ESD-TR-89-148

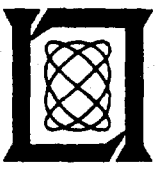
Technical Report
854

Binary Optics Technology: The Theory and Design of Multi-level Diffractive Optical Elements

G.J. Swanson

14 August 1989

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS



Prepared for the Defense Advanced Research Projects Agency under Air Force Contract F19628-85-C-0002.

This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Defense Advanced Research Projects Agency under Air Force Contract F19628-85-C-0002 (ARPA Order 6008).

The ESD Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Hugh L. Southall

Hugh L. Southall, Lt. Col., USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

BINARY OPTICS TECHNOLOGY: THE THEORY AND DESIGN OF MULTI-LEVEL DIFFRACTIVE OPTICAL ELEMENTS

G.J. SWANSON
Group 52

TECHNICAL REPORT 854

14 AUGUST 1989

LEXINGTON MASSACHUSETTS

ABSTRACT

Multi-level diffractive phase profiles have the potential to significantly improve the performance of many conventional lens systems. The theory, design, and fabrication of these diffractive profiles are described in detail. Basic examples illustrate the potential usefulness, as well as the limitations, of these elements.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

The direction of propagation of a light ray can be changed by three basic means: reflection, refraction, and diffraction. Three simple, yet useful, equations that describe these phenomena are the law of reflection, Snell's law, and the grating equation. These three fundamental equations are the foundation for the description of redirecting light rays.

Virtually all optical systems in existence rely on only reflection and refraction to achieve the desired optical transformation. Lens design, based on reflective and refractive elements, is a well-established and refined process. Until recently, diffractive elements have been neglected as viable components of optical systems.

One reason for the lack of interest in using diffractive elements in a lens design is that the process of diffraction does not simply redirect a light ray. Diffraction, unlike reflection and refraction, splits a light ray into many rays — each of which is redirected at a different angle. The percentage of the incident light redirected by the desired angle is referred to as the diffraction efficiency. The diffraction efficiency of a diffractive element is determined by the element's surface profile. If the light that is not redirected by the desired angle is substantial, the result will be an intolerable amount of scatter in the image or output plane of the optical system.

Fortunately, a surface profile exists (in theory) that achieves 100-percent diffraction efficiency at a specified wavelength. The theoretical diffraction efficiency of this surface profile is also relatively insensitive to a change in wavelength. This profile could therefore be used in optical systems operating over finite wavelength bands. Section 2 of this report discusses a theory of this highly efficient diffractive profile. The diffraction efficiency of the simplest example of a diffractive element, a grating, is derived. The section concludes with the extension of the results to diffractive elements having arbitrary phase profiles.

The theoretical existence of a surface profile having high diffraction efficiency has no practical consequences in the design of optical systems unless this profile can be easily determined and readily fabricated. The diffractive surface profile described in Section 2 is not easily fabricated. It is possible, however, to readily fabricate diffractive phase profiles that approximate the ideal diffractive profile. The ideal profile can be approximated in a discrete fashion, similar to the digital representation of an analog function. This discrete representation is called a multi-level profile and is theoretically analyzed in Section 3.

If diffractive surfaces are to become an accepted alternative to reflective and refractive surfaces, a well-defined process of actually fabricating the diffractive surface is needed. Section 4 describes a fabrication process that starts with a mathematical phase description of a diffractive phase profile and results in a fabricated multi-level diffractive surface. The fabrication process is best described in two different steps. The first step, described in detail, is to take the mathematical phase expression and generate from it a set of masks that contains the phase profile information. The second step is to transfer the phase profile information from the masks into the surface of the element specified by the lens design. This particular step is explained in a brief fashion, since the details of this procedure can be found in another report currently in preparation.

A multi-level diffractive phase profile is an additional option that should be seriously considered by lens designers. These profiles are not the solution to all problems; yet, in many instances, they can be used to improve on a design that consists solely of reflective and refractive elements. Section 5 describes some basic examples of cases where a diffractive phase profile can improve on the performance of a completely refractive design. The limitations of these diffractive phase profiles are quantified in order to give the lens designer a sense of the realm of applicability of diffractive profiles.

A lens designer relies heavily on the capabilities of a lens design program in arriving at a suitable solution to a particular problem. If a diffractive phase profile is to be considered in a design, the lens design program must have the capability to insert and optimize these diffractive profiles. Widely distributed lens design programs such as CODE V and ACCOS have the ability to insert diffractive surfaces into lens systems. These programs also have the capability to optimize the phase profiles of the diffractive surfaces in order to attain the best possible performance.

Section 6 describes in detail the process of inserting a diffractive surface in a lens design and the optimization of the diffractive element's phase profile. The format and terminology of the lens design program, CODE V, were chosen as the basis for the description of the procedure. CODE V was chosen because it is not only the most widely available program with the required capability, but also it is the program with which we are most familiar. The reader not familiar with CODE V can gain some insight into the process and peculiarities of designing a lens that contains diffractive surfaces.

2. THEORY

2.1 DIFFRACTION GRATING

The simplest example of a diffractive optical element is a linear grating. A variety of different types of gratings can be categorized based on the way by which the grating modulates the incident light field. Amplitude gratings, for example, modulate the incident light field by transmitting a certain percentage of the incident light and either absorbing or reflecting the rest. Phase gratings, on the other hand, transmit all of the incident light. The modulation is achieved by imparting to the incident light field a periodic phase delay. This periodic phase delay can be accomplished, as in a volume grating, by periodically modulating the refractive index of a material; or it can be accomplished, as in a surface relief grating, by periodically changing the physical thickness of a material.

Phase gratings have the advantage over amplitude gratings in that they can be made to diffract 100 percent of the incident light (of a given wavelength) into one diffraction order. This is a desirable, if not necessary, condition if a diffractive element is to be used in an optical system. Surface relief gratings have the advantage over volume gratings in that the diffraction efficiency falloff as a function of wavelength is minimized. This is a requirement if the element is to be used in an optical system designed to operate over a finite wavelength band. Furthermore, surface relief gratings can be fabricated in a mass-production environment similar to the integrated circuit fabrication process. For these reasons, we have concluded that surface relief phase gratings have the most potential for finding their way into commercial and military optical systems. The rest of this report will focus only on surface relief structures.

A surface relief phase grating is shown in Figure 2-1. The surface relief pattern, familiar to most people, is that of a conventional blazed grating. In order to analyze this structure, we will assume that the grating period T is large enough compared with the wavelength of the incident light so that the scalar approximation to Maxwell's equations can be used. The scalar theory is, in general, accurate when the grating period is greater than five wavelengths. The possible applications of diffractive structures, discussed later in this report, fall well within the regime of validity of the scalar approximation.

In the scalar approximation, the transmittance of the surface relief grating in Figure 2-1 can be described by

$$t(x) = \sum_{m=-\infty}^{\infty} \delta(x - mT) * rect(\frac{x}{T}) \exp(i2\pi\beta x) \tag{2.1}$$

where $\beta = (n-1)d/\lambda T$ and $*$ represents a convolution.

For an incident plane wave traveling in the z-direction, the far-field amplitude distribution is given by the Fourier transform $F(f)$ of the grating transmittance function $t(x)$

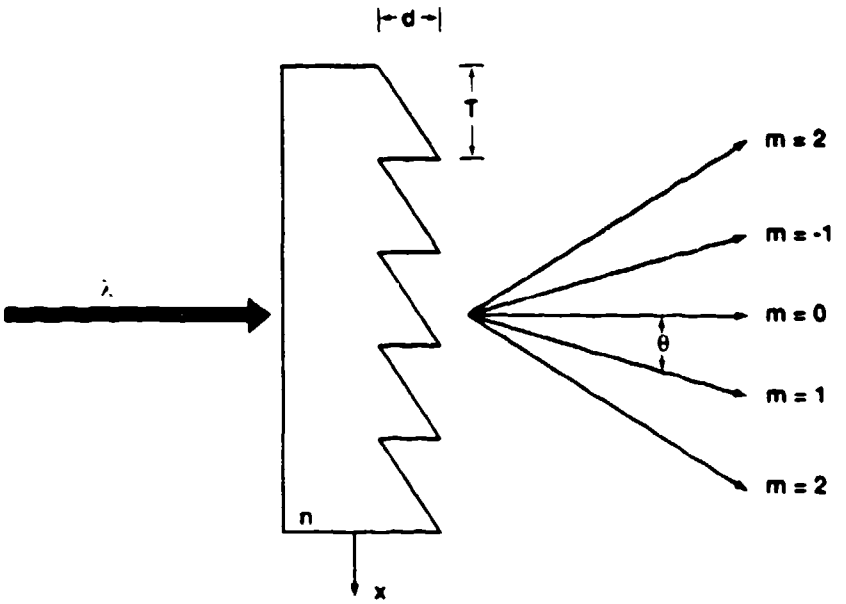


Figure 2-1. Surface relief phase grating.

$$F(f) = \sum_{m=-\infty}^{\infty} \delta(f - \frac{m}{T}) \frac{\sin(\pi T(\beta - f))}{\pi T(\beta - f)} \tag{2.2}$$

where $f = \sin(\theta)/\lambda$. It is apparent from Equation (2.2) that the amplitude of the m^{th} diffraction order is given by

$$a_m = \frac{\sin(\pi T(\beta - \frac{m}{T}))}{\pi T(\beta \frac{m}{T})}. \tag{2.3}$$

The diffraction efficiency of the m^{th} order is the absolute value of the amplitude of the m^{th} order squared

$$\eta_m = \left[\frac{\sin(\pi T(\beta - \frac{m}{T}))}{\pi T(\beta - \frac{m}{T})}\right]^2. \tag{2.4}$$

The diffraction order of interest, in general, is the first diffraction order. Setting m = 1 in Equation (2.4), the diffraction efficiency of the first order is given by

$$\eta_1 = \left[\frac{\sin(\pi(\beta T - 1))}{\pi(\beta T - 1)}\right]^2. \tag{2.5}$$

This equation predicts that, when $\beta = 1/T$, the diffraction efficiency of the first order will be 100 percent. Therefore, a properly constructed surface relief phase grating can diffract all of the incident light of a given wavelength into the first diffraction order.

Equation (2.5) also predicts that the first-order diffraction efficiency is both depth and wavelength dependent. A depth error in the fabrication process will result in a lower diffraction efficiency. Likewise, a change in wavelength will result in a diffraction efficiency decrease.

The depth dependence of the diffraction efficiency can be modeled by assuming a depth error of ϵd. The total grating depth is then

$$d = (1 + \epsilon)\frac{\lambda_0}{(n-1)}. \tag{2.6}$$

Substituting this value of d in Equation (2.5) results in a first- order diffraction efficiency given by

$$\eta_1 = \left[\frac{\sin(\pi\epsilon)}{\pi\epsilon}\right]^2. \tag{2.7}$$

This equation predicts that a ±5-percent depth error results in a diffraction efficiency falloff of less than 1 percent. In the majority of applications, this can be considered negligible. A depth error of ±5 percent corresponds to a physical depth error of approximately ±500 Angstroms for visible light. The etching technology used to fabricate these structures can be controlled to achieve depth tolerances of better than ±500 Angstroms.

The wavelength dependence of these elements becomes a concern when the element is to be used in an optical system operating over a finite wavelength band. There are, in fact, two wavelength-dependent effects unique to these structures. The first well-known effect, apparent from Equation (2.2) is that the first-order diffraction angle is wavelength dependent. Longer wavelengths are diffracted over larger angles than shorter wavelengths. This is a chromatic dispersion effect that will be discussed later in this report. The second effect is the wavelength dependence of the first-order diffraction efficiency.

Returning to Equation (2.5) and assuming that the diffraction efficiency is maximized for a wavelength λ_0 by setting $d = \lambda_0/(n-1)$ result in

$$\eta_1 = \left[\frac{\sin(\pi(\frac{\lambda_0}{\lambda} - 1))}{\pi(\frac{\lambda_0}{\lambda} - 1)}\right]^2. \tag{2.8}$$

This equation expresses the first-order diffraction efficiency at wavelength λ of an element optimized for wavelength λ_0. It is apparent from Figure 2-2, a plot of the diffraction efficiency as a function of wavelength, that the diffraction efficiency falloff is small for wavelengths close to λ_0 and is significant for large wavelength deviations.

