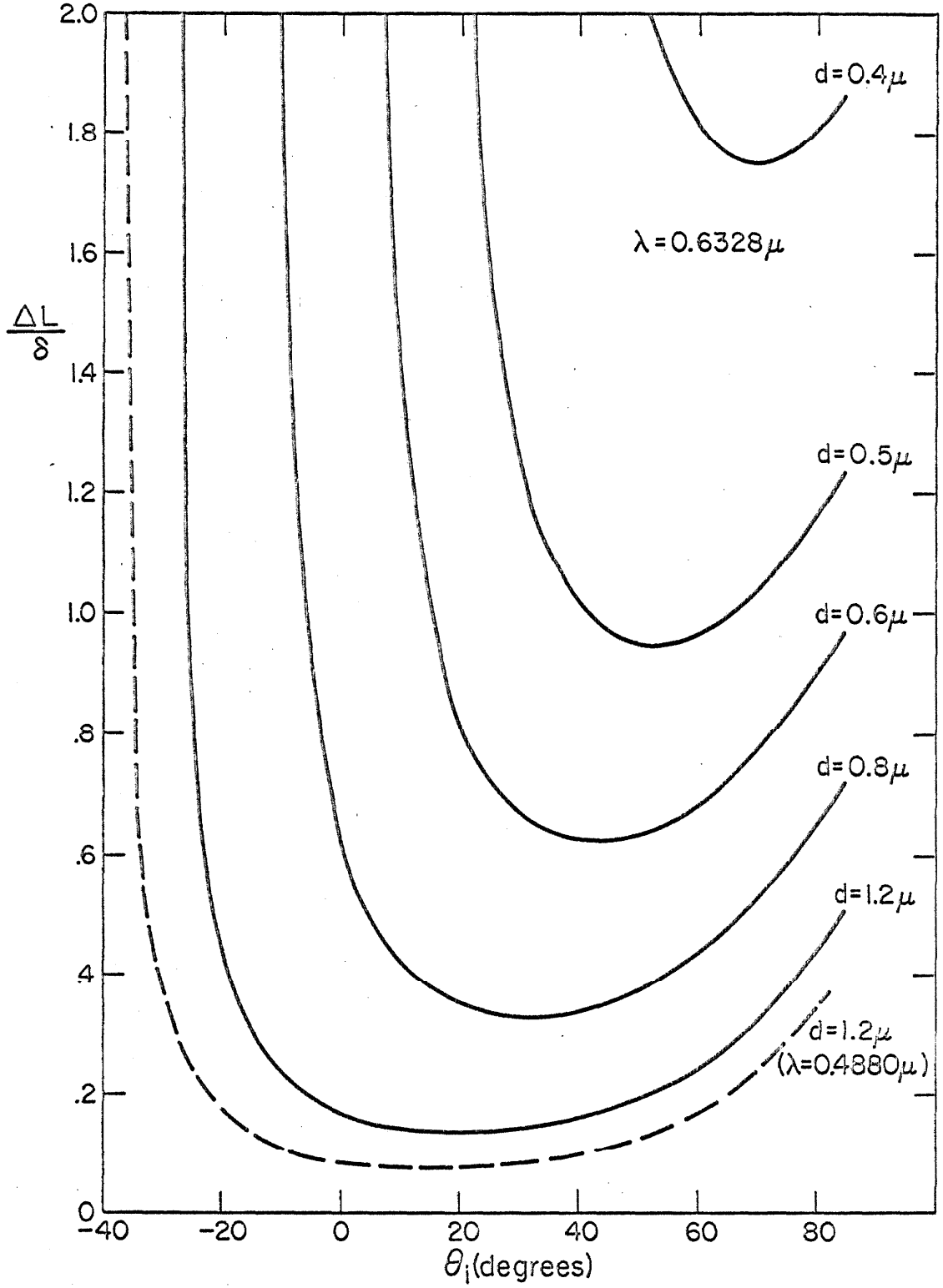


FIGURE 5.4 NORMALIZED PATH LENGTH DIFFERENCES FOR THE N = -1 FIRST-ORDER WAVE AND THE ZEROTH-ORDER WAVE.

The corresponding curves for $N = +1$ are found by replacing θ_i by $-\theta_i$. We observe that for a given film emulsion separation distance δ higher spatial frequencies (smaller d) and longer wavelengths yield larger values of ΔL. In addition, if d is small (of the order of $\bar{\lambda}_o$) then ΔL is strongly dependent on θ_i. It is seen from figure 5.4 that (for $N = -1$) $\Delta\ell$ increases very rapidly as θ_i approaches some minimum value, which is a function of $\bar{\lambda}_o/d$. This minimum value of θ_i corresponds to the smallest value of θ_i from which a solution for θ_d exists (equation 5.8 with $N = -1$), and thus is the solution of

$$\sin \theta_i = \frac{\bar{\lambda}_o}{d} - 1 \quad . \tag{5.12}$$

Similarly, for $N = +1$ there will be a maximum allowable value of θ_i which is specified by

$$\sin \theta_i = \frac{\bar{\lambda}_o}{d} + 1 \quad . \tag{5.13}$$

We observe that if $d < \bar{\lambda}_o$ then there will be no value of θ_i for which both first-order diffracted waves exist, and neither first order will exist for $\theta_i = 0$. Thus, if the illumination wave is brought in at normal incidence (which is the standard procedure for making a "contact print"), then it will be possible to obtain a duplicate hologram only if $d > \bar{\lambda}_o$. This is independent of the coherence length of the source.

On the other hand, if d is large compared with $\bar{\lambda}_o$ then,

for $\eta = 1.0$,

$$\Delta L \simeq \frac{\delta}{2}\left(\frac{\bar{\lambda}_o}{d}\right)^2 \qquad (5.14)$$

and we see that the path length differences involved with either first order decrease very rapidly with increasing d. It is thus clear why holograms involving low spatial frequencies (less than 200 lines/mm) are relatively easy to duplicate using what appears to be a "contact print" process, as reported by Vandewarker and Snow (49).

5.4.2 Early Experiments

As mentioned in section 5.1, the initial experimental work in the field of holography which was done here consisted of duplicating a borrowed hologram,* the subject of the hologram being a model train. The apparatus which was used is shown in figure 5.5, and consisted of a source, a collimating lens system, and a photocopy frame which was used to hold the two film plates in close contact. The illumination beam was incident at $\theta_i = 0$ in all cases, and no special vibration elimination techniques were used.

Both the master hologram and the duplicate film plates were Kodak type 649f film plates, 4" x 5" x .040 size. The master hologram was apparently made without the use of spatial filtering in the reference beam, as it exhibited the characteristic rings and swirls of diffraction patterns caused by dust on the elements of a coherently

---

* The hologram was borrowed from Ivan Courtwright of Spectra Physics Corporation, and the experimental work was done with the assistance of Milton Chang.

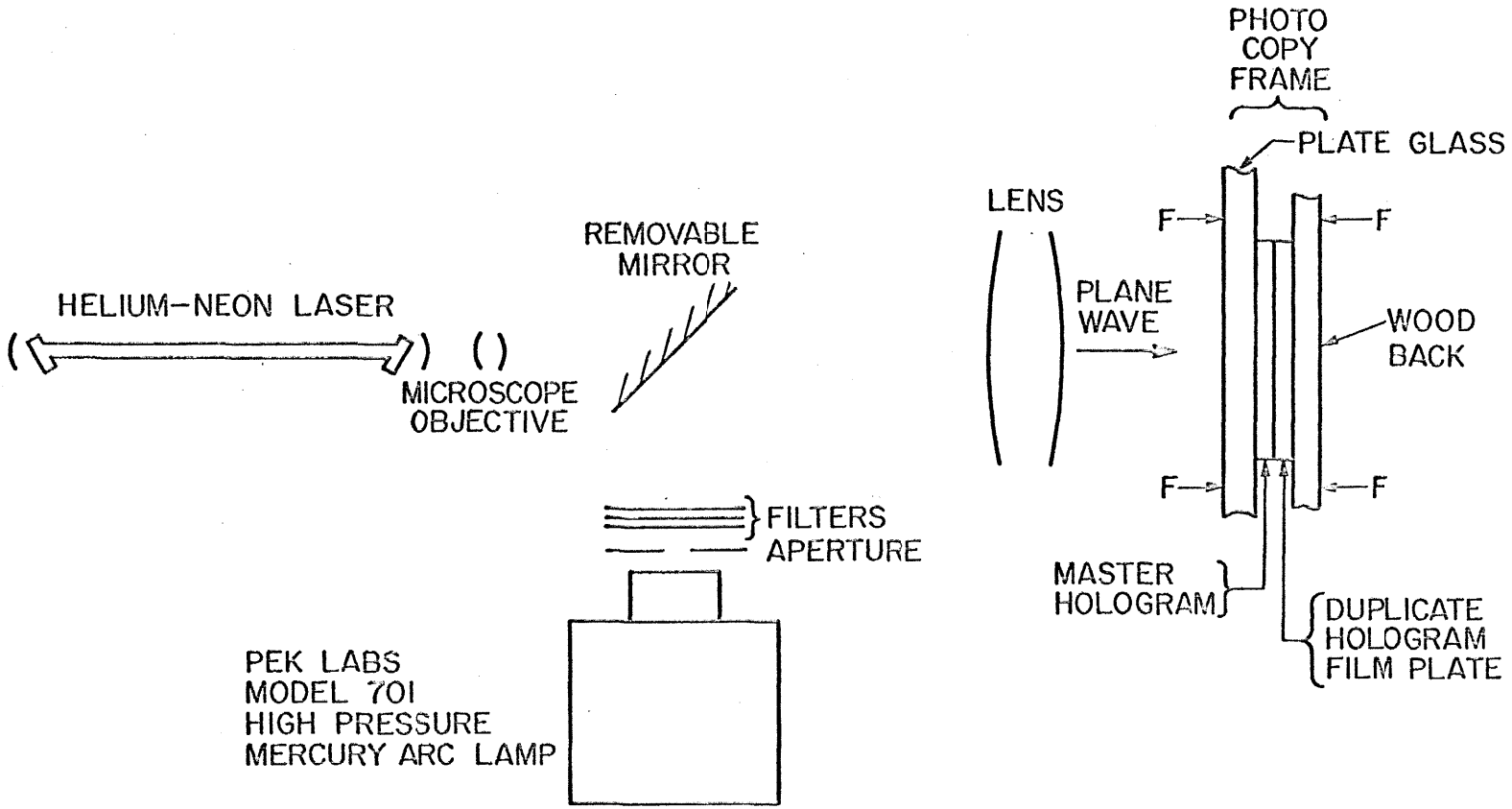


FIGURE 5.5 DUPLICATION APPARATUS

illuminated optical system. This was, in fact, somewhat of an advantage, because these large fluctuations in the transmission of the master hologram made the duplication process less sensitive to the total exposure.