For optical systems designed to operate in finite spectral bands, the integrated diffraction efficiency over the spectral band is the parameter of interest. The average diffraction efficiency over a finite bandwidth $\lambda_0 \pm \Delta\lambda$ is given by

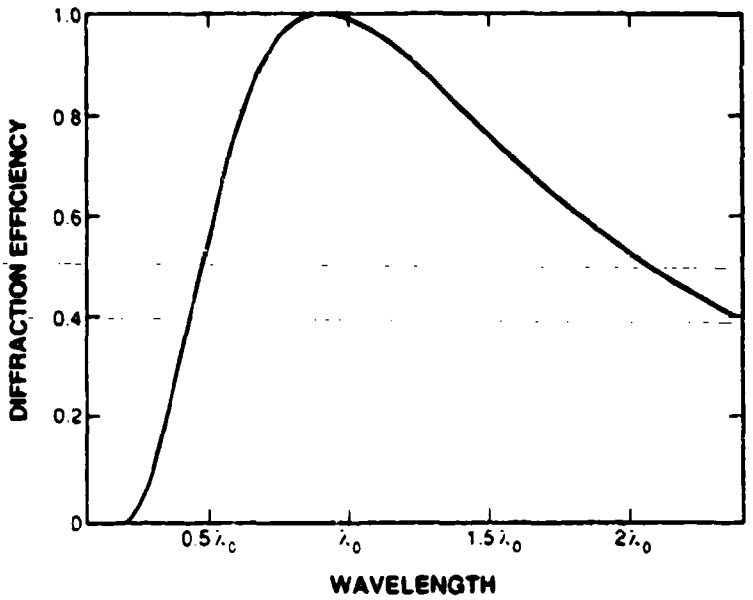


Figure 2-2. Plot of the diffraction efficiency as a function of wavelength.

$$\bar{\eta}_1 = \frac{1}{\Delta\lambda}\int_{\Delta\lambda} \eta_1(\lambda)d\lambda \tag{2.9}$$

which is approximately expressed by

$$\bar{\eta}_1 \approx \left[1 - \left(\frac{\pi\Delta\lambda}{6\lambda_0}\right)^2\right]. \tag{2.10}$$

Table 2-1 lists the average diffraction efficiency over various fractional bandwidths. The average diffraction efficiency remains above 95 percent for fractional bandwidths of up to 40 percent, but falls off rapidly for larger bandwidths. This effect is the most limiting constraint in using a diffractive element in a finite bandwidth system. The decrease in efficiency as a function of bandwidth has to be considered in a system design. The residual light that is not diffracted into the desired order is diffracted into different orders. This light manifests itself as a type of scatter at the image plane of an optical system. The amount of tolerable scatter is particular to the optical system's performance requirements. The lens designer has to establish the advantage or disadvantage of introducing a diffractive element into a design based on the performance goals of the optical system.

The scatter as a function of bandwidth introduced by a diffractive element is unlike the more familiar random scatter caused by inadequate surface polishing or surface defects. The scatter caused by the diffractive surface is deterministic. This scatter, or residual light, propagates in different diffraction orders. The amount of light at any wavelength, and in any given order, can be easily calculated from Equation (2.4). Figure 2-3 shows the amount of light in the various diffraction orders, at various wavelengths, for an element designed to have a 100-percent efficient

TABLE 2-1.

Average Diffraction Efficiency for Various Fractional Bandwidths

Fractional Bandwidth $(\Delta\lambda/\lambda_0)$	Efficiency $\bar{\eta}_1$
0.00	1.000
0.10	0.997
0.20	0.989
0.30	0.975
0.40	0.956
0.50	0.931
0.60	0.901

first order at λ_0. As the wavelength increases from λ_0, the residual light appears most strongly in the zero order. For decreasing wavelengths, the light appears in the second order. This residual light can be traced through an optical system to see how it is distributed at the image plane.

The blazed phase grating analyzed in this section was described by the transmittance function given in Equation (2.1). An equivalent way of expressing the transmittance function of a blazed grating is

$$t(x) = e^{i2\pi \cdot o_0 x}{}_a \tag{2.11}$$

where $|o_0 x|_a$ represents a linear function in x, modulo a, limited to values between $\pm a/2$.

The transmittance of a prism (the refractive counterpart of a grating) can be expressed as

$$t(x) = e^{i\frac{2\pi}{\lambda} o_0 x}. \tag{2.12}$$

The prism and grating phase functions are shown in Figure 2-4.

2.2 ARBITRARY PHASE PROFILE

In the previous section, the theory of a blazed phase grating having 100-percent diffraction efficiency in the first order was given. The effects of a change in depth, or a change in wavelength, were analyzed. However, a diffraction grating is of extremely limited usefulness in optical systems which require elements that focus or reshape the wavefront by desired amounts. In conventional optical systems, this is done refractively or reflectively by employing lenses and mirrors. What is required is the diffractive counterpart to these conventional optical elements.

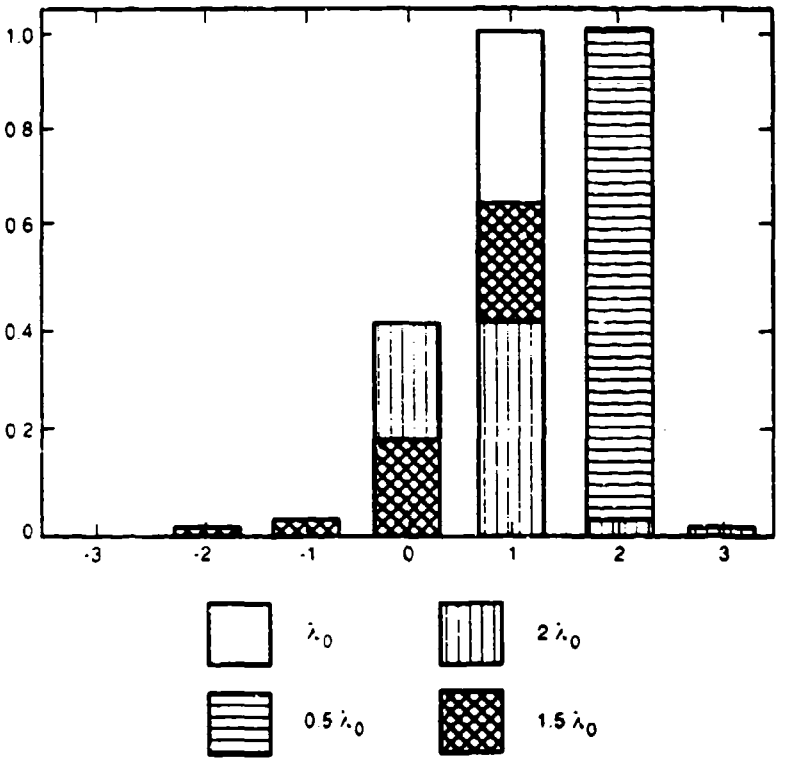


Figure 2-3. Plot of diffraction efficiency as a function of diffraction order for various wavelengths.

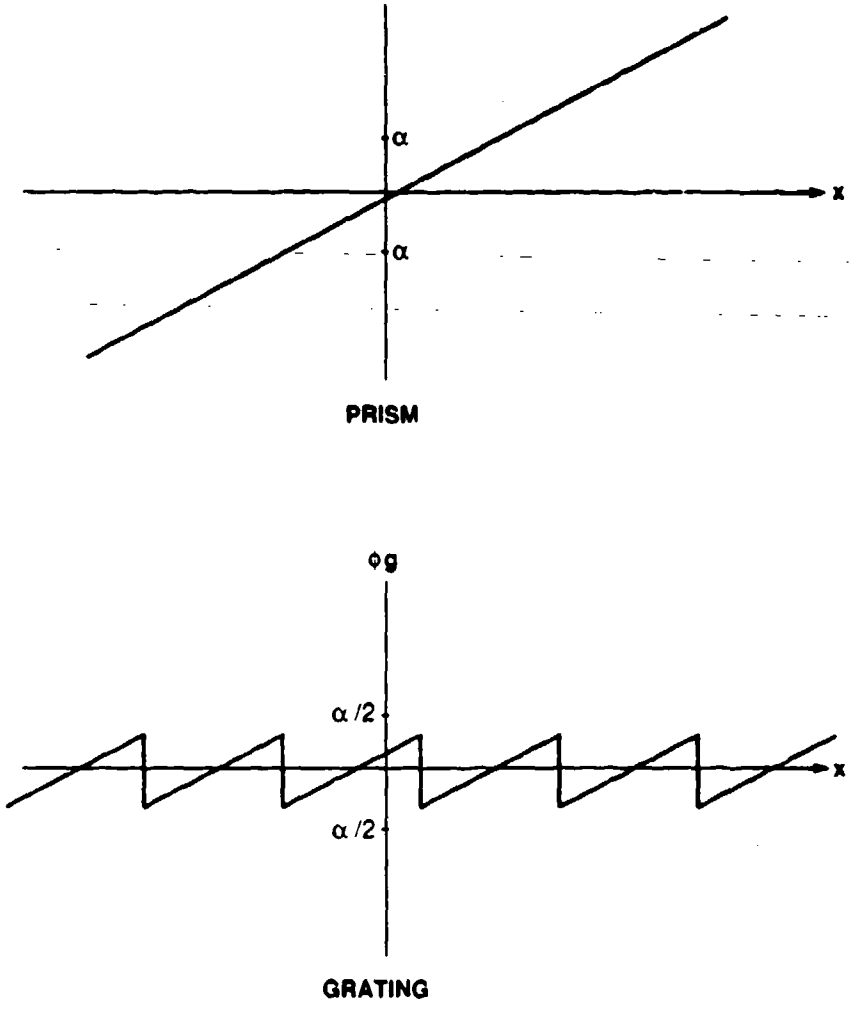


Figure 2-4. Phase functions of a prism and grating.

Consider a refractive element that can be described by a transmittance function

$$t_r(x) = e^{i2\pi\phi(x)} \tag{2.13}$$

where $\phi(x)$ is an arbitrary function of x. Can a general analogy, like the grating-prism analogy of the previous section, be made? What is the behavior of the diffractive counterpart with a transmittance function of

$$t_d(x) = e^{i2\pi\phi'(x)} \tag{2.14}$$

where $\phi'(x) = [\phi(x)]_a$? The refractive and diffractive phase functions for an arbitrary phase are plotted in Figure 2-5.

In order to understand the diffractive transmittance function of Equation (2.14), a nonlinear limiter analysis is used. The diffractive phase $\phi'(x)$ is plotted as a function of the refractive phase $\phi(x)$ in Figure 2-6. The diffractive phase, for generality, has been limited to values between $\pm a/2$. It is apparent from the figure that $\phi'(x)$ is periodic in $\phi(x)$ with a period equal to one. It follows that $\exp[2\pi\phi'(x)]$ is also periodic in $\phi(x)$ and can therefore be written as a generalized Fourier series

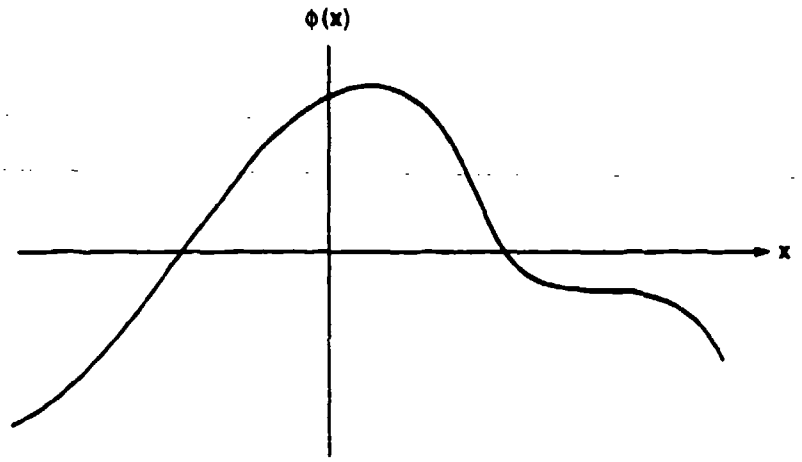
$$e^{i2\pi\phi'(x)} = \sum_{m=-\infty}^{\infty} c_m e^{i2\pi m\phi(x)} \tag{2.15}$$

where the coefficients c_m are given by

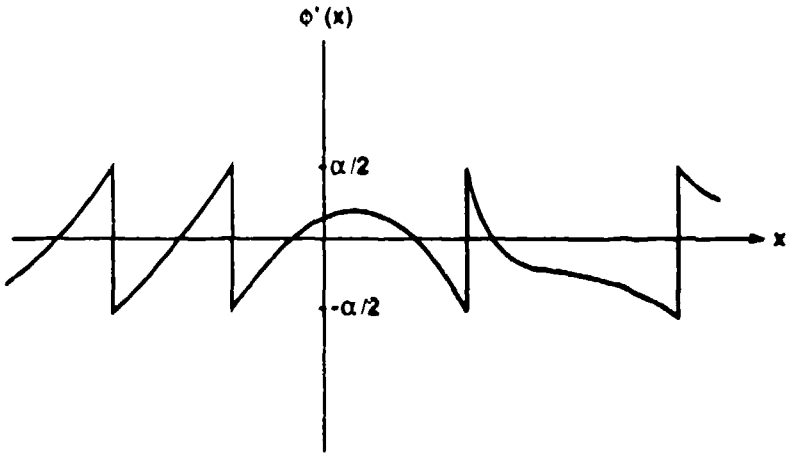
$$c_m = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi(a-m)\phi(x)} d\phi(x) = \frac{\sin(\pi(a-m))}{\pi(a-m)}. \tag{2.16}$$

Therefore, if $a = 1$, c_1 is equal to 1, and all the other c_m coefficients are zero. The exiting wavefront from this diffractive structure is identical to its refractive counterpart. It is interesting to note that the c_m coefficients are identical to the a_m coefficients of Equation (2.3) by setting $a = (n-1)d/\lambda_0$.

The wavelength and depth dependence of diffraction efficiency for any arbitrary phase diffractive structure is identical to the linear phase grating. The arbitrary diffractive phase element has *orders* similar to the grating. These *orders*, instead of being plane waves, take on more complicated wavefront profiles represented on the right-hand side of Equation (2.15).



REFRACTIVE PHASE



DIFFRACTIVE PHASE

124631-7

Figure 2-5. Comparison of the refractive and diffractive phase functions of an arbitrary phase profile.