A number of different sources were used, the first being a helium-neon laser operated at 6328A°. A number of duplicate holograms of a quality comparable to that of the master hologram were obtained with this source, both with the emulsions in contact and with the duplicate film plate turned around (which provided a spacing of .040" between the two emulsions).

The other sources consisted of a high pressure mercury arc lamp (PEK LABS MODEL 701) used with a variety of filters, as noted below:

(a) Spectrolab No. 2412 (7A° wide at 6843A°)

(b) Spectrolab No. 1709 (100A° wide at 6328A°)

(c) Corning 2-63 and 1-69 (band pass 5900A° to 9000A°)

The coherence lengths of (a) and (b) are found from equation 5.7 and are equal to 670 microns and 40 microns, respectively. To compute the coherence length of source number (c) we must take into account the fact that the 649f type emulsion is only sensitive out to a wavelength of about 7000A°, which would make $\Delta\lambda$ in this case equal to 1100A°, and hence $\ell_c = 4\mu$ for source (c).

All three of the above sources yielded duplicate holograms of a quality comparable with that of the original hologram when the

emulsion layers of the master hologram and duplicate hologram were placed in contact (and held there by the spring loaded photocopy frame). It is thus evident that there is little difficulty in reducing the separation distance between the two emulsion layers to a value sufficiently small such that the duplication can be done with conventional sources.

5.5 Duplication of Reflection Holograms

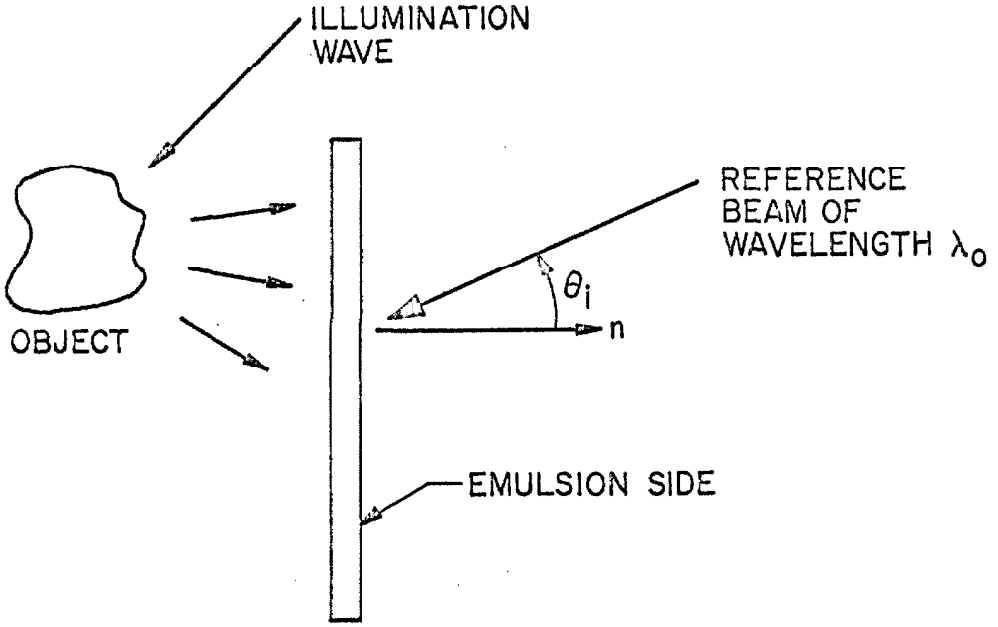
In this section we will consider the case where one wishes to duplicate a reflection hologram. It is clear, from the analysis and discussion of reflection holograms in Chapter One, that it would be completely meaningless to talk about making a "contact print" of such a hologram. However, if one views the duplication process as that of "making a hologram of a hologram," it is quite straightforward to demonstrate that a duplicate reflection hologram can indeed be produced. What is required, as is the case with transmission holograms, is to illuminate the duplicate film plate with essentially the same field as was used to produce the master hologram. This can be done by illuminating the master hologram plate so that it yields a reconstruction of the original signal beam, and then placing the emulsion layer of the duplicate film plate in the region where the interference pattern generated by the illumination wave and the reconstructed signal beam exists.

In the case where we are duplicating a transmission hologram, this region exists on the side of the master film opposite that which

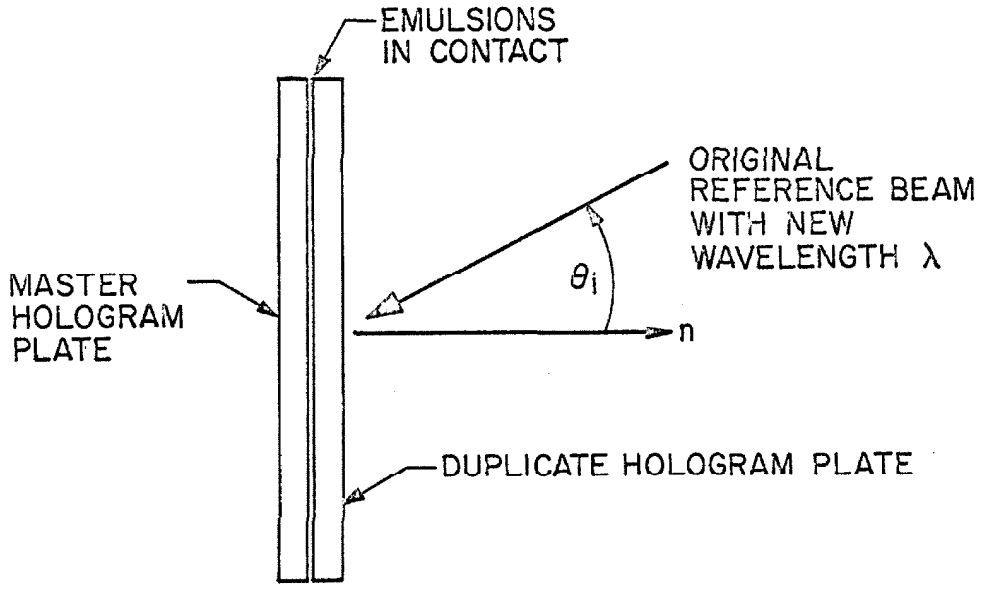
is first illuminated by the illumination beam, and hence the duplicate film plate is placed "behind" the master hologram plate, as if making a contact print. With a reflection hologram, however, this region exists in front of the master hologram plate, and hence we must place the duplicate film plate between the master hologram plate and the illumination beam, as shown in figure 5.6b. The illumination wave thus passes through the emulsion layer of the duplicate film plate first, prior to striking the master hologram plate. The transmitted portion of the illumination wave then illuminates the master hologram plate, producing a reconstruction of the signal beam in reflection, which then illuminates the duplicate film plate. An examination of the situation shows that if the illumination wave is essentially the same as the original reference beam, and if it produces a reconstruction of the original signal beam, then the field which exposes the duplicate film plate is essentially the same as that which produced the master hologram, and hence a duplicate hologram will be obtained.

5.5.1 Source Requirements

We recall from the discussion of reflection holograms given in Chapter One that there are a number of fundamental differences between reflection holograms and transmission holograms and these differences will be reflected in the duplication process. We recall that for a transmission hologram, a change in the wavelength of the illumination wave doesn't preclude the production of a duplicate hologram, as a reconstruction can usually be obtained over a wide range of



(a) PRODUCTION OF A MASTER REFLECTION HOLOGRAM



(b) DUPLICATION OF A REFLECTION HOLOGRAM

FIGURE 5.6 PRODUCTION AND DUPLICATION OF A REFLECTION HOLOGRAM.

wavelengths with a transmission hologram. With the reflection hologram, however, there is only a narrow band of wavelengths for which a reconstruction of the signal beam can be obtained. This means that the illumination wave used in the duplication process must have a wavelength within this same band.

5.5.2 Emulsion Shrinkage Effects

In the absence of any shrinkage of the emulsion layer during the processing of the master hologram film plate, the center wavelength of the reflection band of the master hologram will be at about the same wavelength as that of the original exposing field. If there is emulsion shrinkage, then the reflection band will be shifted toward shorter wavelengths. Fleisher et al (27), in an article dealing with an optically accessed memory using the Lippmann process, have reported recording standing wave interference patterns using Kodak 649-f film plates in which emulsion shrinkage shifted the reflection band from 5461A° to 4450A°. Upatnieks et al (24), who also used Kodak 649-f film plates, report similar large shifts in the reflection band of reflection holograms, stating that a reflection hologram made with red light requires green light for the reconstruction. They also note, however, that emulsion shrinkage can be reduced considerably by eliminating the fixing step of the development process.

It is apparent that emulsion shrinkage may prevent the use of the same laser source in the duplication process as was used to produce the original reflection hologram, as while the reflection band may be

of the order of 50A° wide (27), shifts of 1000A° may occur. If we are forced to use a laser source of a different wavelength in the duplication process (assuming that one exists with the approximate λ), then it is clear that the fields which expose the duplicate film plate will not be the same as those that exposed the original master hologram plate. Thus the duplicate hologram will differ from that of the original. The most striking difference will be that the duplicate hologram will have a different reflection band, being shifted again towards shorter wavelengths due to emulsion shrinkage in the processing of the duplicate film plate.

5.5.3 Use of Non-Laser Sources

It may well be that in some cases no suitable laser source will exist for the duplication process, or more likely, that none will be available. If such is the case, then a conventional source would have to be used, and one would need to consider coherence lengths and path length differences. A discussion of these factors for the case of the reflection hologram would be quite similar to that given in section 5.4 for transmission holograms, and would add little new insight into the problem, and hence will not be considered here. We should perhaps note, however, that with reflection holograms we may be dealing with very thick emulsions and hence the path length differences associated with the path lengths within the master hologram plate should be given more attention than they were in section 5.4.

If a conventional source is to be used in the duplication

process, the question may well be raised as to whether or not the wavelength selectivity of the reflection hologram itself can be used to take the place of a narrow band light source. The basic idea of such a scheme would be that since the reflection hologram reflects only in a narrow portion of the spectrum, the reflected light would be narrow band, regardless of the spectral width of the illumination wave. The reflected wave could thus interfere with that portion of the illumination wave that lies with the band of reflected wavelengths. There is one obvious objection to this scheme and that is that all of the illumination wave passes through the duplicate film plate prior to reaching the reflection hologram, and hence those wavelengths not of interest would produce an undesirable level of background exposure. It may be, however, that if this background level is not too high, bleaching of the emulsion layer as described in Chapter Four may effectively remove it.

5.5.4 Efficiency

We shall limit our discussion of efficiency of duplicate reflection holograms to a brief discussion of the implications of having the illumination beam pass through the duplicate hologram film plate first, prior to striking the master reflection hologram. We observe that if the master hologram has a very low efficiency then the power in the "reference beam" will be much greater than the power in the "signal beam" which exposes the duplicate film plate. This should result in a very inefficient duplicate hologram. The situation is quite different when we duplicate an inefficient transmission hologram, when the in-

efficiency is due to overexposure, as although the power in the reconstructed images may be low due to the attenuation by the high background grain density, the "reference beam" is likewise attenuated, and thus a favorable power ratio may still be obtained.

5.6 Discussion

In this chapter we have described the duplication process from the point of view of making a hologram of a hologram, rather than in terms of making a "contact print." Described in these terms, it becomes clear that the production of a duplicate hologram involves the recording of interference patterns, just as is the case when one produces a hologram by conventional means. In the case where the master hologram is a transmission hologram, there are essentially two such interference patterns that need be considered, namely those generated by each of the two first-order "images" and the transmitted portion of the illumination wave. On the other hand, there is only one such interference pattern that is recorded when we duplicate a reflection hologram, as a reflection hologram only yields one "image" when it is illuminated.

We have seen that the nature of the duplicate hologram depends primarily on the nature of the interference pattern (or patterns) that are produced when the master hologram is illuminated. In the case where the master hologram is a "thick" transmission hologram, the nature of the two interference patterns are highly dependent on the geometrical characteristics of the illumination wave. A careful study was made of

the effects of varying the angle of incidence of the illumination wave when the master hologram was a holographic diffraction grating. From this study we are able to conclude that varying the geometrical characteristics of the illumination wave will not affect the form of the reconstructed images produced by the duplicate grating, but will affect the overall efficiency of the duplicate hologram. Furthermore, if the emulsion layer of the duplicate hologram is itself "thick," then the way in which the brightness of the images reconstructed by the duplicate hologram vary with the geometrical characteristics of the illumination beam will be strongly dependent on the geometrical characteristics of the illumination wave used in the duplication process. We recall from Chapter Four that in general the most efficient reconstruction of the signal beam is produced when the reference beam is used to illuminate the hologram. The same principle applies in the case of the duplicate hologram, only now the reference beam referred to is the illumination wave that was used in the duplication process. The effect of using a non-laser source was also considered, and it results in a localization of the interference patterns in those regions where the path length differences are less than the coherence length of the source.

Transmission holograms involving low spatial frequencies can also be treated from the point of view developed in this chapter, although most of the interesting effects predicted by this approach become negligible in the limit of very low spatial frequencies. For

example, we observe that as the spatial frequencies diminish to low values, the sensitivity to the geometrical characteristics and coherence length decrease accordingly (the path length differences decreases as f^2 for low spatial frequencies). As the spatial frequencies decrease to the point where the variations with depth are unimportant, the duplication process is quite adequately described in terms of the transmittance approach described in Chapter Two. In such a case, the duplication process can be viewed as that of making a "contact print." It is clear, however, that this will never be the case with a reflection hologram, as in this case it is the variations with depth that produce the hologram.

We observe that the mechanical stability required in a hologram duplication apparatus such as shown in Figure 5.1 is far less than what would be required in a conventional hologram apparatus, such as shown in Figure 4.1. The basic requirement for mechanical stability in either case arises from the requirement that the interference pattern that is being recorded remain fixed with respect to the recording media during the duration of the exposure. In the case where we are duplicating a hologram, the interference pattern is fixed with respect to the master hologram plate, and hence all we need do is to be sure that the duplicate film plate remains fixed with respect to the master hologram plate. Furthermore, the allowable relative motion of the two film plates can be fairly large if the spatial frequencies in the master hologram are low. The situation is quite different when we are recording a "master" hologram. In this case, the interference pattern

and film plate are fixed with respect to a third item, the table on which the apparatus is mounted. The sensitivity to motion of the film plate is similar to that of the duplication process, being proportional to the spatial frequencies being recorded. On the other hand, changes in the path lengths involved in the interference pattern on the order of $\lambda/2$ will completely wash out the recording of the interference pattern, irrespective of the spatial frequencies involved. Such changes could be produced by motions of the reflecting elements in the optical system occurring after the beam is divided into two portions.

In addition to the greatly reduced requirements for mechanical stability, a duplication apparatus can employ a source having a relatively short coherence length. Thus, although great effort may be required to produce a master hologram having a very large depth of field, such a hologram can be duplicated with no more effort than is required to duplicate a hologram having a very limited depth of field.

CHAPTER SIX

SUMMARY AND CONCLUSIONS

A general analytical method has been formulated for computing the diffracted field that is produced when a volume hologram is illuminated. The diffracted field is computed in terms of the initial exposing field, the characteristics of the recording media (assumed to be film), and the illumination field. The analysis allows for a careful accounting of the response of the recording media, and is applicable to both transmission and reflection holograms.

In the formulation of the analysis, it is assumed that the exposing and illumination fields are known and can be specified in the region of space occupied by the hologram in the form of a sum of plane or quasi-plane waves. The diffracted field is computed in the immediate vicinity of the hologram plate, and is also expressed in the form of a sum of plane or quasi-plane waves. By expressing the fields in this form, and neglecting multiple scattering, we are able to compute each of the diffracted waves independently of the others by solving a variation of the same basic problem, that of computing the amplitudes, directions, and phases of the two-first order waves that are produced when a three-dimensional sinusoidal array of scattering particles is illuminated by a plane wave.

The directions and phases of the diffracted waves produced by transmission holograms were found to be independent of the three-dimensional nature of the recording media, and are a function

only of the direction, wavelength, and phase of the illumination wave and the periodicities of the recorded interference patterns in the plane of the emulsion surface. General expressions were derived (equations 1.26, 1.28, and 1.34) for the directions and phases of these diffracted waves, which are equivalent to the equations used by Offner (14), and these expressions are shown (in Chapter Two) to be equivalent to the expressions used by Gabor (5) in his formulation of the theory of holography.

The equivalence of the two approaches with respect to the computation of the directions and phases of the diffracted waves stems from the fact that these quantities are independent of the three-dimensional nature of the recording media, and thus the characterization of the emulsion layer by the amplitude transmittance as done by Gabor, which implicitly neglects variations with depth, still yields correct results for the directions and phases of the diffracted waves - even when the concept of amplitude transmittance becomes questionable, as with thick transmission holograms involving high spatial frequencies.

The transmittance approach, however, is not applicable to reflection holograms or to the computation of the amplitudes of the diffracted waves. Reflection holograms are treated using the analysis formulated here, and it is shown that a reconstruction of the signal beam is obtained when the illumination wave is the reference beam, but that no real image beam accompanies the reconstruction of the signal beam, or "virtual image." Reflection holograms are then briefly

discussed and compared with transmission holograms.

In computing the amplitudes of the two first-order waves diffracted by the sinusoidal array of scattering particles, the individual scattered waves are summed coherently, neglecting multiple scattering. Reflection losses as well as attenuation within the emulsion layer are taken into account, and the illumination wave is allowed to have any direction or wavelength. The resulting expression for the diffracted power (equation 4.57) is shown to reduce to the results of Leith et al (8), who neglect attenuation and reflection losses and consider the case where the wave vector of the illumination wave has no component along the direction of the grating lines.

Supporting and extending this analytical work was an experimental study of the holographic diffraction grating. The ratio of the power diffracted into each of the first-order waves to the illumination power was measured as a function of the direction of the illumination wave for different gratings, using different polarizations and wavelengths. Comparison of computer generated curves with measured data showed that the theoretical and experimental results were generally in good agreement. It was seen that for thick transmission holograms, the power diffracted into the virtual or real images is highly dependent on the direction of the illumination wave, and that the power diffracted into the virtual image is a maximum when the illumination wave is the reference beam.

The effect of having a reference beam which consists of a sum of plane or quasi-plane waves was investigated (Chapter Three),

and it was shown that in order to obtain a reconstruction of the signal beam, the hologram must be illuminated with almost the identical reference beam that was used to expose it. In practice this usually requires that the experimental apparatus that was used to expose the hologram be left undisturbed, and that the developed hologram be exactly repositioned in the experimental setup. The power diffracted into the virtual image was computed as a function of error in repositioning the hologram plate for a specific experiment and then measured, and the experimental and theoretical results were found to be in agreement.

The fact that the reference beam consists of a series of waves rather than a single wave was seen to imply that the reconstruction of the signal beam is accompanied by a "background noise." A signal to noise ratio was defined and computed, and found to approach unity as the number of waves in the reference beam becomes large.

Complex spatial filtering and character recognition operations were interpreted in terms of multiple wavefront reference beam Fourier transform holography, and the effects of translations of the transparency and hologram were investigated, both with plane wave and diffuse illumination.

In Chapter Five the general analytical method for computing the diffracted field was applied to the problem of the duplication of holograms. It was shown that the duplication process should be viewed as that of recording a hologram of a hologram,

rather than that of making a contact print. Experimental evidence was presented to support this point of view, and the effect of varying the characteristics of the illumination wave was described. In addition, a simple method for duplicating reflection holograms was proposed and discussed.

APPENDIX I

VECTOR INTERFERENCE OF TWO PLANE WAVES - PREFERRED POLARIZATION FOR THE CASE WHERE THE RECORDING MEDIUM IS FILM

In this appendix we shall examine the interference pattern generated by two plane waves of the same frequency. We shall compute the time average of the Poynting vector, energy in the electric field, energy in the magnetic field and total energy. We shall then point out that it makes a difference which quantity is used to characterize the interference pattern and discuss the reasons why, in the case where the recording medium is film, that the time average energy in the electric field must be used to characterize the interference pattern. We will then show that there is a preferred polarization with respect to the recording of the interference pattern.