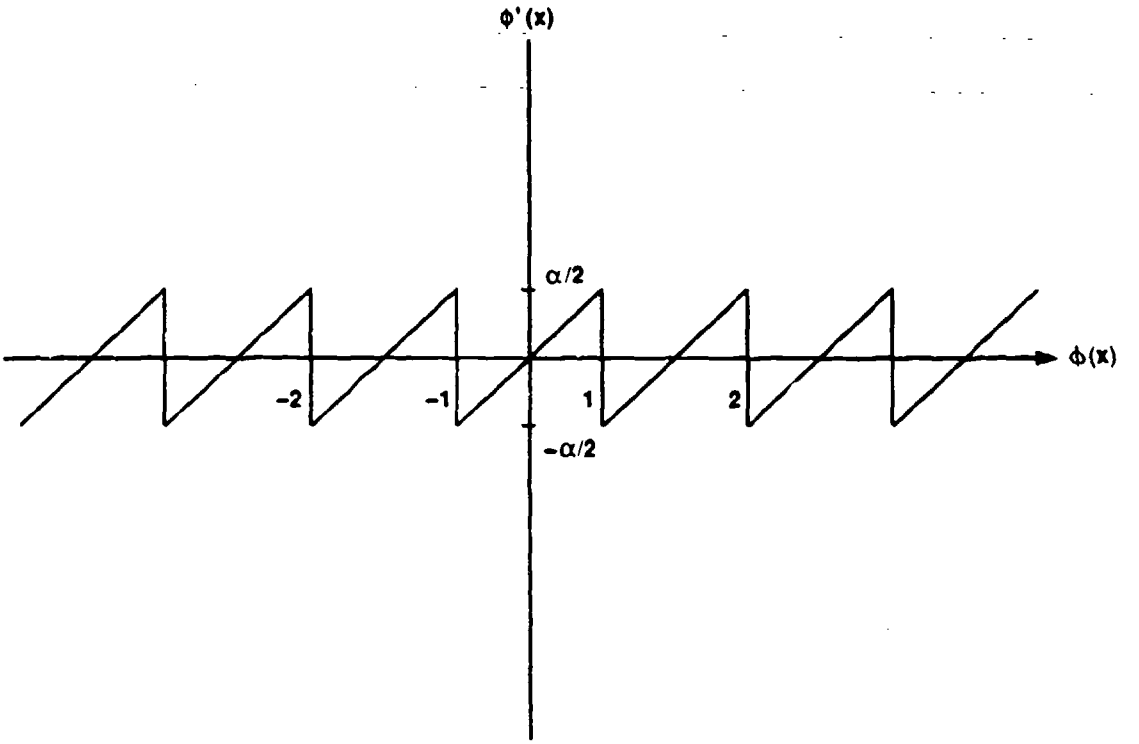


Figure 2-6. The diffractive phase $\phi'(x)$ plotted as a function of the refractive phase $\phi(x)$.

3. MULTI-LEVEL STRUCTURES

In Section 2 we showed that an arbitrary wavefront can be produced from a diffractive structure with 100-percent diffraction efficiency at the design wavelength. Unfortunately, this diffractive structure has a surface relief depth which varies continuously over every 2π phase interval. This phase profile, with a continuous depth, is not easily fabricated with any existing technology. A compromise has to be made between achievable diffraction efficiency and ease of fabrication.

A compromise that results in relatively high diffraction efficiency and ease of fabrication is a multi-level phase structure. Figure 3-1 shows a continuous phase grating profile compared with phase gratings with 2, 4, and 8 discrete phase levels. It is apparent from the figure that the larger the number of discrete phase levels, the better the approximation to the continuous phase profile. These multi-level phase profiles can be fabricated using standard semiconductor fabrication techniques. The fabrication process of multi-level structures will be described later in this report. The first question to be answered is the extent of the sacrifice in diffraction efficiency as a function of the number of discrete phase levels.

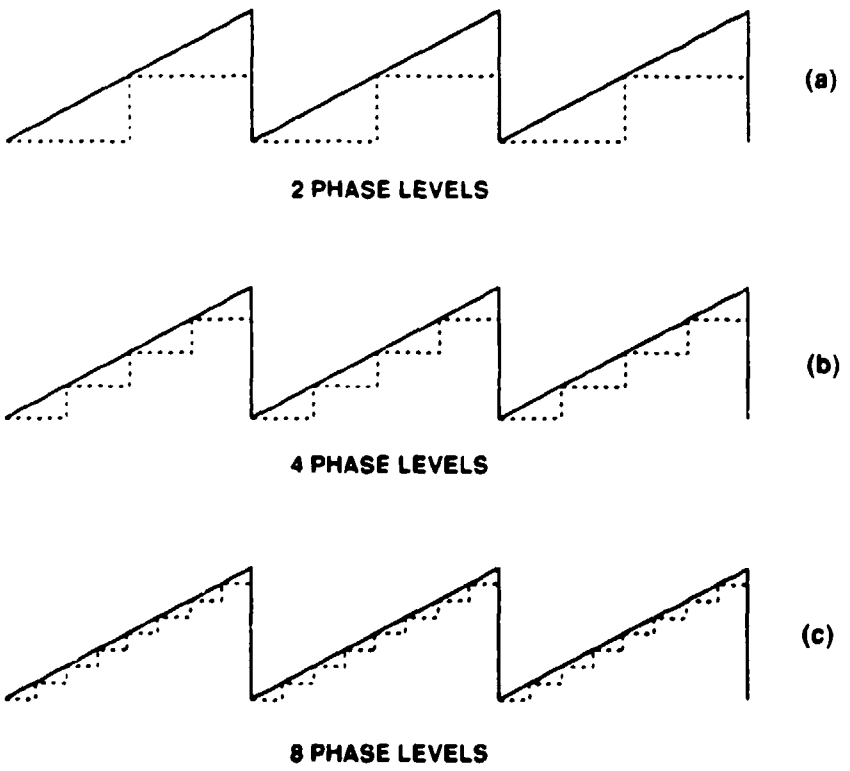
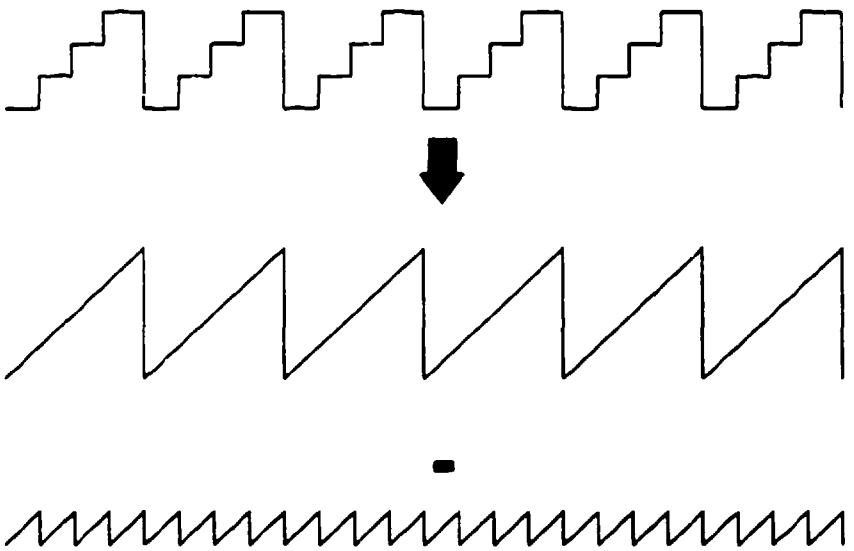


Figure 3-1. A continuous phase grating compared with 2, 4, and 8 discrete phase levels.

The diffraction efficiency of a multi-level structure can be simply derived by considering the multi-level structure as being equal to the desired continuous profile minus an error phase profile. The diffraction efficiency of the multi-level structure is then the efficiency of the desired continuous phase profile multiplied by the zero-order efficiency of the error phase structure. Figure 3-2 illustrates this concept for a 4-level structure If the number of discrete phase levels of the multi-level structure is N, then the error phase structure to be subtracted has a depth of d/N (where d is the desired continuous phase depth) and a periodicity of 1/N times that of the ideal structure. The resulting diffraction efficiency of the multi-level structure is then easily obtained by using Equation (2.4) and is given by

$$\eta_m^N = \left[\frac{\sin(\pi(\frac{(n-1)d}{\lambda} - m))}{\pi(\frac{(n-1)d}{\lambda} - m)}\right]^2 \left[\frac{\sin(\pi(\frac{(n-1)d}{\lambda N}))}{\pi(\frac{(n-1)d}{\lambda N})}\right]^2 . \tag{3.1}$$

This equation can be used to determine the diffraction efficiency of any multi-level profile at any wavelength and for any diffraction order.



124931-4

Figure 3-2. A multi-level phase structure can be analyzed by representing it as the difference of two continuous phase profiles.

As an example of how the number of phase levels affects the diffraction efficiency, consider a continuous phase structure designed to achieve 100-percent diffraction efficiency in the first order at the design wavelength. Equation (3.1) then reduces to the expression

$$\eta_1^N = \left[\frac{\sin(\pi/N)}{\pi/N}\right]^2 . \tag{3.2}$$

The continuous phase profile would achieve 100-percent diffraction efficiency, whereas a multi-level structure with N levels is reduced to that given by Equation (3.2). Table 3-1 lists the diffraction efficiency of a multi-level structure for various values of the number of phase levels. Two things to notice in the table are that for 16 phase levels the diffraction efficiency at the design wavelength is 99 percent, and that values of diffraction efficiency are highlighted for multi-level structures with N equal to a power of 2. The reason multi-level structures with a number of levels equal to a power of 2 are highlighted will become apparent in the next section.

TABLE 3-1.

Multi-level Diffraction Efficiency for Various Numbers of Phase Levels

Number of Levels N	First-Order Efficiency η_1^N
2	0 41
3	0.68
4	0.81
5	0.87
6	0.91
8	0.95
12	0.98
16	0.99

A 16-phase level structure achieving 99-percent diffraction efficiency is an element that could have advantageous implications in the design of many optical systems. The residual 1 percent of the light is diffracted into higher orders and manifests itself as scatter. In many optical systems, this is a tolerable amount of scatter. The fabrication of a 16-phase level structure, described in the following section, is relatively efficient due to the fact that only four processing iterations are required to produce the element.

4. MULTI-LEVEL FABRICATION

The fabrication of a multi-level diffractive element requires the same technology used in the production of integrated circuits. This fabrication process will be outlined in order to describe in a general fashion the steps involved in producing a multi-level diffractive element. A separate report describing in detail all the equipment and processing steps used in fabrication of multi-level elements is in preparation.

The first step involved in fabricating a multi-level element is to mathematically describe the ideal diffractive phase profile that is to be approximated in a multi-level fashion. The simplest case, for example, is a grating of period T, designed to operate at a wavelength λ_0. The phase function for this grating can be mathematically described by

$$\phi(x) = \frac{2\pi}{\lambda_0} ax \tag{4.1}$$

where $a = \lambda_0/T$.

A phase function having more complexity than a simple grating can be mathematically described in a general way by expanding it in a power series

$$\phi(x,y) = \frac{2\pi}{\lambda_0} \sum_{n,m} a_{nm} x^n y^m. \tag{4.2}$$

This equation represents a general phase function in the spatial coordinates (x,y). The number of terms retained in the power series determines how well of an approximation the series is to the actual phase desired. The values of the a_{nm} coefficients are optimized to make the series expansion best approximate the desired phase. For example, the grating phase of Equation (4.1) is represented by Equation (4.2) where all the a_{nm} coefficients are zero, except for a_{10} which is equal to λ_0/T.

The majority of cases in optical design require phase functions that are circularly symmetric; these phase functions can also be described by a power series expansion

$$\phi(r) = \frac{2\pi}{\lambda_0} \sum_{p} a_p r^p \tag{4.3}$$

where r is the radial coordinate. The optical axis of the lens system is at the radial coordinate r = 0. The values of the a_p coefficients determine the functional form of the radially dependent phase.

The next step in the fabrication process, once the phase function is mathematically determined, is to create a set of lithographic masks which are produced by standard pattern generators used in the integrated circuit industry. Pattern generators, either optical or electron beam, expose a thin layer of photoresist which resides on a chrome-covered quartz substrate. The exposed photoresist is then washed off the chrome-coated substrate, leaving the pattern in the remaining unexposed photoresist. The pattern is then transferred to the chrome by etching away the chrome that is not covered by the remaining photoresist. Once the chrome has been patterned, the remaining

photoresist is washed away, resulting in a finished lithographic mask. The final product is a binary amplitude mask that transmits light where the pattern was exposed, and reflects any incident light where there was no exposure.

4.1 GRATING FABRICATION

To illustrate the purpose of these lithographic masks in the fabrication of a multi-level element, consider the simplest case of a grating. The final grating that is desired has the surface relief profile shown in Figure 3-1(a). The coarsest approximation to the desired grating profile, also shown in Figure 3-1(a), is a binary phase profile. A lithographic mask can be easily produced with a binary amplitude grating pattern having the desired period and a 50-percent duty cycle (i.e., 50 percent of the light is transmitted). What remains is to transfer the amplitude pattern contained on the lithographic mask onto an optically transmissive substrate, and convert the amplitude pattern into a surface relief pattern.

Figure 4-1 illustrates the process of fabricating a binary surface relief grating, starting with the binary amplitude lithographic mask. A substrate of the desired material is coated with a thin layer of photoresist. The lithographic mask is then placed in intimate contact with the substrate and illuminated from above with an ultraviolet exposure lamp. The photoresist is developed, washing away the exposed resist and leaving the binary grating pattern in the remaining photoresist. This photoresist will act as an etch stop, like in the lithographic mask process, except that now the substrate material has to be etched instead of chrome.

The most reliable and accurate way to etch many optical substrate materials is to use reactive ion etching. The process of reactive ion etching anisotropically etches materials at very repeatable rates. The desired etch depth can be obtained very accurately. The anisotropic nature of the process assures a vertical etch, resulting in a truly binary surface relief profile. Once the substrate has been reactively ion etched to the desired depth, the remaining photoresist is stripped away, leaving a binary phase surface relief grating. In the case of a binary phase profile, Equation (3.2) predicts a maximum first-order diffraction efficiency of 40.5 percent for an etch depth of $d = \lambda_0/2(n-1)$.

Imagine repeating the process described above on the same substrate, except this time using a lithographic mask having twice the period of the first mask. Figure 4-2 illustrates this process. The binary phase element is recoated with photoresist and exposed using the lithographic mask #2 that has a period twice that of the first mask. After developing and washing away the exposed photoresist, the substrate is reactive ion etched to a depth half that of the first etch [i.e., $d = \lambda_0/4(n-1)$]. Removal of the remaining photoresist results in a 4-level approximation to the desired profile. The 4-level phase element has a first-order diffraction efficiency, predicted by Equation (3.2), of 81 percent. One can imagine repeating the process a third and fourth time with lithographic masks having periods of one-quarter and one-eighth that of the first mask, and etching the substrate to depths of one-quarter and one-eighth the depth of the first etch. The successive etches result in elements having 8 and 16 phase levels. The first-order diffraction efficiency, after the third and fourth etches, is predicted from Equation (3.2) to be 95 and 99 percent, respectively.

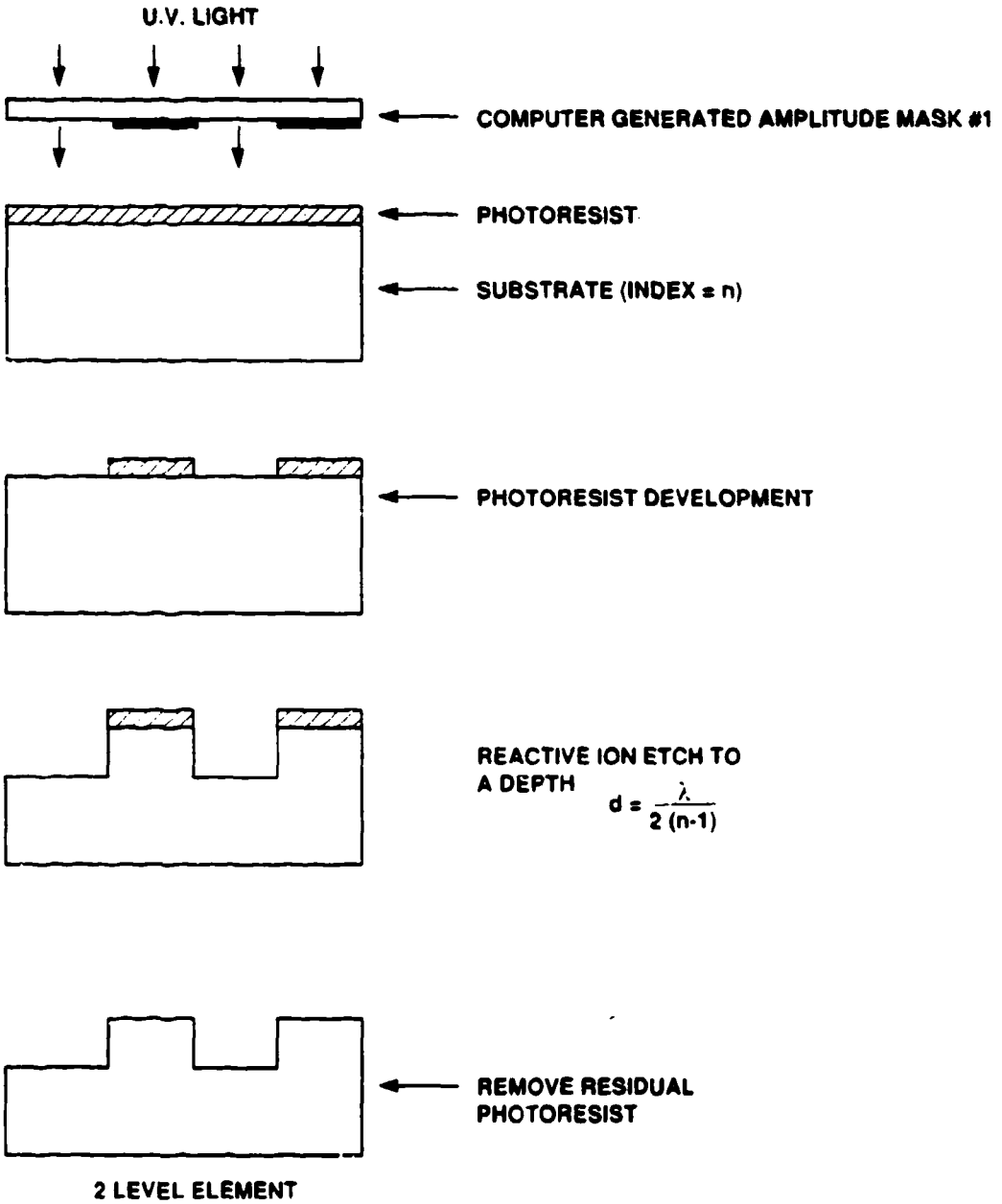


Figure 4-1. Illustration of the fabrication of a binary surface relief grating.

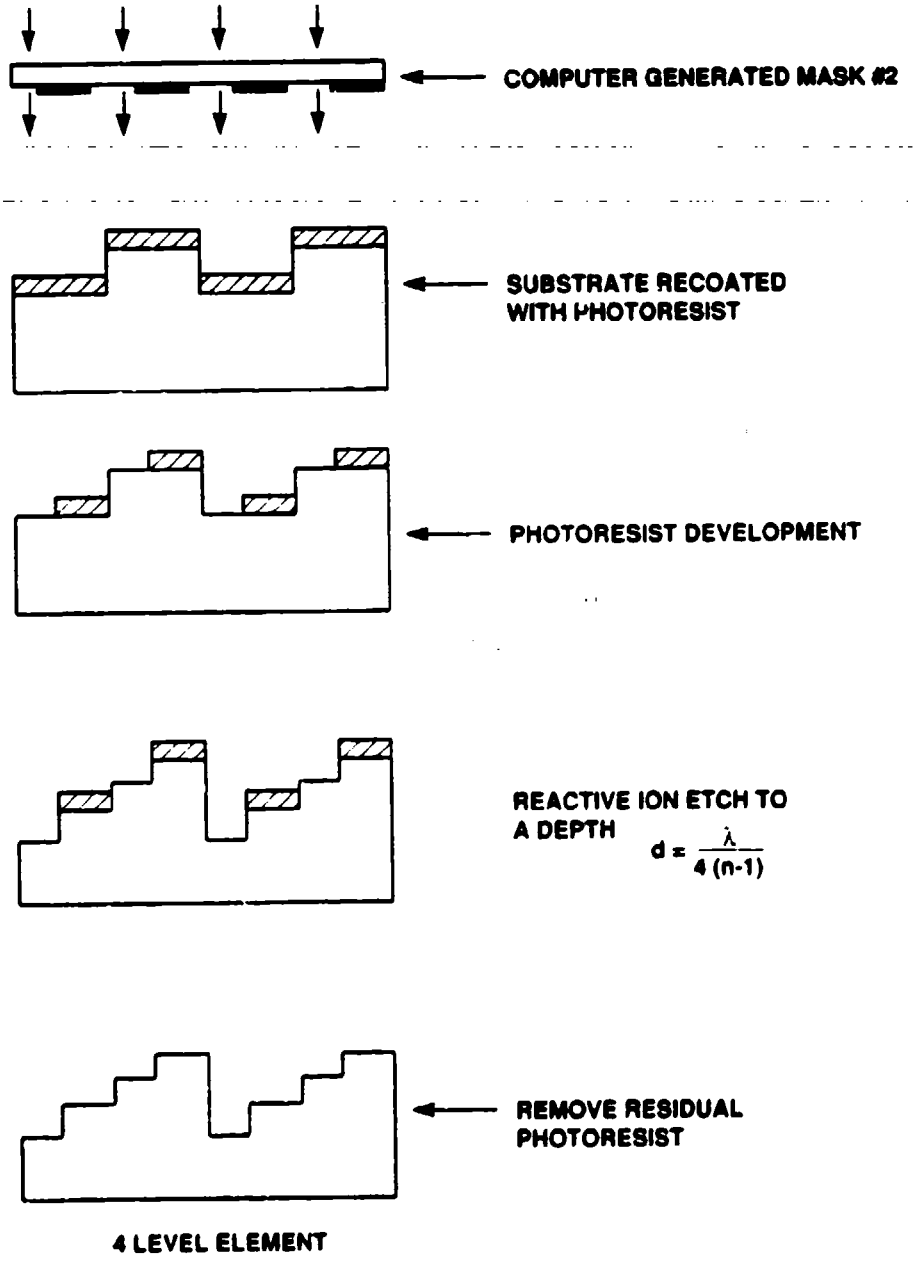


Figure 4-2. Illustration of the fabrication of a 4-level surface relief grating.

After only four processing iterations, a 16-phase level approximation to the continuous case can be obtained. A similar iteration process to that described here is used in the fabrication of integrated circuits. The process can be carried out in parallel, producing many elements simultaneously, in a cost-effective manor.

The one major difference between fabricating a binary phase element and a multi-level element is that, after the first etching step, the second and subsequent lithographic masks have to be accurately aligned to the existing pattern on the substrate. Alignment is accomplished using another tool standard to the integrated circuit industry, a mask aligner. Mask aligners are commercially available with varying degrees of sophistication. The majority of aligners can place in registry the mask and substrate to submicron tolerances. This degree of alignment accuracy allows for the fabrication of many useful multi-level diffractive elements.

Some instances in a lens design may require or prefer the diffractive surface to reside on a substrate surface that is not flat. The process, as described, necessitated a flat substrate to obtain intimate contact between the lithograghic masks and the substrate. The condition of intimate contact can be relaxed, depending on the feature sizes of the lithographic masks. If the mask and substrate are not in intimate contact, the ultraviolet exposure light will diffract from the mask, blurring the pattern in the photoresist. In many instances, particularly for diffractive elements designed for use in the infrared, the mask's feature sizes are large enough to allow for a significant distance between the mask and the substrate. This is a point the lens designer must be aware of in a system design.

In summary, the fabrication of a multi-level surface relief grating requires a set of lithographic masks and standard integrated circuit fabrication equipment. A set of M properly designed lithographic masks results in a multi-level surface relief grating with 2^M phase levels. The optimum etch depth for the M^{th} mask pattern is $d_M = \lambda_0 / 2^M (n-1)$.

4.2 ARBITRARY PHASE FABRICATION

The design of a set of lithographic masks used in the fabrication of a multi-level grating was easy to visualize. Mask #M simply had a 50-percent duty cycle and a period equal to one-half that of mask #(M - 1). The design of a set of lithographic masks used in the fabrication of multi-level structures approximating the general phase functions described by Equations (4.2) and (4.3) is slightly more involved.

Let us consider the case of designing a set of masks to be used in the fabrication of a circularly symmetric phase function described by Equation (4.3). The coarsest approximation to this phase profile is again a binary phase profile. The lithographic mask needed to produce this binary phase element will have a circularly symmetric amplitude profile (i.e., a set of alternately transmitting and reflecting annuli). The positions and widths of these annuli are determined from Equation (4.3). The phase at the center of the pattern is zero. By stepping out in radius, the phase $\phi(r)$ either increases or decreases. The magnitude of the phase will reach π at some value of r which is the first radial position on the lithographic mask where an amplitude transition occurs.

Continuing the process of stepping out in radius results in radial locations where the phase function $\phi(r)$ takes on values that are integer multiples of π. These are the subsequent radial positions where amplitude transitions occur on the first lithographic mask. The process of stepping out in radius is continued until the maximum radial value of the pattern to be written is reached.

The resulting set of radial values is sufficient information to write the lithographic mask. A computer program called Mann 53, and written by the Binary Optics Group at Lincoln Laboratory, is able to take the set of radial values and properly format the data such that they can be read by a Mebes electron beam pattern generator. The Mann 53 program creates a data tape that can be sent to various lithographic mask vendors for mask fabrication.

The process of designing mask #M, where M is greater than 1, is carried out in a fashion similar to the first mask. Again, the phase function $\phi(r)$ is monitored as a function of radius. At some radial distance, the magnitude of the phase will reach the value π/M. This is the radial position where the first amplitude transition will occur on mask #M. Subsequent amplitude transition points for mask #M will occur at radial values where the phase is an integer multiple of π/M. The process of stepping out in radius is continued until the maximum radius of the pattern is reached. The Mann 53 program is used to format these radial values and produce data tapes in a manner similar to the case of mask #1.