Thus, let us consider the case where two plane waves exist in the same region of space, where their propagation vectors are given by

$$\bar{k}_1 = \frac{2\pi}{\lambda} (\bar{e}_x \sin\theta + \bar{e}_z \cos\theta) \qquad \text{(I-1)}$$

and

$$\bar{k}_2 = \frac{2\pi}{\lambda} (-\bar{e}_x \sin\theta + \bar{e}_z \cos\theta) \quad . \qquad \text{(I-2)}$$

The two waves will be assumed to have arbitrary eliptical polarizations and thus their electric field vectors can be written in the form (using complex notation and suppressing the factor $e^{-i\omega t}$)

$$\bar{E}_1 = [E_{y1} e^{i\delta_{y1}} \bar{e}_y + E_{10} e^{i\delta_{10}} (\bar{e}_x \cos\theta - \bar{e}_z \sin\theta)] e^{i\bar{k}_1 \cdot \bar{r}} \tag{I-3}$$

and

$$\bar{E}_2 = [E_{yz} e^{i\delta_{yz}} \bar{e}_y + E_{20} e^{i\delta_{20}} (\bar{e}_x \cos\theta + \bar{e}_z \sin\theta)] e^{i\bar{k}_2 \cdot \bar{r}} \tag{I-4}$$

where E_{y1}, E_{y2}, E_{10}, E_{20}, δ_{y1}, δ_{y2}, δ_{10} and δ_{20} are real constants. The corresponding magnetic field vectors are found from a straightforward application of Maxwell's equations, using

$$\bar{H} = \frac{1}{i\omega\mu} \nabla \times \bar{E} \tag{I-5}$$

We find

$$\bar{H}_1 = [E_{y1}\sqrt{\frac{\varepsilon}{\mu}} e^{i\delta_{y1}} (-\bar{e}_x \cos\theta + \bar{e}_z \sin\theta) + E_{10}\sqrt{\frac{\varepsilon}{\mu}} e^{i\delta_{10}} \bar{e}_y] e^{i\bar{k}_1 \cdot \bar{r}} \tag{I-6}$$

and

$$\bar{H}_2 = [E_{y2}\sqrt{\frac{\varepsilon}{\mu}} e^{i\delta_{y2}} (-\bar{e}_x \cos\theta - \bar{e}_z \sin\theta) + E_{20}\sqrt{\frac{\varepsilon}{\mu}} e^{i\delta_{20}} \bar{e}_y] e^{i\bar{k}_2 \cdot \bar{r}} \tag{I-7}$$

The total electric field $\bar{E} = \bar{E}_1 + \bar{E}_2$ can be decomposed into two orthogonal fields $\bar{E}_a$ and $\bar{E}_b$, with corresponding magnetic fields $\bar{H}_a$ and $\bar{H}_b$, which are given by

$$\bar{E}_a = \bar{e}_y [E_{y1} e^{i(\bar{k}_1 \cdot \bar{r} + \delta_{y1})} + E_{y2} e^{i(\bar{k}_2 \cdot \bar{r} + \delta_{y2})}] \tag{I-8}$$

$$\bar{H}_a = -\bar{e}_x \sqrt{\frac{\varepsilon}{\mu}} \cos\theta \left[E_{y1} e^{i(\bar{k}_1\cdot\bar{r}+\delta_{y1})} + E_{y2} e^{i(\bar{k}_2\cdot\bar{r}+\delta_{y2})}\right]$$

$$+ \bar{e}_z \sqrt{\frac{\varepsilon}{\mu}} \sin\theta \left[E_{y1} e^{i(\bar{k}_1\cdot\bar{r}+\delta_{y1})} - E_{y2} e^{i(\bar{k}_2\cdot\bar{r}+\delta_{y2})}\right] \qquad (I\text{-}9)$$

and

$$\bar{E}_b = \bar{e}_x \cos\theta \left[E_{10} e^{i(\bar{k}_1\cdot\bar{r}+\delta_{10})} + E_{20} e^{i(\bar{k}_2\cdot\bar{r}+\delta_{20})}\right]$$

$$- \bar{e}_z \sin\theta \left[E_{10} e^{i(\bar{k}_1\cdot\bar{r}+\delta_{10})} - E_{20} e^{i(\bar{k}_2\cdot\bar{r}+\delta_{20})}\right] \qquad (I\text{-}10)$$

$$\bar{H}_b = \bar{e}_y \sqrt{\frac{\varepsilon}{\mu}} \left[E_{10} e^{i(\bar{k}_1\cdot\bar{r}+\delta_{10})} + E_{20} e^{i(\bar{k}_2\cdot\bar{r}+\delta_{20})}\right] \qquad (I\text{-}11)$$

We observe $\bar{E}_a$ is perpendicular to the plane of incidence of the two waves (the plane y = constant) and $\bar{E}_b$ is in the plane of incidence.

We are interested in computing the time average of the Poynting vector, electric energy density, magnetic energy density and total energy density. Papas (56) shows that these quantities are given by

$$\bar{S} = \frac{1}{2} \mathrm{Re}\,\{\bar{E} \times \bar{H}^*\} \qquad (I\text{-}12)$$

$$W_e = \frac{1}{4} \varepsilon\, \bar{E}\cdot\bar{E}^* \qquad (I\text{-}13)$$

$$W_m = \frac{1}{4} \mu\, \bar{H}\cdot\bar{H}^* \qquad (I\text{-}14)$$

and

$$W = W_e + W_m \tag{I-15}$$

respectively. It is clear from an examination of the above four equations that due to the orthogonality of the fields $\bar{E}_a$ and $\bar{E}_b$, the above quantities can be computed for each of these two fields separately. We find

$$\bar{S}_a = \bar{e}_z \sqrt{\frac{\varepsilon}{\mu}} \frac{\cos\theta}{2} [E_{y1}^2 + E_{y2}^2 + 2E_{y1} E_{y2} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{y1} - \delta_{y2})]$$

$$+ \bar{e}_x \sqrt{\frac{\varepsilon}{\mu}} \frac{\sin\theta}{2} [E_{y1}^2 - E_{y2}^2 - 2E_{y1} E_{y2} \sin(\frac{2\pi x}{\lambda} \sin\theta + \delta_{y1} - \delta_{y2})] \tag{I-16}$$

$$W_{ea} = \frac{1}{4} \varepsilon [E_{y1}^2 + E_{y2}^2 + 2E_{y1} E_{y2} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{y1} - \delta_{y2})] \tag{I-17}$$

$$W_{ma} = \frac{1}{4} \varepsilon [E_{y1}^2 + E_{y2}^2 + 2\cos(2\theta) E_{y1} E_{y2} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{y1} - \delta_{y2})] \tag{I-18}$$

and

$$W_a = \frac{1}{2} \varepsilon [E_{y1}^2 + E_{y2}^2 + 2\cos^2(\theta) E_{y1} E_{y2} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{y1} - \delta_{y2})] \tag{I-19}$$

Similarly, we find

$$\bar{S}_b = \bar{e}_z \sqrt{\frac{\varepsilon}{\mu}} \frac{\cos\theta}{2} [E_{10}^2 + E_{20}^2 + 2E_{10} E_{20} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{10} - \delta_{20})]$$

$$+ \bar{c}_x \sqrt{\frac{\varepsilon}{\mu}} \frac{\sin\theta}{2} [E_{10}^2 - E_{20}^2 + 2E_{10} E_{20} \sin(\frac{2\pi x}{\lambda} \sin\theta + \delta_{10} - \delta_{20})] \quad \text{(I-20)}$$

$$W_{eb} = \frac{1}{4} \varepsilon [E_{10}^2 + E_{20}^2 + 2\cos(2\theta) E_{10} E_{20} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{10} - \delta_{20})] \quad \text{(I-21)}$$

$$W_{mb} = \frac{1}{4} \varepsilon [E_{10}^2 + E_{20}^2 + 2E_{10} E_{20} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{10} - \delta_{20})] \quad \text{(I-22)}$$

and

$$W_b = \frac{1}{2} \varepsilon [E_{10}^2 + E_{20}^2 + 2\cos^2(\theta) E_{10} E_{20} \cos(\frac{2\pi x}{\lambda} \sin\theta + \delta_{10} - \delta_{20})] \quad \text{(I-23)}$$

It is clear from equations I-16 to I-23 that it makes a considerable difference as to which quantity is taken to characterize the interference pattern. This, of course, will depend on which of the quantities is important in the recording or measuring process. In holography, when the recording process involves film as the recording medium, it is the time-average energy in the electric field that is important. To understand why this is so, it is necessary to consider

some aspects of the formation of an image in the photographic process. When the emulsion layer of the film is exposed to light of a sufficiently short wavelength, certain changes take place which result in the formation of what is termed the "latent image." When the film bearing this latent image is chemically processed during the development procedure, an image composed of metallic silver grains is formed, corresponding to that which initially existed in the latent image.

What is of interest here is the interaction of the electromagnetic field with the film emulsion in the formation of the latent image. The nature of the process as it is presently understood is discussed in detail in treatises on photographic chemistry and photography (57, 58). The essential point with regard to this discussion is that the process involves the interaction of the electromagnetic field with a bromide ion (BR^-) in a silver bromide crystal within the emulsion layer, with the extra electron being raised to a higher energy state. Thus, the interaction of interest is essentially that of the interaction of an electromagnetic field with a nearly free electron. This problem has been treated in detail, both from a classical point of view (59) and from a quantum mechanical point of view (60). One finds that it is the electric field that is important in the interaction, and not the magnetic field.

Thus, since the recording of the interference pattern depends primarily on W_e, we should expect different results for the two different polarizations, $\bar{E}_a$ and $\bar{E}_b$. Comparing equations I-17 and I-21, we observe that for small θ, $\cos 2\theta \approx 1$ and both cases are the same.

However, as θ increases, the visibility of the fringes will decrease for the case where the electric field vector is in the plane of incidence, going to zero at $2\theta = 90^{\circ}$. On the other hand, the visibility of the fringes will be independent of θ when the electric field vector is perpendicular to the plane of incidence (Case a). These observations were experimentally confirmed by Wiener (61) and are discussed in detail by Born and Wolf (62).

It is clear then, in holographic experiments where the reference beam is brought in at a different angle from the signal beam, that it is best to have the polarization of the signal beam and reference beam be perpendicular to the plane defined by the wave vectors of the signal and reference beams.

APPENDIX II

GRAIN DENSITY EQUATION

In this appendix an expression for the grain density is derived in terms of the initial exposing field and film constants C_o, C_1, --- . The field within the emulsion during exposure of the film plate is given by (equation 1.6)

$$\bar{E} = \bar{E}_o\, e^{i\bar{k}_o\cdot\bar{r}}\, e^{-i\omega t} + \sum_{n=1}^{M} E_n\, e^{i(\bar{k}_n\cdot\bar{r} + \phi_n)}\, e^{-i\omega t} \qquad \text{(II-1)}$$

The grain density D is expressed in terms of powers of $|E|^2$ by (equation 1.7)

$$D = C_o + C_1|E|^2 + C_2|E|^4 + \text{---} \qquad \text{(II-2)}$$

where $|E|^2$ is given by

$$|E|^2 = \bar{E}\cdot\bar{E}^* \qquad \text{(II-3)}$$

Substitution of $\bar{E}$ from equation II-1 into equation II-3 yields

$$\bar{E}\cdot\bar{E}^* = E_o^2 + e^{i\bar{k}_o\cdot\bar{r}} \sum_n \bar{E}_o\cdot\bar{E}_n\, e^{-i(\bar{k}_n\cdot\bar{r}+\phi_n)} + e^{-i\bar{k}_o\cdot\bar{r}} \sum_n \bar{E}_o\cdot\bar{E}_n\, e^{i(\bar{k}_n\cdot\bar{r}+\phi_n)}$$

$$+ \sum_{n,m} \bar{E}_n\cdot\bar{E}_m\, e^{i(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + i(\phi_n - \phi_m)} \qquad \text{(II-4)}$$

It is convenient to define the real numbers b_n and C_{nm} as follows

$$b_n = \frac{1}{|\bar{E}_o|}\,\bar{E}_o\cdot\bar{E}_n \qquad \text{(II-5)}$$

$$C_{nm} = \bar{E}_n\cdot\bar{E}_m \quad . \qquad \text{(II-6)}$$

Then $\bar{E}\cdot\bar{E}^*$ can be written in the form

$$\bar{E}\cdot\bar{E}^* = E_o^2 + E_o\Sigma\, b_n\, e^{i(\bar{k}_o-\bar{k}_n)\cdot\bar{r}-i\phi_n} + E_o\Sigma_n\, b_n\, e^{-i(\bar{k}_o-\bar{k}_n)\cdot\bar{r}+i\phi_n}$$

$$+ \sum_{n,m} C_{nm}\, e^{i(\bar{k}_n-\bar{k}_m)\cdot\bar{r}+i(\phi_n-\phi_m)} \quad . \qquad \text{(II-7)}$$

It is readily seen that the second and third terms in equation II-7 can be combined as follows:

$$E_o\Sigma_n\, b_n\left(e^{i(\bar{k}_o-\bar{k}_n)\cdot\bar{r}-i\phi_n} + e^{i(\bar{k}_o-\bar{k}_n)\cdot\bar{r}+i\phi_n}\right) = 2E_o\Sigma_n\, b_n\cos[(\bar{k}_o-\bar{k}_n)\cdot\bar{r}-\phi_n] \quad . \qquad \text{(II-8)}$$

Noting that $C_{nm} = C_{mn}$ it can be shown that the last term in equation II-7 can be written in the form

$$\sum_{n,m} C_{nm}\, e^{i(\bar{k}_n-\bar{k}_m)\cdot\bar{r}+i(\phi_n-\phi_m)} = \sum_{n,m} C_{nm}\cos[(\bar{k}_n-\bar{k}_m)\cdot\bar{r}+\phi_n-\phi_m] \quad . \qquad \text{(II-9)}$$

Upon substitution of equations II-8 and II-9 into II-7 and regrouping, one obtains

$$\bar{E}\cdot\bar{E}^* = E_o^2 + 2E_o \sum_n b_n \cos[(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - \phi_n] + \sum_{n,m} C_{nm} \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m] \; . \tag{II-10}$$

The computation of $E^4 = (|E|^2)^2$ is somewhat lengthy but will nevertheless be carried out in detail here. It is more profitable to use the expression for $|E|^2$ given in equation II-7 rather than the one given in equation II-10 , as it makes the eventual grouping of the terms easier. The expression for $|E|^2$ given by equation II-7 is of the form of a sum of four terms. Recalling

$$(a+b+c+d)^2 = a^2+b^2+c^2+d^2 + 2ab + 2ac + 2ad + 2bc + 2bd + 2cd$$

it is seen that E^4 is of the form (written in the order

2ab, 2ac, 2ad, a^2, b^2, c^2, d^2, 2bc, 2bc, 2cd)

$$E^4 = 2E_o^2E^2 - E_o^4 + E_o^2 \sum_{n,m} b_n b_m e^{i(2\bar{k}_o - \bar{k}_n - \bar{k}_m)\cdot\bar{r}} e^{-i(\phi_n + \phi_m)}$$

$$+ E_o^2 \sum_{n,m} b_n b_m e^{-i(2\bar{k}_o - \bar{k}_n - \bar{k}_m)\cdot\bar{r}} e^{i(\phi_n + \phi_m)}$$

$$+ \sum_{n,m,p,q} C_{nm} C_{pq} e^{i(\bar{k}_n - \bar{k}_m + \bar{k}_p - \bar{k}_q)\cdot\bar{r}} e^{i(\phi_n - \phi_m + \phi_p - \phi_q)}$$

$$+ 2E_o^2 \sum_{n,m} b_n b_m e^{-i(\bar{k}_n - \bar{k}_m)\cdot\bar{r}} e^{i(\phi_n - \phi_m)}$$

$$+ 2E_o \sum_{n,p,q} b_n C_{pq} e^{i(\bar{k}_o - \bar{k}_n + \bar{k}_p - \bar{k}_q)\cdot\bar{r}} e^{-i(\phi_n - \phi_p + \phi_q)}$$

$$+ 2E_o \sum_{n,p,q} b_n C_{qp} e^{-i(\bar{k}_o - \bar{k}_n - \bar{k}_q + \bar{k}_p)\cdot\bar{r}} e^{i(\phi_n + \phi_q - \phi_p)} \quad . \qquad \text{(II-11)}$$

This can be written in the form (recalling $C_{pq} = C_{qp}$)

$$E^4 = 2E_o^2E^2 - E_o^4 + 2E_o^2 \sum_{n,m} b_n b_m \cos[(2\bar{k}_o - \bar{k}_n - \bar{k}_m)\cdot\bar{r} - \phi_n - \phi_m]$$

$$+ 2E_o^2 \sum_{n,m} b_n b_m \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m]$$

$$+ 4E_o \sum_{n,p,q} b_n C_{pq} \cos[(\bar{k}_o - \bar{k}_n + \bar{k}_p - \bar{k}_q)\cdot\bar{r} - \phi_n + \phi_p - \phi_q]$$

$$+ \sum_{n,m,p,q} C_{nm} C_{pq} e^{i(\bar{k}_n - \bar{k}_m + \bar{k}_p - \bar{k}_q)\cdot\bar{r}} e^{i(\phi_n - \phi_m + \phi_p - \phi_q)} \quad . \qquad \text{(II-12)}$$

The expressions for $|E|^2$ and $|E|^4$ given in equations II-10 and II-12 are substituted in equation II-2 to yield

$$D = C_o + C_1E_o^2 + C_2E_o^4 + (2C_1E_o + 4C_2E_o^3) \sum_n b_n \cos[(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - \phi_n]$$

$$+ (C_1 + 2C_2E_o^2) \sum_{n,m} C_{nm} \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m]$$

$$+ 2C_2E_o^2 \sum_{n,m} b_n b_m \cos[(\bar{k}_n - \bar{k}_m)\cdot\bar{r} + \phi_n - \phi_m]$$

$$+ 2C_2E_o^2 \sum_n b_n^2 \cos[2(\bar{k}_o - \bar{k}_n)\cdot\bar{r} - 2\phi_n]$$

$$+ 2C_2E_o^2 \sum_{\substack{n,m \\ n\neq m}} b_n b_m \cos[(2\bar{k}_o - \bar{k}_n - \bar{k}_m)\cdot\bar{r} - \phi_n - \phi_m]$$

$$+ 4C_2E_o \sum_{n,p,q} b_n C_{pq} \cos[(\bar{k}_o - \bar{k}_n + \bar{k}_p - \bar{k}_q)\cdot\bar{r} - \phi_n + \phi_p - \phi_q]$$

$$+ C_2 \sum_{n,m,p,q} C_{nm} C_{pq} e^{i(\bar{k}_n - \bar{k}_m + \bar{k}_p - \bar{k}_q)\cdot\bar{r}} e^{i(\phi_n - \phi_m + \phi_p - \phi_q)}$$

$$+ \text{higher order terms} . \qquad \text{(II-13)}$$

APPENDIX III

TRANSFORMATION EQUATIONS

In this appendix we shall list the transformation equations for the coordinates and direction cosines that exist between the x'y'z', xyz, and x"y"z" coordinate systems, as referred to in Chapter One.

The xyz system is formed by a rotation of Φ (equation 1.16) about the z' axis. The corresponding transformation equation are

$$x' = x \cos\Phi - y \sin\Phi \qquad \text{(III-1)}$$

$$y' = x \sin\Phi + y \cos\Phi \qquad \text{(III-2)}$$

$$z' = z \qquad \text{(III-3)}$$

and

$$\ell' = \ell \cos\Phi - m \sin\Phi \qquad \text{(III-4)}$$

$$m' = \ell \sin\Phi + m \cos\Phi \qquad \text{(III-5)}$$

$$n' = n \qquad \text{(III-6)}$$

the x"y"z" system is formed by a rotation of β (equation 1.55) about the x axis. The corresponding transformation equations are

$$x = x" \qquad \text{(III-7)}$$

$$y = y" \cos\beta - z" \sin\beta \qquad \text{(III-8)}$$

$$z = y" \sin\beta + z" \cos\beta \qquad \text{(III-9)}$$

and

$$\ell = \ell" \qquad \text{(III-10)}$$

$$m = m" \cos\beta - n" \sin\beta \qquad \text{(III-11)}$$

$$n = m" \sin\beta + n" \cos\beta \qquad \text{(III-12)}$$

APPENDIX IV

TRANSLATION SENSITIVITY CALCULATIONS

In this appendix we shall compute the effect of a translation of the film plate on the phase of one of the diffracted waves which contributes to the reconstruction of the m^{th} signal wave in equation 3.11. There are P such waves which contribute to the m^{th} signal wave, and we shall consider the j^{th} one of these. This wave is produced by the j^{th} reference wave interacting with the periodicity corresponding to

$$C_{jm} \cos[(\bar{k}_j - \bar{k}_m) \cdot \bar{r} + \phi_j - \phi_m]$$

The loci of points of maximum grain density in this periodicity is given by

$$(\bar{k}_j - \bar{k}_m) \cdot \bar{r} + \phi_j - \phi_m = 2M\pi \quad . \tag{IV-1}$$

Defining, as before, the x y z coordinate system to be formed by a rotation of Φ about the z' axis, where

$$\Phi = \tan^{-1}\left(\frac{\ell'_m - \ell'_j}{m'_j - m'_m} \right) \tag{IV-2}$$