The process described above is straightforward and applicable to any radially symmetric phase profile, except for a certain subset of phase profiles where an anomaly can occur. This subset of phase functions is one where the first derivative of the phase with respect to r is zero at some radial position that corresponds to a transition point. For this subset of functions, a little more care has to be taken in locating the transition points. If the first derivative $d\phi/dr$ is zero at a transition point, the question arises whether or not this point should correspond to a transition. The way to determine whether a transition should or should not occur is to check the second derivative. If $d^2\phi/dr^2$ is also zero at the radial point in question, the phase is at an inflection point and a transition should be located there. On the other hand, if the second derivative is zero, the point corresponds to a local maximum or minimum of the phase, and a transition should not be located there.

Figure 4-3 illustrates an example of a phase function that contains a local maximum at a radial position r_3: this position also happens to correspond to a value of the phase equal to 3π. Since the second derivative is not zero at this point, a transition point should not be located there.

The procedure described above is also applicable for noncircularly symmetric phase functions described by Equation (4.2). The basic idea is the same, yet the determination of the transition point locations can become quite computationally intensive. Perkin-Elmer Corp. has devised a software package that can take an arbitrary two-dimensional phase profile and construct from it a set of proper lithographic masks.

Once the proper set of lithographic masks is designed and constructed, the fabrication of the multi-level diffractive phase element is identical to the fabrication process of the multi-level phase grating described above. The first mask pattern is reactively ion etched to a π phase depth.

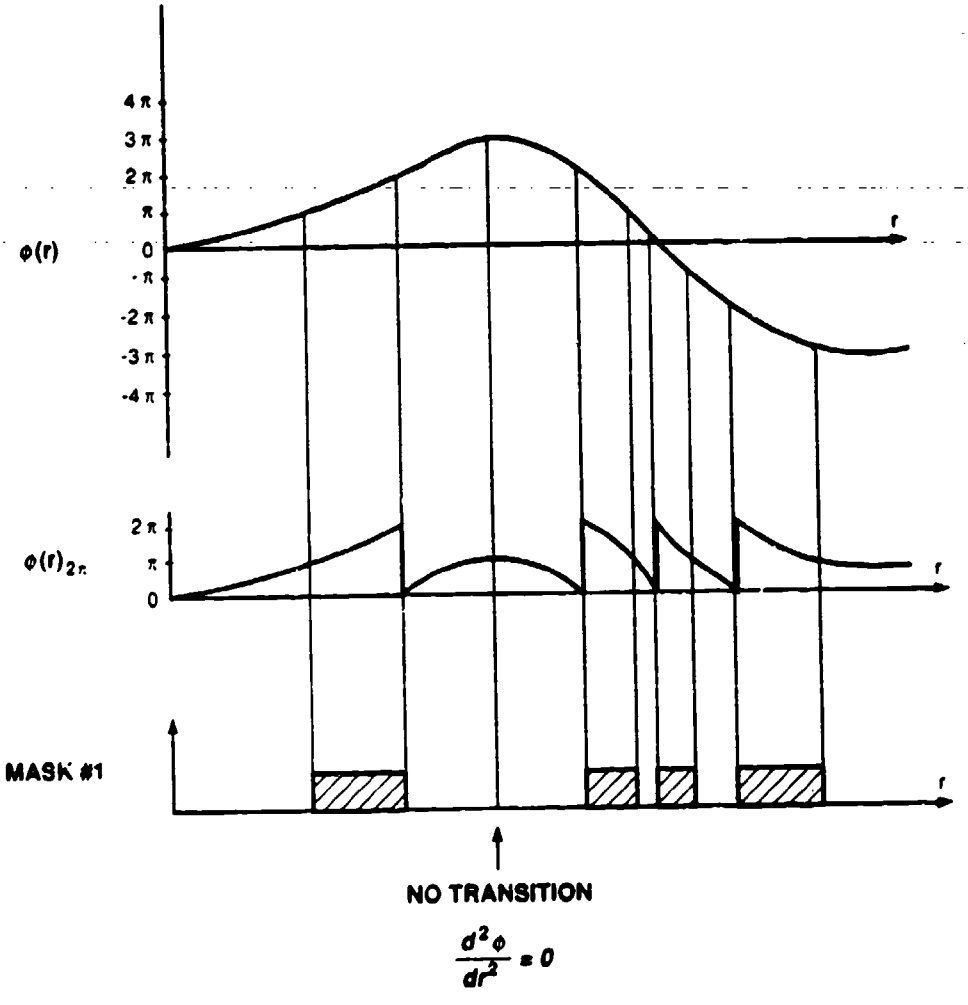


Figure Example of a phase function that contains a local maximum.

Subsequent mask patterns are aligned and etched to a phase depth of $\pi/2^M$. The procedure for determining transition point location and etch depth is summarized in Figure 4-4.

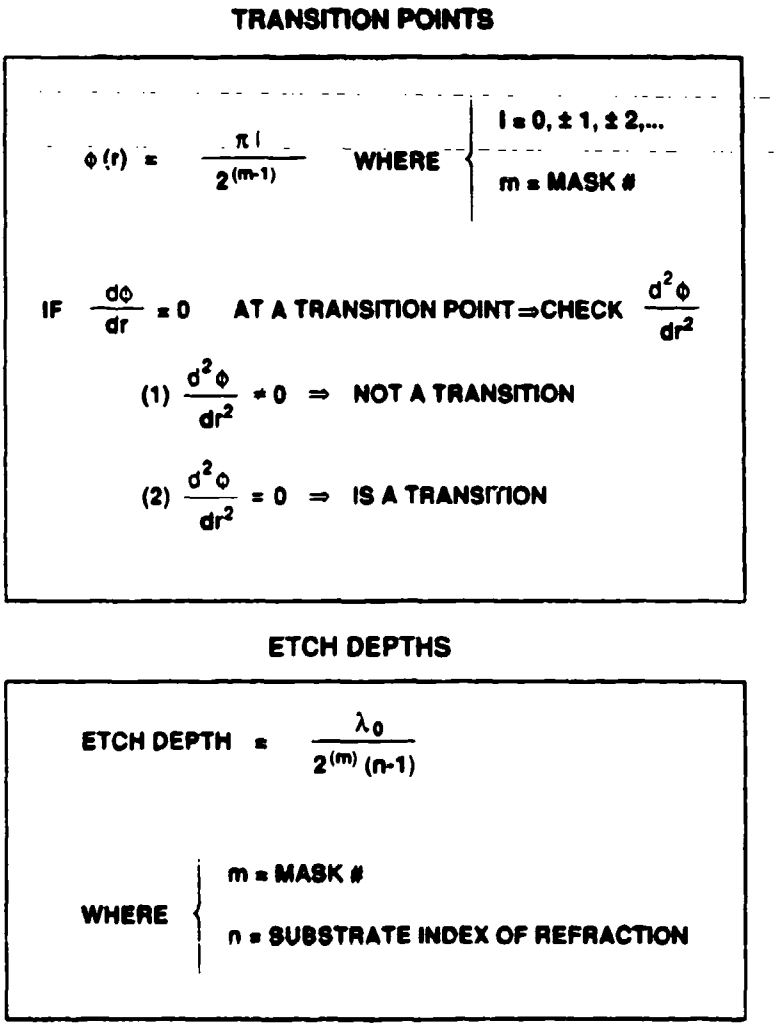


Figure 4-4. Summary of the procedure for determining transition point locations and etch depths.

5. APPLICATIONS OF MULTI-LEVEL DIFFRACTIVE PROFILES

The fabrication of multi-level diffractive phase profiles has been described in the previous section. Here, we attempt to elucidate the potential as well as the limitations of using a diffractive surface in the design of an optical system. Hopefully, a lens designer will be able to determine whether or not a multi-level diffractive surface will be advantageous in any particular design.

We begin with a description of the focusing properties of a completely diffractive lens (Section 5.1). A completely diffractive lens is shown to suffer from severe chromatic aberration, limiting its usefulness in any optical system that has to operate over a finite wavelength band.

In Section 5.2 we show how the chromatic dispersion of the diffractive lens can be used to one's advantage by combining it with a refractive lens element. The combination of a refractive lens and a diffractive profile is shown to be a very powerful concept in the design of optical elements.

The idea of using a diffractive profile to correct for the inherent spherical aberration of a single spherical lens is described in Section 5.3. For a monochromatic system, the spherical aberration can be completely eliminated; for a finite waveband system, it can be reduced. The amount of correction possible is shown to depend on the fractional operating bandwidth of the system.

Finally, in Section 5.4 we show how the diffractive pattern that corrects for spherical aberration can be combined with the diffractive pattern that corrects for chromatic aberration. The result of this combination is shown to be a single diffractive pattern that corrects for both spherical and chromatic aberration.

5.1 DIFFRACTIVE LENS

The simplest example of a diffractive phase profile, other than a linear grating, is a quadratic phase profile. In the paraxial approximation, a quadratic phase profile is a lens. A one-dimensional diffractive lens, having a quadratic phase profile, is illustrated in Figure 5-1. The lens has a focal length F_0 for wavelength λ_0, and forms an image at z_i of an object located at z_0. The transmittance function for this lens, assuming 100-percent diffraction efficiency in the first order for wavelength λ_0, is described by

$$t(x_0) = e^{-i\pi c x_0^2} \tag{5.1}$$

where $c = 1/\lambda_0 F_0$ is a constant. By performing a Fresnel diffraction calculation, it is shown that the first-order lens equation for a diffractive lens is

$$\frac{1}{z_i} = \lambda c - \frac{1}{z_0}. \tag{5.2}$$

From this equation it is apparent that the image distance z_i is strongly dependent on wavelength. Setting $c = 1/\lambda_0 F_0$ in Equation (5.2) results in the expression

$$\frac{1}{F(\lambda)} = \frac{1}{z_0} + \frac{1}{z_i} \tag{5.3}$$

where $F(\lambda) = \lambda_0 F_0/\lambda$. This expression looks conspicuously like the first-order lens equation for refractive lenses. The only difference is that the focal length of the lens, instead of being constant, depends inversely on the wavelength. The result of this wavelength dependence is severe chromatic aberration.

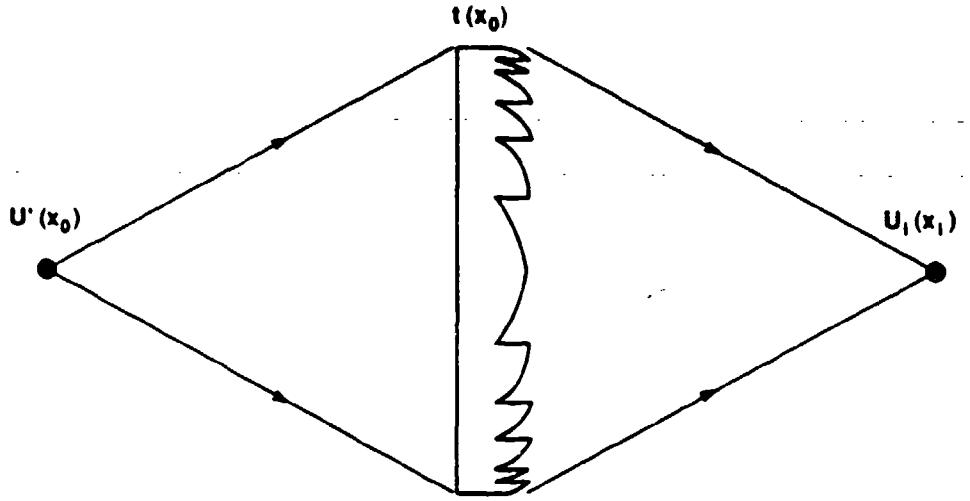


Figure 5-1. Illustration of a one-dimensional diffractive lens.

The amount of chromatic aberration in a diffractive lens can be quantified by setting the object distance to infinity. The output wavefront from the diffractive lens is then described by the lens transmittance function of Equation (5.1). An ideal wavefront, having no chromatic aberration, is given by

$$U_i(x) = e^{-i\frac{\pi x^2}{\lambda F_0}}. \tag{5.4}$$

The output wavefront from the diffractive lens can be rewritten as

$$U_0(x) = U_i(x)U_a(x) = e^{-i\frac{\pi x^2}{\lambda F_0}} e^{-i\frac{\pi}{F_0}(\frac{1}{\lambda_0} - \frac{1}{\lambda})x^2} \tag{5.5}$$

where U_i is the ideal wavefront and U_a is the aberration component of the wavefront. The phase function ϕ_ϵ of U_a,

$$\phi_\epsilon(x) = \frac{1}{2F_0}(\frac{1}{\lambda_0} - \frac{1}{\lambda})x^2 \tag{5.6}$$

is the chromatic phase aberration of a diffractive lens. Note that, for the wavelength $\lambda = \lambda_0$, the phase aberration is zero. For any wavelength other than λ_0, the phase aberration is nonzero.

An expression for the maximum amount of chromatic phase aberration present in a diffractive lens of aperture diameter A, and operating over a bandwidth $\Delta\lambda$ centered at λ_0, can be written as

$$max_{\epsilon} = \frac{A}{8(F/\#)}\left(\frac{\Delta\lambda}{\lambda^2}\right) \tag{5.7}$$

where F/# is the f-number of the lens. Note that the amount of chromatic aberration is proportional to both the fractional bandwidth and the number of wavelengths across the aperture. As an example, consider an F/2 lens, operating over the 8- to 12-μm wavelength band, and having a 75-mm aperture. The maximum phase error due to chromatic aberration of this lens is 234 waves! This is an intolerable amount of chromatic aberration. Clearly, the usefulness of a completely diffractive lens in an optical system operating over a finite wavelength band is limited.