(ℓ',m',n' being direction cosines in the x' y' z' system), equation III-1 can be written in the form (in the $z = 0$ plane)

$$(m_j - m_m)y + \phi_j - \phi_m = 2\pi M \quad . \tag{IV-3}$$

A displacement of the film plate an amount Δy in the y direction means that we replace y by $y - \Delta y$ in equation IV-3. The loci of points of maximum grain density are then specified by

$$(m_j - m_m)(y - \Delta y) + \phi_j - \phi_m = 2\pi M \quad . \tag{IV-4}$$

We specify the phase of the wave diffracted by the periodicity under consideration by requiring that the interference pattern generated by the transmitted portion of the illumination wave and the diffracted wave "match up" with the recorded interference pattern as specified by equation IV-4. Since we are re-illuminating with one of the two original waves which produced the periodicity under consideration, one of the first-order waves will have the direction of the other initial wave. The re-illumination interference pattern which is of interest is thus specified by (in the $z = 0$ plane)

$$(m_j - m_m)y + \phi_j - \phi_d = 2\pi M \quad . \tag{IV-5}$$

We specify ϕ_d by requiring that the loci of points of minimum electric field in the above re-illumination pattern coincide with the locii of points of maximum grain density. Thus we set $M = 0$ in equation IV-4, solve for y, and substitute this value of y in equation IV-5. We then solve equation IV-5 for ϕ_d with $M = 1/2$. Thus

$$y = \frac{\phi_m - \phi_j}{(m_j - m_m)} + \Delta y \tag{IV-6}$$

and

$$(m_j - m_m)\left[\frac{\phi_m - \phi_j}{(m_j - m_m)} + \Delta y\right] + \phi_j - \phi_d = \pi \tag{IV-7}$$

or

$$\phi_d = \phi_m - \pi + \Delta y(m_j - m_m) \quad . \tag{IV-8}$$

This is equivalent to the result specified by equation 3.15, which is

slightly more general since it allows for a translation of the film plate in the z direction, while IV-8 only allows for a translation in the x'y' plane. To show this equivalence we express equation 3.15 in the x y z coordinate system defined by equation IV-2. Recalling that $\ell_j = \ell_m$ in this system, and taking $\bar{r}_o = \Delta x\, \bar{e}_x + \Delta y\, \bar{e}_y$, equation 3.14 becomes

$$\phi_{jm} = (m_j - m_m)\, \Delta y \qquad \text{(IV-9)}$$

which corresponds to having

$$\phi_d = \phi_m - \pi + \Delta y(m_j - m_m) \qquad \text{(IV-10)}$$

APPENDIX V

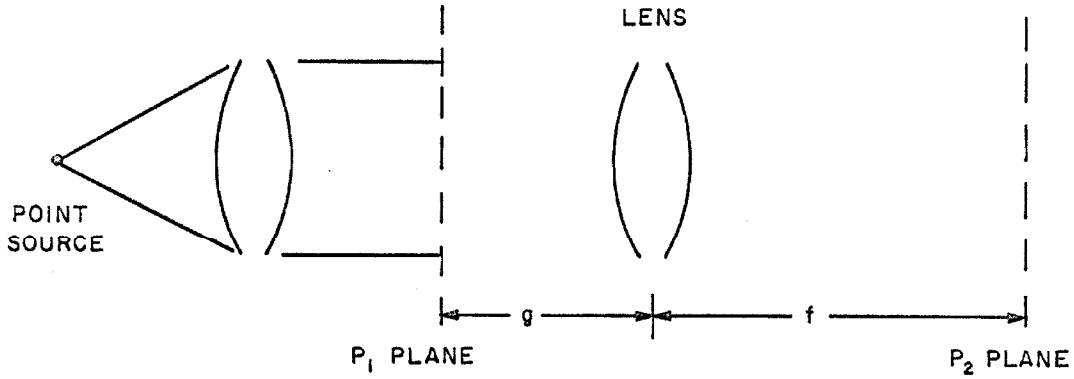
TRANSFORM RELATIONS IN COHERENT OPTICAL SYSTEMS

In this appendix we shall review the subject of transform relations in coherent optical systems. We shall derive the relationship between the field amplitudes in the front and back focal plane of an ideal lens. The derivation which we shall give will be essentially that given by Cutrona et al (29). Other derivations can be found in articles by Champagne (30) and by Vander Lugt (44).

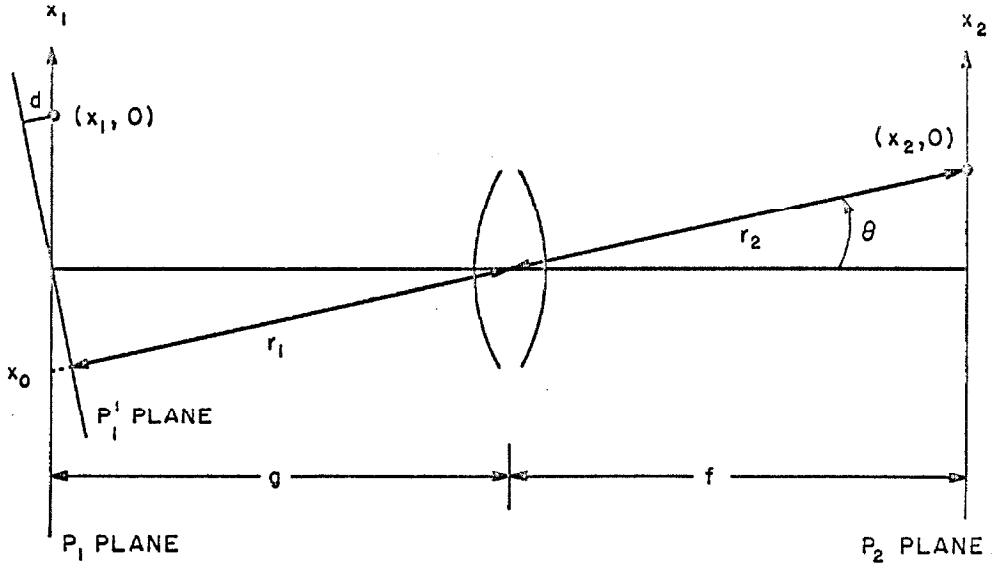
Let us consider the system shown in Figure V-1a. Some system, such as a point source in the focal plane of a lens, produces a collimated beam which has some complex amplitude distribution $E_1(x_1,y_1)$ in plane P_1. We wish to compute the resulting distribution $E_2(x_2,y_2)$ in P_2, where the plane P_2 is taken to be the back focal plane of lens 1.

The method that will be used will be to apply Fresnel-Kirchhoff diffraction theory, treating plane P_1 as a large diffracting aperture with complex transmittance $E_1(x_1,y_1)$, illuminated by a plane wave of unit amplitude at normal incidence. Thus, applying the results of Fresnel-Kirchhoff diffraction theory (64), we write

$$E_2(x_2,y_2) = -\frac{i}{2\lambda}\iint E_1(x_1,y_1)\,\frac{e^{ikr}}{S}\,[1+\cos\psi]\,dx_1dy_1 \qquad (V\text{-}1)$$



(a) OPTICAL SYSTEM



(b) GEOMETRY

FIGURE V-1 COHERENT OPTICAL SYSTEM

where the effect of the lens has been taken into account by replacing the distance S in the exponent by r, the optical path length between (x_1,y_1) and (s_2,y_2).

The quantity $\frac{1+\cos\psi}{S}$ will be removed from under the integral and set equal to $2/f$. Note that we set $1/S = 1/f$ not $1/f+g$, as we neglect the amplitude attenuation due to the distance g between plane P_1 and the lens, since we have a collimated beam to the left of the lens. Thus

$$E_2(x_2,y_2) = -\frac{i}{\lambda f}\iint E_1(x_1,y_1)\, e^{ikr}\, dx_1 dy_1 \quad . \tag{V-2}$$

The next problem is to compute r, the optical path length. This will be done for the special two-dimensional case that results if we set $y_1 = y_2 = 0$. Consider the diagram shown in Figure V-1b. Since plane P_2 is the back focal plane of the lens, a plane wave making an angle Θ with the normal to the P_1 (with wave vector $\bar{k}$ in the xz plane) is brought to focus at the point $(x_2,0)$ in plane P_2. Any point on the P_1' plane which is taken to be a plane perpendicular to $\bar{k}$ is the same optical distance $c = r_1 + r_2$ from x_2. The optical path length between $(x_1,0)$ and $(x_2,0)$ is thus seen to be

$$r(x_1,0,x_2,0) = r_1 + r_2 - d \quad . \tag{V-3}$$

Now

$$r_1 + r_2 = \sqrt{g^2 - x_o^2\cos^2\Theta} + \sqrt{x_2^2 + f^2} \tag{V-4}$$

or

$$r_1 + r_2 = g\sqrt{1 - (\frac{x_o}{g})^2 \cos^2 \theta} + f\sqrt{1 + (\frac{x_2}{f})^2} \quad . \qquad (V\text{-}5)$$

By similar triangles, it is seen that

$$\frac{x_o}{g} = \frac{x_2}{f} \quad . \qquad (V\text{-}6)$$

It will be assumed that θ is small, which implies that we can take $\cos \theta \simeq 1$ and $\frac{x_2}{f} << 1$. Hence

$$r_1 + r_2 \simeq g\sqrt{1 - (\frac{x_2}{f})^2} + f\sqrt{1 + (\frac{x_2}{f})^2} \quad . \qquad (V\text{-}7)$$

Expanding the square root using $(1 \pm \alpha)^n \simeq 1 \pm n\alpha$ yields

$$r_1 + r_2 \simeq g[1 - \frac{1}{2} (\frac{x_2}{f})^2] + f[1 + \frac{1}{2} (\frac{x_2}{f})^2] \qquad (V\text{-}8)$$

or

$$r_1 + r_2 \simeq g + f + (1 - \frac{g}{f}) \frac{x_2^2}{2f} \quad . \qquad (V\text{-}9)$$

The distance d is seen to be (from Figure V-1b)

$$d = x_1 \sin \theta \qquad (V\text{-}10)$$

but

$$\sin \theta = \frac{x_2}{f} \qquad (V\text{-}11)$$

hence

$$d = \frac{x_1 x_2}{f} \quad . \tag{V-12}$$

Thus

$$r(x_1,0,x_2,0) = g + f + (1 - \frac{g}{f}) \frac{x_2^2}{2f} - \frac{x_1 x_2}{f} \quad . \tag{V-13}$$

A similar, but more lengthy, computation yields for the general case

$$r(x_1,y_1,x_2,y_2) = g + f + (1 - \frac{g}{f}) (\frac{x_2^2 + y_2^2}{2f}) - \frac{x_1 x_2}{f} - \frac{y_1 y_2}{f} \quad . \tag{V-14}$$

The constant term $g + f$ will be suppressed since it merely adds a constant phase factor. Thus, if we define $\beta(x_2,y_2)$ as

$$\beta(x_2,y_2) = (1 - \frac{g}{f}) (\frac{x_2^2 + y_2^2}{2f}) \tag{V-15}$$

then

$$E_2(x_2,y_2) = - \frac{i}{\lambda f} e^{ik\beta(x_2,y_2)} \iint E_1(x_1,y_1) e^{- \frac{i2\pi}{\lambda f} (x_1 x_2 + y_1 y_2)} dx_1 dy_1 \, . \tag{V-16}$$

Defining the "spatial frequencies" ξ and η as

$$\xi = \frac{2\pi}{\lambda f} x_2 \tag{V-17}$$

and

$$\eta = \frac{2\pi}{\lambda f} y_2 \tag{V-18}$$

equation 2.85 becomes

$$E_2 = - \frac{i}{\lambda f} e^{ik\beta} \iint E_1(x_1,y_1) e^{-i(\xi x_1 + \eta y_1)} dx_1 dy_1 \, . \tag{V-19}$$

The above formulas were derived assuming a time variation of the form $e^{-i\omega t}$. If a time variation of the form $e^{i\omega t}$ is used, then the result is a change in sign of k and hence ξ and η are defined as

$$\xi = -\frac{2\pi}{\lambda f} x_2 \tag{V-20}$$

$$\eta = -\frac{2\pi}{\lambda f} y_2 \tag{V-21}$$

and the expression for E_2 becomes

$$E_2 = \frac{i}{\lambda f} e^{-ik\beta} \iint E_1(x_1,y_1)\, e^{-i(\xi x_1 + \eta y_1)}\, dx_1 dy_1 \quad . \tag{V-22}$$

In the following we will use the $e^{-i\omega t}$ time convention so equations V-16 and V-19 will be applicable.

It is observed that when plane P_1 is the front focal plane of the lens, then $g = f$ and $\beta = 0$. Thus, for this case, (apart from a constant) E_1 and E_2 form a Fourier transform pair. Hence E_1 can be found from E_2 by an inverse Fourier transform

$$E_1(x_1,y_1) = -\left(\frac{1}{2\pi}\right)^2 \iint \left(\frac{\lambda f}{i}\right) E_2(\xi,\eta)\, e^{i(\xi x_1 + \eta y_1)}\, d\xi\, d\eta \tag{V-23}$$

or, since $\xi = \frac{2\pi}{\lambda f} x_2$ and $\eta = \frac{2\pi}{\lambda f} y_2$.

$$E_1(x_1,y_1) = \frac{i}{\lambda f} \iint E_2(x_2,y_2)\, e^{i\frac{2\pi}{\lambda f}(x_1x_2 + y_1y_2)}\, dx_2 dy_2 . \tag{V-24}$$

We can check to see if the constant $i/\lambda f$ is correct by applying Parseval's theorem and the conservation of energy. This is done as follows: Conservation of energy requires that

$$E_1^2(x_1,y_1)dx_1dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_2^2(x_2,y_2)dx_2dy_2 \qquad (V\text{-}25)$$

where $E^2 = EE^*$.

We have shown that E_2 is given in terms of E_1 by

$$E_2 = C \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_1(x_1,y_1)\, e^{-i\xi x_1}\, e^{-i\eta y_1}\, dx_1dy_1 \qquad (V\text{-}26)$$

or

$$\frac{1}{C} E_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_1\, e^{-i(\xi x_1 + \eta y_1)}\, dx_1dy_1 \equiv f\{E_1\} \ . \qquad (V\text{-}27)$$

Thus E_2/C and E_1 form a Fourier transform pair. Applying Parseval's theorem

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_1^2\, dx_1dy_1 \equiv \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{E_2^2}{C^2}\,(\xi,\eta)\, d\xi\, d\eta \qquad (V\text{-}28)$$

or

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_1^2\, dx_1dy_1 = \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{E_2^2}{C^2}\left(\frac{2\pi}{\lambda f}\right)\left(\frac{2\pi}{\lambda f}\right) dx_2dy_2 \ . \qquad (V\text{-}29)$$

We observe conservation of energy requires $\frac{1}{c^2}\left(\frac{1}{\lambda f}\right)^2 \equiv 1$, hence

$$|C| = \frac{1}{\lambda f} \quad . \tag{V-30}$$

APPENDIX VI

SINUSOIDAL GRAIN DENSITY - ABSENCE OF THE SECOND AND HIGHER ORDER WAVES

In this appendix we shall show that if the spatial variation of the grain density is sinusoidal, then only the first-order diffracted waves are produced when the grains are illuminated with a plane wave. Let us begin by summing the waves scattered by the grains located in the $z = 0$ plane (we shall use the xyz coordinate system defined in Section 1.3.3). The grain density in this plane is of the form

$$D(x,y,0) = D_o \left[1 + \cos\left(\frac{2\pi y}{d}\right)\right] \qquad \text{(VI-1)}$$

where d is the periodicity or fringe spacing in the $z = 0$ plane (the plane of the emulsion surface). The directions of the diffracted waves are specified (see Section 1.3.4) by requiring that there be no phase difference between the waves scattered by grains with the same y coordinate and that there be a linear phase shift of

$$e^{\frac{i2\pi Ny}{d}}$$

as we move in the y direction, where N is the diffracted order under consideration ($N = \pm 1, \pm 2$, etc.). The amplitude of the particular diffracted wave under consideration is proportional to

$$A = \int_{xy} D(x,y,0)\, e^{\frac{i2\pi Ny}{d}}\, dx\, dy \qquad \text{(VI-2)}$$

where the integration over x and y corresponds to integrating over the transverse extent of the hologram grating. Because of the periodic nature of $D(x,y,0)$ we need only consider the integration over $y = 0$ to $y = d$. It is straightforward to show that

$$\int_{y=0}^{d} \left[1 + \cos\left(\frac{2\pi y}{d}\right)\right] e^{\frac{i2\pi N y}{d}} \, dy$$

vanishes for all integer values of N except $N = \pm 1$, and thus we conclude that only the two-first order waves will be produced by the grains in the $z = 0$ plane. The same conclusion is reached for the grains located in any plane $z =$ constant and hence the grating will produce only the two first-order waves.

APPENDIX VII

PATH LENGTH DIFFERENCES

In this appendix we shall compute the path length differences involved in the duplication of a holographic diffraction grating. Illumination of the holographic diffraction grating as shown in Figure VII-1 with a plane wave produces two first-order diffracted plane waves and a zeroth-order plane wave. The directions θ_d of the diffracted waves are found from

$$\sin \theta_d = \sin \theta_i + \frac{\tilde{N}\tilde{\lambda}_o}{d} \qquad \text{(VII-1)}$$

We wish to consider the case where arrangement of the master and duplicate plates is as shown in Figure VII-2, where the region between the two emulsion layers is filled with a fluid having an index of refraction η. We shall compute the optical path length difference between the zeroth-order wave and either of the two first order waves at an arbitrary point P in the plane of the emulsion surface of the duplicate film plate. We shall assume that the diffraction by the holographic diffraction grating takes place at the emulsion-fluid interface (i.e., we neglect the thickness of the emulsion layer) and hence, as seen from Figure VII-2, the optical path length difference ΔL is given by

$$\Delta L = \eta_e d_1 + \eta d_2 - \eta d_3 \quad . \qquad \text{(VII-2)}$$

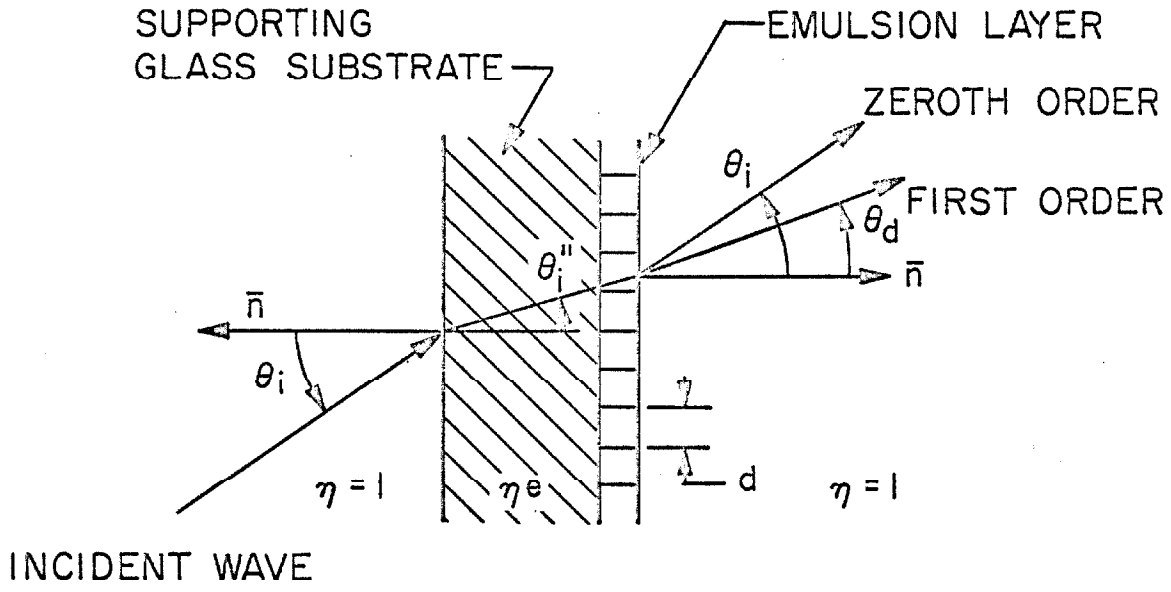


Figure VII-1 ILLUMINATION OF A HOLOGRAPHIC DIFFRACTION GRATING

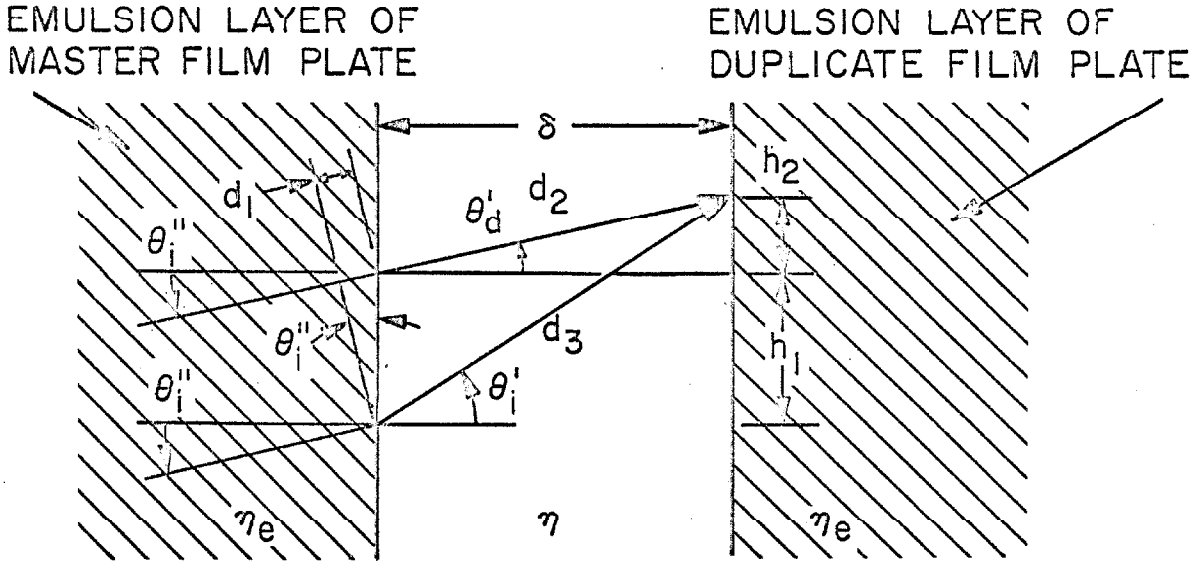


Figure VII-2 DUPLICATION GEOMETRY

We observe that

$$d_1 = h_1 \sin \theta_i'' \qquad \text{(VII-3)}$$

and that

$$\tan \theta_i = \frac{h_2 + h_1}{\delta} \quad . \qquad \text{(VII-4)}$$

Now

$$h_1 = \delta \tan \theta_i' - h_2 \qquad \text{(VII-5)}$$

and

$$h_2 = \delta \tan \theta_d' \qquad \text{(VII-6)}$$

and hence

$$d_1 = \delta(\tan \theta_i' - \tan \theta_d') \sin \theta_i'' \quad . \qquad \text{(VII-7)}$$