5.2 REFRACTIVE/DIFFRACTIVE ELEMENTS

The previous analysis made it quite apparent that completely diffractive lenses cannot be used in finite wavelength band systems. A solution to this dilemma is to combine a refractive lens with a diffractive lens profile. It should be pointed out that this is not a detraction from using diffractive lens profiles in an optical system — rather, it is an attraction. A completely diffractive lens would have to reside on an optically flat substrate to retain good performance. The cost differential between an optically flat refractive substrate and a refractive lens with spherical surfaces is negligible. There is no cost advantage in using completely diffractive lens elements. Furthermore, the smaller the F/# of a diffractive lens, the finer the features become in the diffractive profile and the more difficult the element becomes to construct. By combining a refractive lens and a diffractive lens, the refractive lens can do the majority of the focusing, substantially increasing the feature sizes required in the diffractive lens profile.

The most compelling reason to consider refractive/diffractive elements is that the chromatic dispersion of the diffractive surface can be used to negate the chromatic dispersion of refractive lenses. The etching of a properly designed diffractive lens profile on a surface of a dispersive refractive lens can result in a single-lens element that has virtually no dispersion. This concept is very powerful, especially in wavelength regions where the number of available materials that have suitable transmittance characteristics is limited.

The index of refraction of a refractive lens can be modeled in a linear approximation as

$$n(\lambda) = n_0 - D(\lambda - \lambda_0) \tag{5.8}$$

where D is the dispersion constant and λ_0 is the center wavelength of the wavelength band. The focal length of this dispersive refractive lens, F_r, is given by

$$\frac{1}{F_r(\lambda)} = \frac{1}{F_{r0}} - \frac{D(\lambda - \lambda_0)}{(n_0 - 1)F_{r0}} \tag{5.9}$$

where F_{r0} is the focal length of wavelength λ_0. The focal length of a diffractive lens F_d was previously determined to be

$$F_d(\lambda) = \frac{\lambda_0}{\lambda} F_{d0}. \tag{5.10}$$

A refractive/diffractive combination of the lenses described by Equations (5.9) and (5.10) results in a lens with a focal length $F(\lambda)$ given by

$$\frac{1}{F(\lambda)} = \frac{1}{F_r(\lambda)} + \frac{1}{F_d(\lambda)}. \tag{5.11}$$

Substituting Equations (5.9) and (5.10) into Equation (5.11) results in

$$\frac{1}{F(\lambda)} = \frac{\lambda}{\lambda_0 F_{d0}} + \frac{1}{F_{r0}} - \frac{D(\lambda - \lambda_0)}{(n_0 - 1)F_{r0}}. \tag{5.12}$$

Now, if the ratio of the focal length of the diffractive surface to that of the refractive lens is set to

$$\frac{F_{d0}}{F_{r0}} = \frac{(n_0 - 1)}{\lambda_0 D} \tag{5.13}$$

the resulting focal length of the combined refractive/diffractive element is given by

$$\frac{1}{F(\lambda)} = \frac{1}{F_{r0}} + \frac{1}{F_{d0}}. \tag{5.14}$$

The result of Equation (5.14) is an element that has a focal length independent of wavelength. By satisfying the condition of Equation (5.13), a refractive/diffractive lens can be made that, in a linear approximation, has no chromatic dispersion. It is important to note that both the refractive and diffractive components of the combination element have focal powers of the same sign. This is unlike a conventional achromatic doublet lens made from two refractive lenses of different materials. In a conventional refractive achromat, the two lenses must have focal powers of opposite sign. This difference between conventional and refractive/diffractive achromats is due to the fact that the focal length of a diffractive lens is shorter for longer wavelengths, while all refractive lenses have focal lengths that are longer for longer wavelengths.

Consider an example of the utility of a refractive/diffractive lens in correcting for chromatic aberration. KrFl lasers are fast becoming useful tools in microlithography and medicine. KrFl lasers emit ultraviolet light in a 2-nm wavelength band centered at 248 nm. The only durable material that can be polished into refractive lenses at this wavelength is fused silica. However, fused silica is very wavelength dispersive at 248 nm. A conventional refractive achromatic doublet is difficult to fabricate due to the lack of materials other than fused silica.

A refractive/diffractive achromatic can be readily fabricated. The dispersion constant of fused silica at 248 nm is $D = 6 \times 10^{-4}$ nm^{-1}. Using this value of D in Equation (5.13) results in

$$\frac{F_{d0}}{F_{r0}} = 3.4. \tag{5.15}$$

Therefore, if a diffractive lens is etched into the surface of a fused silica lens such that Equation (5.15) is satisfied, the resulting combination will have minimum chromatic dispersion. Figure 5-2(a) shows the phase aberration in waves across the aperture of a fused silica lens. The lens has a 1-in-diam. aperture and a 9-in focal length. Approximately 3 waves of chromatic aberration are present at the edge of the aperture over a 2-nm bandwidth. The placement of a diffractive lens profile, that satisfies Equation (5.15), on a surface of the fused silica lens results in the chromatic phase error shown in Figure 5-2(b). The maximum chromatic phase error has been reduced from 3 waves of aberration to less than 0.02 wave. This is a 150-fold improvement in wavefront error!

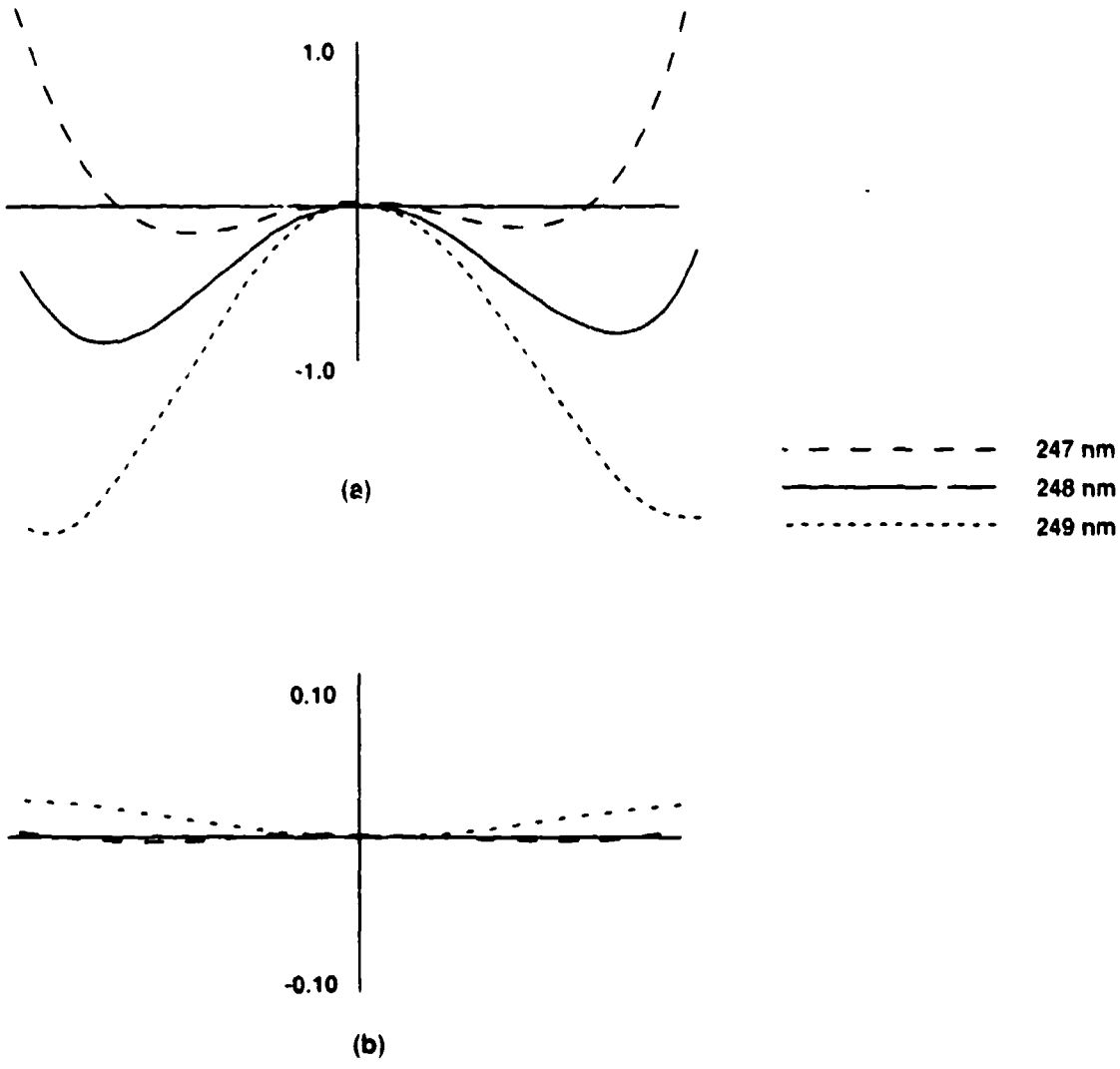


Figure 5-2. The phase aberration (a) of a refractive fused silica lens, and (b) of the same lens with diffractive aberration correction.

The concept of a refractive/diffractive achromat has been experimentally verified in the visible region of the spectrum. A 1-in-diam. fused silica (quartz) lens, with a 6-in focal length, was used to image an Air Force resolution target illuminated with a source emitting from 450 to 700 nm. An identical lens, with the properly designed diffractive profile etched into one of its surfaces, was also tested. The results are shown in Figure 5-3. The diffractively corrected lens is obviously far superior in performance than the refractive lens. The figure also shows that the refractive/diffractive combination is far superior for off-axis points. This can be understood by realizing that the amount of lateral chromatic aberration depends on the separation of the two lens components. In the case of a refractive/diffractive achromat, the two lens components are placed as close in proximity as possible, thus minimizing lateral chromatic effects.

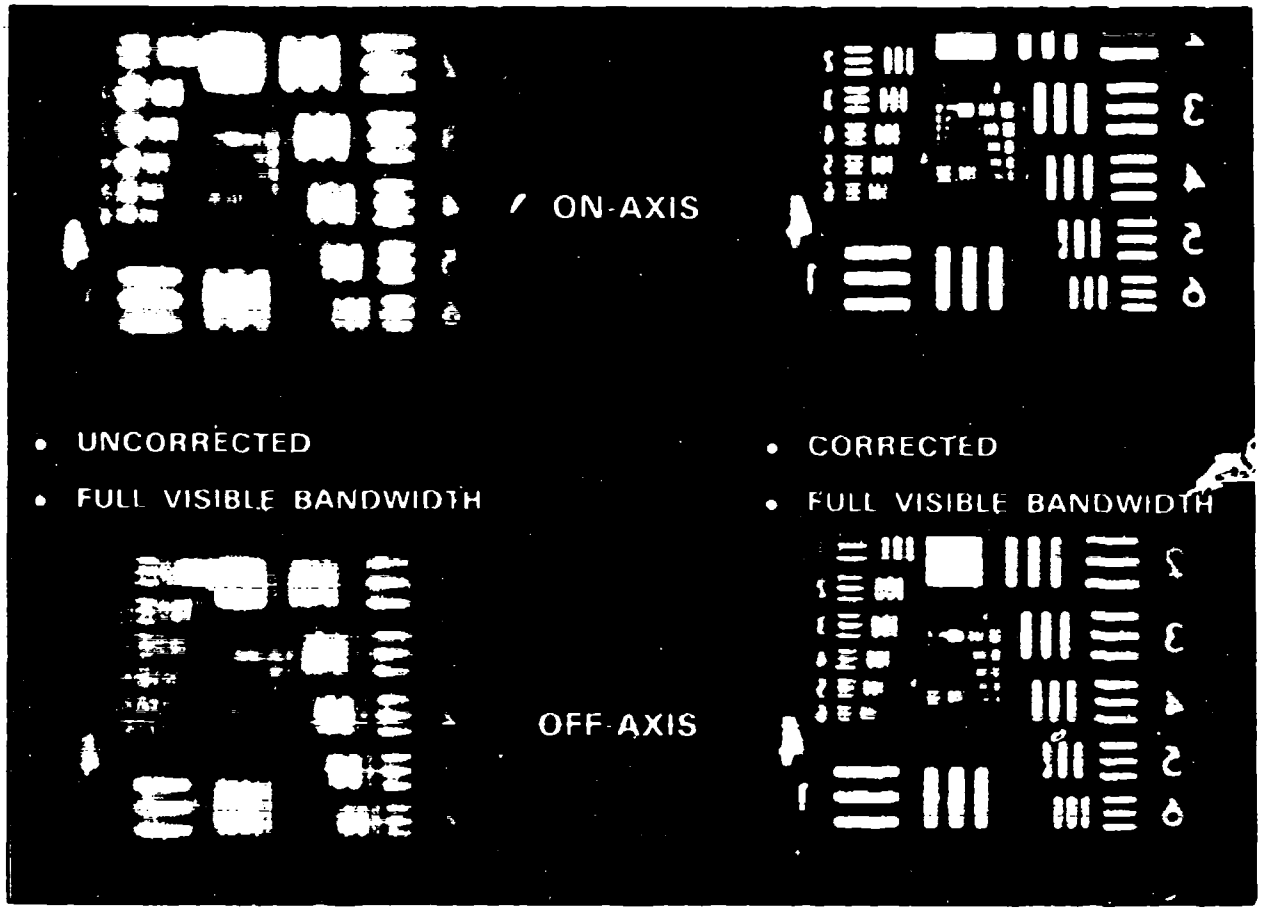


Figure 5-3. Experimental imaging results of the fused silica lens, with and without diffractive aberration correction.

Experimental verification of chromatic aberration correction using a refractive/diffractive element has been shown not only in the visible portion of the spectrum, but in the far-infrared (8 to 12 μm), the mid-infrared (3 to 5 μm), and the ultraviolet (0.246 to 0.248 μm) as well.

A refractive/diffractive achromat does not completely eliminate all of the chromatic aberration because the refractive index variation as a function of wavelength is not exactly described by Equation (5.8) which is a linear approximation to the true dispersive properties of refractive materials. In reality, the dispersion has a small nonlinear component that cannot be compensated for by a diffractive element. This nonlinear component is, in terms of lens design, called the secondary spectrum. Fortunately, the secondary spectrum is small in the majority of materials.

5.3 SPHERICAL ABERRATION CORRECTION

In the previous section we showed how a properly designed diffractive lens profile could be used to correct for the chromatic aberration of a refractive lens. Diffractive profiles can be used to correct for the monochromatic aberrations of refractive lenses as well. Here we will discuss the particular case of spherical aberration.

In the majority of cases, refractive lenses have spherical surfaces. A lens with spherical surfaces inherently suffers from spherical aberration. The spherical aberration of a single refractive lens element can be minimized by the proper choice of the radii of curvature of the two surfaces of the lens, but cannot be completely eliminated.

Two conventional solutions exist to eliminate spherical aberration. One is to use multiple lenses instead of a single lens. The number of lenses needed depends on the required performance of the lens system. This solution to the problem results in added weight, lower light throughput, and greater system complexity. The other conventional solution to the problem is to place an aspheric surface on the lens. This solution suffers from the fact that, in general, aspheric surfaces are very costly to produce.

The approach described here is to employ a diffractive surface to eliminate the spherical aberration of a refractive lens. For simplicity, the following analysis shows how a diffractive phase profile can eliminate third-order spherical aberration from a lens. A diffractive surface can correct for higher-order spherical aberration as well. Consider the wavefront, of wavelength λ_0, exiting from the back surface of a lens. Ideally, this wavefront would be a spherical wave converging to the focal point F and described by the phase profile

$$\phi_1 = -\frac{2\pi}{\lambda_0} R \tag{5.16}$$

where $R = (r^2 + F^2)^{\frac{1}{2}}$. A second-order approximation to this ideal wavefront can be made by expanding R in a power series, resulting in

$$\phi_{13} = -\frac{\pi r^2}{\lambda_0 F} + \frac{\pi r^4}{4\lambda_0 F^3}. \tag{5.17}$$

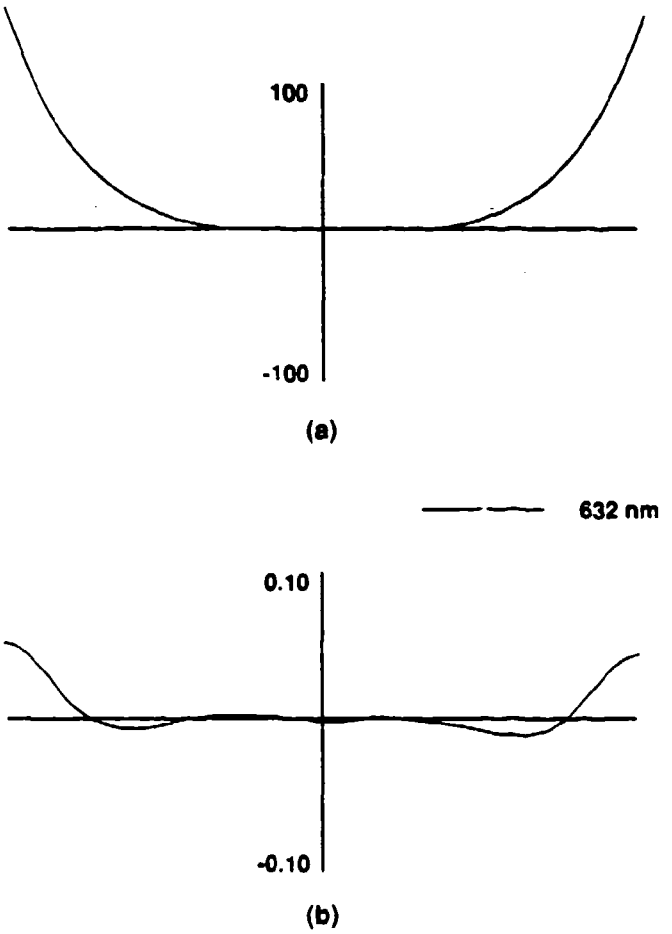


Figure 5-4. Theoretical phase error due to spherical aberration of a fused silica lens with and without diffractive correction.

This equation is the expression, to fourth order in r, of an ideal wavefront. A refractive lens with spherical surfaces cannot produce this wavefront. The wavefront from any particular lens, depending on the design, will vary. To simplify things, let us assume that the wavefront exiting the refractive lens is quadratic and given by

$$o_i = -\frac{\pi r^2}{\lambda_0 F}. \tag{5.18}$$

The third-order spherical aberration of this lens would then be

$$o_a = -\frac{\pi r^4}{4\lambda_0 F^3}. \tag{5.19}$$

This third-order spherical aberration can be negated by simply etching a diffractive phase profile into the back surface of the lens that has a phase profile of

$$o_d = \frac{\pi r^4}{4\lambda_0 F^3}. \tag{5.20}$$

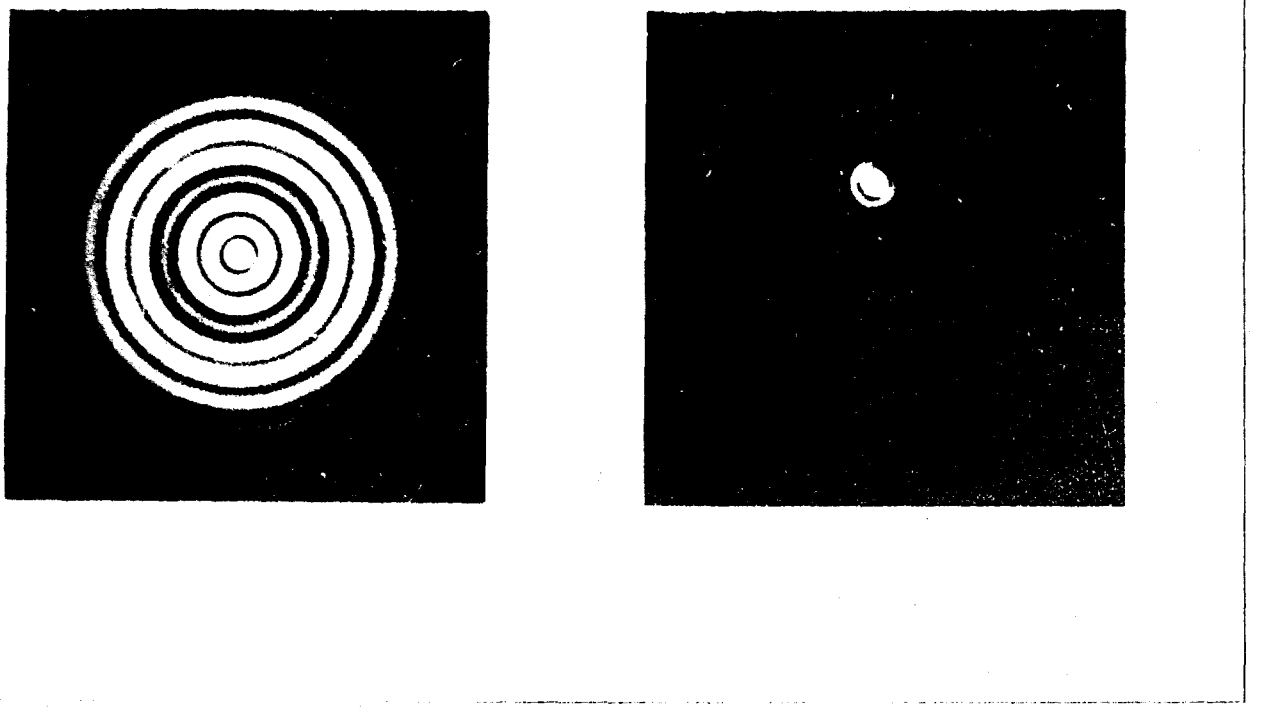
The resultant wavefront would be given by Equation (5.17) and suffer no third-order spherical aberration. This refractive/diffractive element, that has no spherical aberration at the wavelength λ_0, behaves very much like a conventional aspheric element.

A demonstration of the concept of spherical aberration correction has been performed using a fused silica single-element lens at the HeNe laser wavelength of 0.6328 μm. The fused silica lens was plano-convex and had a 1-in aperture and a 2-in focal length. This lens suffered from severe spherical aberration, having a maximum phase error of close to 100 waves, as shown in Figure 5-4(a). When placed on the lens, the properly designed diffractive surface had a theoretical phase error of less than 0.1 wave [see Figure 5-4(b)].

The diffractively corrected refractive lens was fabricated and tested. Figures 5-5(a) and (b) are images of the focal spot produced from the uncorrected and corrected lenses. Figure 5-5(a) clearly shows the expected light distribution associated with spherical aberration. The diffractively corrected focal point of Figure 5-5(b) is essentially diffraction limited. The experimentally verified improvement is enormous. It would take three conventional spherical lenses in tandem to achieve the same performance.

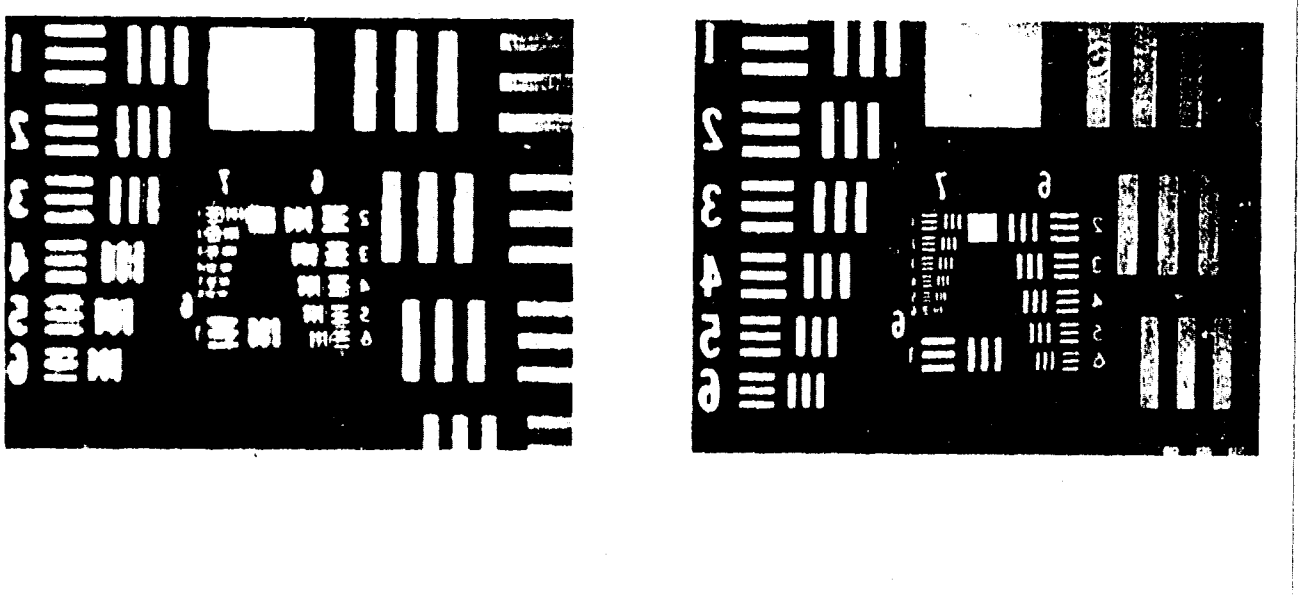
As further verification, the lenses were used to image an Air Force resolution test pattern. The results of the imaging experiments are shown in Figure 5-6. The uncorrected lens [Figure 5-6(a)] has a resolution as expected from theory. The diffractively corrected image [Figure 5-6(b)] has a resolution that is essentially diffraction limited.

The Binary Optics Group at Lincoln Laboratory has also demonstrated spherical aberration correction of lens elements at wavelengths in the far-infrared, the mid-infrared, and the ultraviolet regions of the spectrum.



125929 2

Experimentally measured focal points of the fused silica lens with and



125929 3

for the fused silica lens with and without diffractive spherical

5.4 LIMITATIONS OF REFRACTIVE/DIFFRACTIVE ELEMENTS

In Sections 5.2 and 5.3 we showed that it is possible to diffractively correct for the primary chromatic aberration and spherical aberration, at a specified wavelength, of a refractive lens. All single-element refractive lenses with spherical surfaces suffer from both chromatic and spherical aberration. The question arises as to how well a diffractive surface can correct for both chromatic and spherical aberrations over a finite wavelength band.