We see that

$$d_2 = \frac{\delta}{\cos \theta_d'} \qquad \text{(VII-8)}$$

and

$$d_3 = \frac{\delta}{\cos \theta_i'} \qquad \text{(VII-9)}$$

and thus

$$\Delta L = \delta \left[\eta_e \sin \theta_i'' (\tan \theta_i' - \tan \theta_d') + \frac{\eta}{\cos \theta_d'} - \frac{\eta}{\cos \theta_i'}\right] \qquad \text{(VII-10)}$$

Now, using Snell's Law

$$\eta_e \sin \theta_i'' = \eta \sin \theta_i' \quad , \qquad \text{(VII-11)}$$

equation VII-10 becomes

$$\Delta L = \eta\delta\left[(\tan\theta_i' - \tan\theta_d')\sin\theta_i' + \frac{1}{\cos\theta_d'} - \frac{1}{\cos\theta_i'}\right]. \quad \text{(VII-12)}$$

REFERENCES

1. D. Gabor, "A New Microscope Principle", Nature 161, 777-778 (1948).

2. E. N. Leith and J. Upatnieks, "Reconstructed Wavefronts and Communication Theory", J. Opt. Soc. Am. 52, 1123-1130 (1962).

3. E. N. Leith and J. Upatnieks, "Wavefront Reconstruction with Continuous Tone Objects", J. Opt. Soc. Am. 53, 1377-1381 (1963).

4. E. N. Leith and J. Upatnieks, "Wavefront Reconstruction with Diffused Illumination and Three-Dimsnional Objects", J. Opt. Soc. Am. 54, 1295-1301 (1964).

5. D. Gabor, "Microscopy by Reconstructed Wavefronts", Proc. Roy. Soc. A197, 454-487 (1949).

6. A. A. Friesem, "Holograms on Thick Emulsions", Appl. Phys. Lett. 7, 102-103 (1965).

7. K. S. Pennington and L. H. Lin, "Multicolor Wavefront Reconstruction", Appl. Phys. Lett. 7, 56-57 (1965).

8. E. N. Leith, A. Kozma, J. Upatnieks, J. Marks and N. Massey, "Holographic Data Storage in Three-Dimensional Media", Applied Optics 5, 1303-1311 (1966).

9. N. George and J. W. Matthews, "Holographic Diffraction Gratings", Appl. Phys. Lett. 9, 212-215 (1966).

10. Yu. N. Denisyuk, "On the Reproduction of the Optical Properties of an Object by the Wave Field of its Scattered Radiation", Opt. Spect. 15, 279-284 (1963).

11. Yu. N. Denisyuk, "On the Reproduction of the Optical Properties of an Object by the Wave Field of its Scattered Radiation. II", Opt. Spect. 18, 152-157 (1965).

12. P. J. van Heerden, "A New Optical Method of Storing and Retrieving Information", Applied Optics 2, 387-392 (1963).

13. P. J. van Heerden, "Theory of Optical Information Storage in Solids", Applied Optics 2, 393-400 (1963).

14. A. Offner, "Ray Tracing Through a Holographic System", J. Opt. Soc. Am. 56, 1509-1512 (1966).

15. P. Glafkides, Photographic Chemistry Volumes 1 and 2, Fountain Press (1958).

16. C. B. Neblette, Photography its Materials and Processes, D. Van Nostrand Co. (1962).

17. A. K. Rigler, "Wavefront Reconstruction by Reflection", J. Opt. Soc. Am. 55, 1693 (1965).

18. J. H. Altman, "Pure Relief Images on Type 649-F Plates", Applied Optics 5, 1689-1690 (1966).

19. J. C. Urbach and R. W. Meier, "Thermoplastic Xerographic Holography", Applied Optics 5, 666-667 (1966).

20. J. P. Kirk, "Hologram on Photochromic Glass", Applied Optics 5, 1684-1685 (1966).

21. G. Toraldo di Francia, NBS Circ. 526, 165 (1954).

22. G. W. Stroke and A. E. Labeyrie, "White-Light Reconstruction of Holographic Images Using the Lippmann-Bragg Diffraction Effect", Phys. Lett. 20, 368-370 (1966).

23. L. H. Lin, K. S. Pennington, G. W. Stroke and A. E. Labeyrie, "Multicolor Holographic Image Reconstruction with White-Light Illumination", Bell System Tech. J. XLV, 659-660 (1966).

24. J. Upatnieks, J. Marks and R. Fedorowicz, "Color Holograms for White-Light Reconstruction", Appl. Phys. Lett. 8, 286-287 (1966).

25. G. W. Stroke and R. G. Zech, "White-Light Reconstruction of Color Images from Black and White Volume Holograms Recorded on Sheet Film", Appl. Phys. Lett. 9, 215-217 (1966).

26. C. Kittel, Introduction to Solid State Physics, J. Wiley and Sons, Inc. (1959).

27. H. Fleisher, et al, "An Optically Accessed Memory Using the Lippmann Process for Information Storage", (Chapter 1) in Optical and Electro Optical Information Processing., Edited by J. T. Tippett, et al, MIT Press (1965).

28. M. Born and E. Wolf, Principles of Optics , The Macmillan Co. (1964).

29. L. J. Cutrona, E. N. Leith, C. J. Palermo and L. J. Porcello, "Optical Data Processing and Filtering System", IRE Trans. Inform. Theory IT-6, 386-400 (1960).

30. E. B. Champagne, "Transform Relations in Coherent Systems", Applied Optics 5, 1088 (1966).

31. J. A. Armstrong, "Fresnel Holograms: Their Imaging Properties and Aberrations", IBM J. Res. Develop. 9, 171-178 (1965).

32. E. N. Leith, J. Upatnieks and K. A. Haines, "Microscopy by Wavefront Reconstruction", J. Opt. Soc. Am. 55, 981-986 (1965).

33. R. W. Meier, "Magnifications and Third Order Aberrations in Holography", J. Opt. Soc. Am. 55, 987-992 (1965).

34. J. T. Winthrop and C. R. Worthington, "Fresnel-Transform Representation of Holograms and Hologram Classification", J. Opt. Soc. Am. 56, 1362-1368 (1966).

35. G. W. Stroke, R. Restrick, A. Funkhouser and D. Brumm, "Resolution-Retrieving Compensation of Source Effects by Correlative Reconstruction in High Resolution Holography", Phys. Lett. 18, 274-275 (1965).

36. K. S. Pennington and R. J. Collier, "Hologram-Generated Ghost Image Experiments", Appl. Phys. Lett. 8, 14-16 (1966).

37. R. J. Collier and K. S. Pennington, "Ghost Imaging by Holograms Formed in the Near Field", Appl. Phys. Lett. 8, 44-46 (1966).

38. G. W. Stroke, An Introduction to Coherent Optics and Holography , Academic Press (1966).

39. G. W. Stroke, "Lensless Fourier-Transform Method for Optical Holography", Appl. Phys. Lett. 6, 201-203 (1965).

40. G. W. Stroke, D. B. Brumm and A. Funkhouser, "Three Dimensional Holography with Lensless Fourier-Transform Holograms and Coarse P/N Polaroid Film", J. Opt. Soc. Am. 55, 1327-1328 (1965).

41. A. Vander Lugt, "Signal Detection by Complex Spatial Filtering", IEEE Trans. Inform. Theory IT-10, 139-145 (1964).

42. A. Vander Lugt, F. B. Rotz and A. Klooster, Jr., "Character-Reading by Optical Spatial Filtering", (Chapter 7) in Optical and Electro Optical Information Processing., Edited by J. T. Tippett, et al, MIT Press (1965).

43. D. Gabor, "Character Recognition by Holography", Nature 208, 422-423 (1965).

44. A. Vander Lugt, "Operational Notation for the Analysis and Synthesis of Optical Data Processing Systems", Proc. IEEE 54, 1055-1063 (1966).

45. F. S. Harris, Jr., G. C. Sherman and B. H. Billings, "Copying Holograms", Applied Optics 5, 665-666 (1965).

46. M. J. Landry, "Copying Holograms", Appl. Phys. Lett. 9, 303-304 (1966).

47. F. B. Rotz and A. A. Friesem, "Holograms with Nonpseudoscopic Real Images", Appl. Phys. Lett. 8, 146-148 (1966).

48. D. B. Brumm, "Copying Holograms", Applied Optics 5, 1946-1947 (1966).

49. R. Vandewarker and K. Snow, "Low Spatial Frequency Holograms of Solid Objects", Appl. Phys. Lett. 10, 35-36 (1967).

50. H. C. van de Hulst, Light Scattering by Small Particles, J. Wiley and Sons, Inc. (1957).

51. P. Glafkidis, op. cit., 51.

52. M. Born and E. Wolf, op. cit., Chapter 12.

53. C. B. Burckhardt, "Diffraction of a Plane Wave at a Sinusoidally Stratified Dielectric Grating", J. Opt. Soc. Am. 56, 1502-1509 (1966).

54. M. Born and E. Wolf, op. cit., 42-43.

55. J. F. Asmus, Scientific Report No. 3 under Air Force Contract AF49(638)-1322, 1964 (unpublished); Ph.D. Thesis, California Institute of Technology, 1964 (unpublished).

56. C. H. Papas, Theory of Electromagnetic Wave Propagation, McGraw-Hill (1965).

57. P. Glafkides, op. cit., Chapter 3.

58. C. B. Neblette, op. cit., Chapter 12.

59. M. Born and E. Wolf, op. cit., Chapter 1.

60. L. I. Schiff, Quantum Mechanics, McGraw-Hill (1955), Chapter 10.

61. C. Wrener, "Stehende Lichtwellen und die Schwingungsrichtung Polarisirten Lichtes", Ann. d. Physik, 40 203 (1890).

62. M. Born and E. Wolf, op. cit., 279.

63. R. K. Mueller and N. K. Sheridon, "Sound Holograms and Optical Reconstructions", Appl. Phys. Lett. 9, 328-329 (1966).

64. M. Born and E. Wolf, op. cit., Chapter 8.

65. M. Born and E. Wolf, Ibid., 316-323.

66. Private Communication with R. Anwayl of Eastman Kodak Co., (Rochester, New York).

67. C. Kittel, op. cit. 45.

68. R. P. Chambers and J. S. Courtney-Pratt, "Bibliography on Holograms", J. SMPTE 75, 373-435 (1966).