In order to get an estimate on the capability of a diffractive surface to correct for both chromatic and spherical aberration over a finite wavelength band, a model of the phase error of a refractive lens will be assumed to be

$$o_r(r) = \frac{2\pi}{\lambda}\left[A(\lambda_0 - \lambda)r^2 - B(\lambda_0 - \lambda)^2 r^2 + Cr^4\right]. \tag{5.21}$$

The first term on the right-hand side of this equation represents the primary chromatic aberration of the lens; the next term is a representation of the secondary spectrum; and the last term represents the spherical aberration of the lens. The values of the constants A, B, and C determine the amounts of primary chromatic aberration, secondary spectrum, and spherical aberration present.

A diffractive profile can be added to the refractive lens that imparts a phase given by

$$o_d(r) = \frac{2\pi}{\lambda}\left[A\lambda r^2 - C(\frac{\lambda}{\lambda_0})r^4\right]. \tag{5.22}$$

The resulting wavefront, from the refractive/diffractive lens, will have a residual phase error given by

$$o_e(r) = \frac{2\pi}{\lambda}\left[B(\lambda_0 - \lambda)^2 r^2 + C(1 - \frac{\lambda}{\lambda_0})r^4\right] \tag{5.23}$$

which is the phase error of Equation (5.21) minus the phase correction of Equation (5.22). The residual phase error of Equation (5.23) reveals, as expected, that the secondary spectrum of the refractive lens cannot be corrected. Furthermore, the additional term in Equation (5.23) represents the inability of a diffractive surface to completely correct for spherical aberration over a finite wavelength band. This residual term is commonly referred to as "spherochromatism," which is the amount of spherical aberration present in the image as a function of wavelength. For a center wavelength λ_0, the spherochromatism term in Equation (5.23) is zero, as expected. For wavelengths other than λ_0, the spherochromatism term is nonzero.

A diffractive surface is not able to completely correct for spherical aberration over a finite wavelength band. The amount of correction obtainable, as described in Equation (5.23), is proportional to the fractional bandwidth over which the lens has to operate. In many cases, the amount of correction is sufficient to justify the use of a diffractive surface.

As an example, consider an F/2 single-element silicon lens with a 100-mm focal length and an operating bandwidth from 3 to 5 μm. This bandwidth represents a 50-percent fractional bandwidth

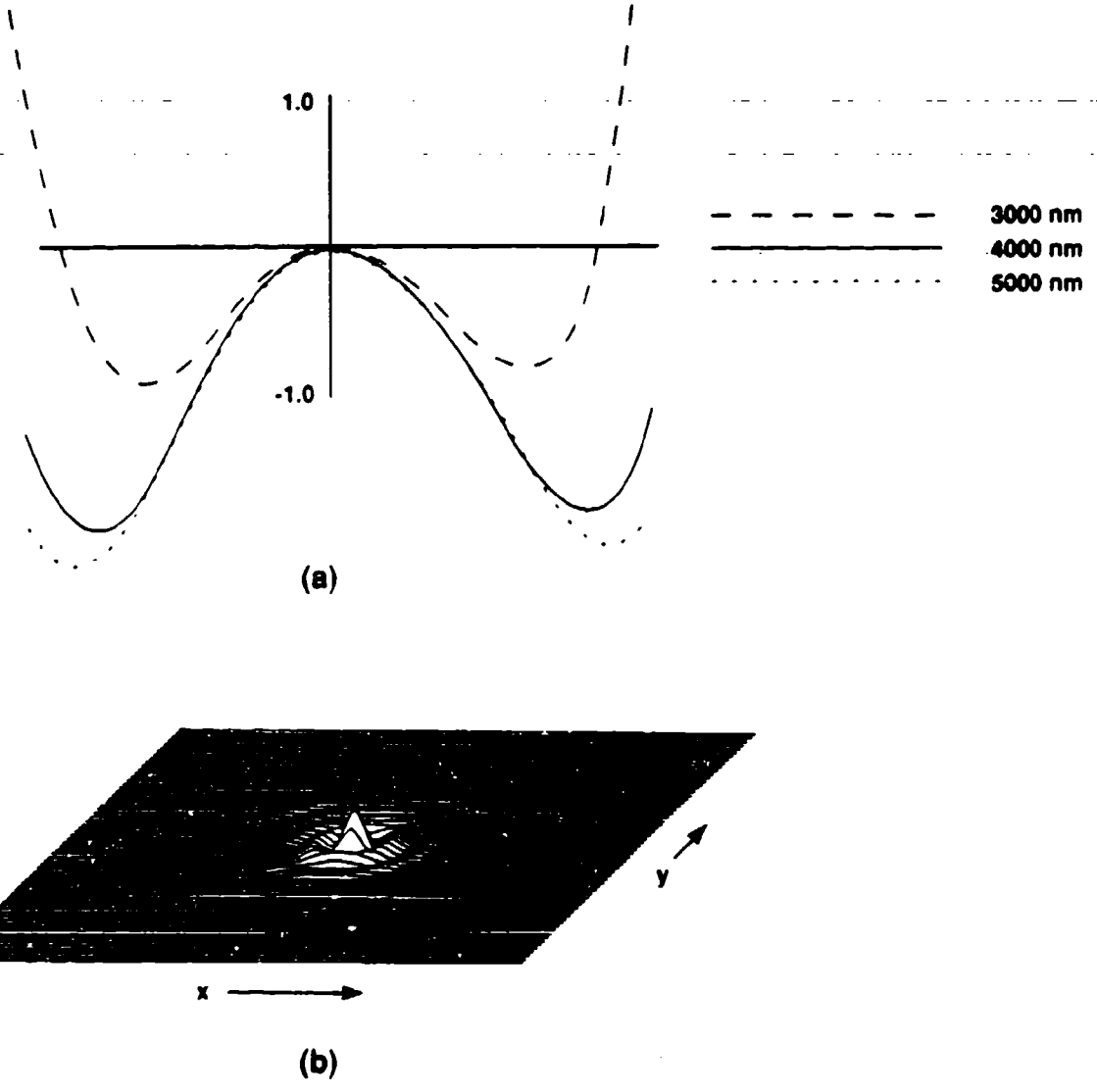


Figure 5-7. (a) The phase aberration and (b) point spread function of a refractive silicon lens.

which is representative of, or larger than, the majority of finite bandwidth systems. The phase aberration of the best-designed, spherical surface, refractive element is shown in Figure 5-7(a), and the light distribution at the focal point (i.e., point spread function) is shown in Figure 5-7(b). It is evident from Figure 5-7(a) that the single-element lens has both chromatic and spherical aberration. A diffractive phase profile, described by Equation (5.22) and placed on the back surface of the refractive silicon lens, results in the phase aberration shown in Figure 5-8(a) and the point spread function shown in Figure 5-8(b). The primary chromatic aberration of the refractive lens has been eliminated, as has the spherical aberration at the center wavelength (4 μm). The residual spherochromatism, that cannot be corrected, has a maximum phase error of 0.2 wave. This is a significant improvement over the maximum phase error of the refractive lens (3 waves).

The silicon lenses described above, with and without the diffractive phase profile, were fabricated and tested. The experimentally measured modulation transfer function (MTF) of both lenses is plotted in Figure 5-9. The resolving capability of the diffractively corrected lens is far superior to that of the completely refractive lens. The discrepancy between the theoretical and experimental performance of the diffractively corrected lens is attributable to the fact that the theoretical prediction assumed a 100-percent efficient diffractive surface for all wavelengths. The experimentally tested lens was an 8-phase level structure with a maximum efficiency, at 4 μm, of only 95 percent. In any case, the diffractively corrected lens far exceeded the completely refractive lens in performance.

The spherochromatism term in Equation (5.23) can be averaged over the operating fractional bandwidth $\Delta\lambda/\lambda_0$, resulting in an expression for the average residual spherochromatism

$$\bar{o}_\epsilon(r) = C\left(\frac{\Delta\lambda}{2\lambda_0}\right) r^4. \tag{5.24}$$

This equation reveals that the ratio of the residual spherochromatism of a diffractively corrected lens to the spherical aberration of the refractive lens is equal to one-half the fractional bandwidth.

Figure 5-10 illustrates the chromatic and spherical aberration reduction capability of a diffractive profile. Examples are shown for common operating wavelength regions extending from the far-infrared to the ultraviolet. Notice that, in all the examples, the residual rms phase error is less than 0.1 wave.

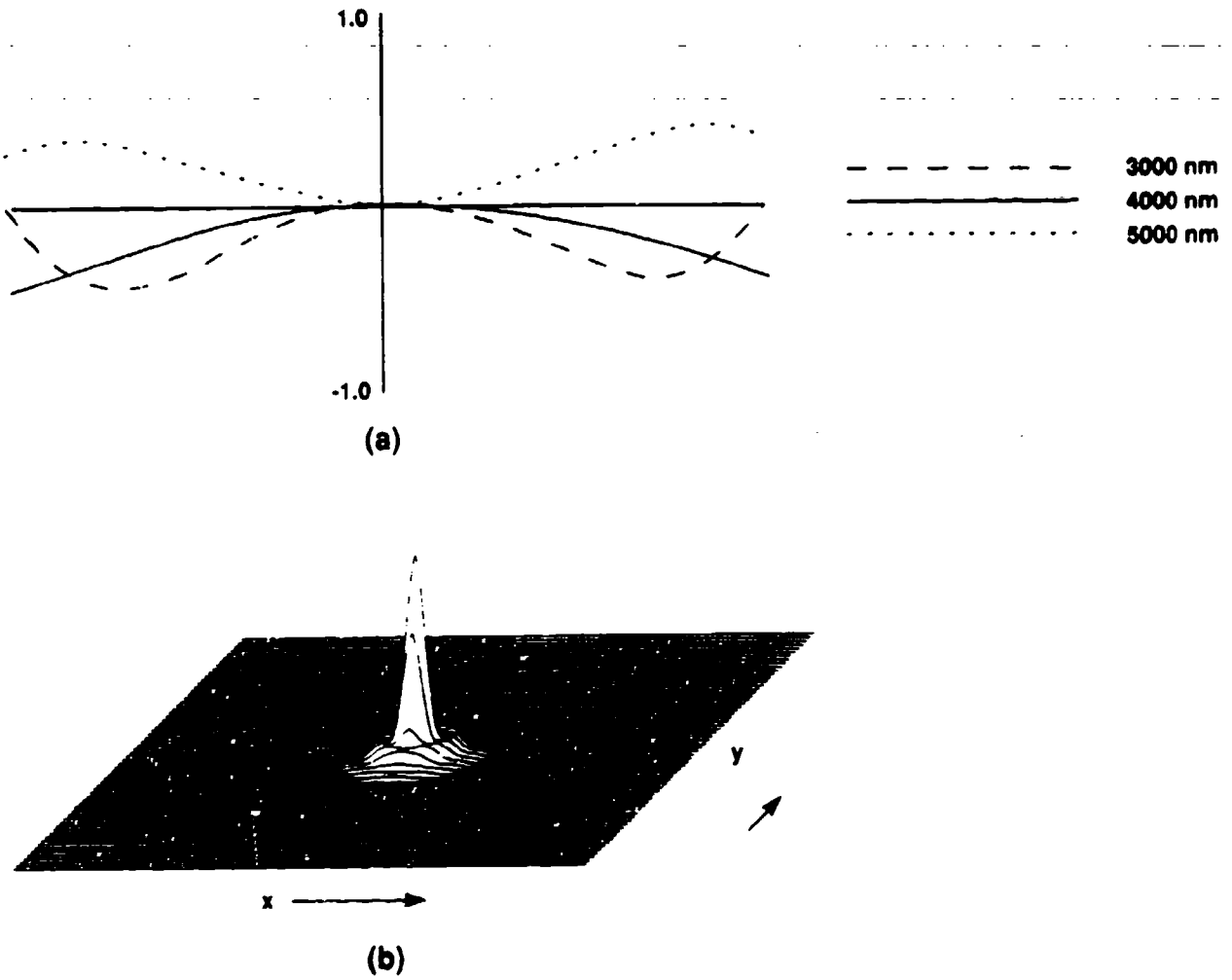
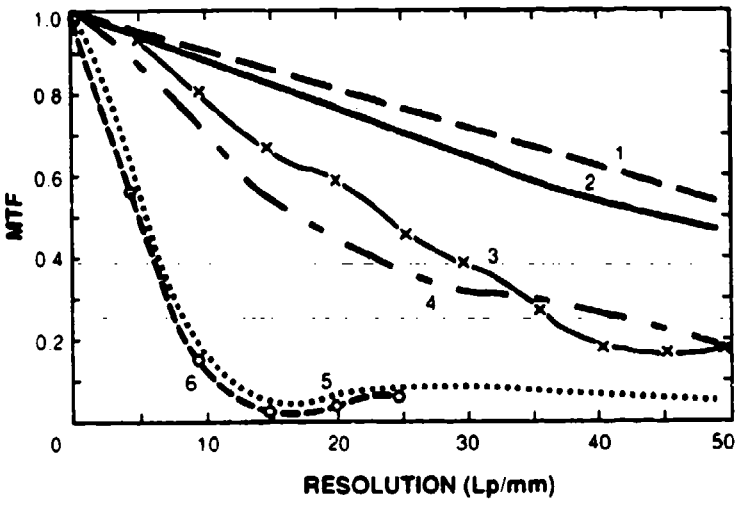


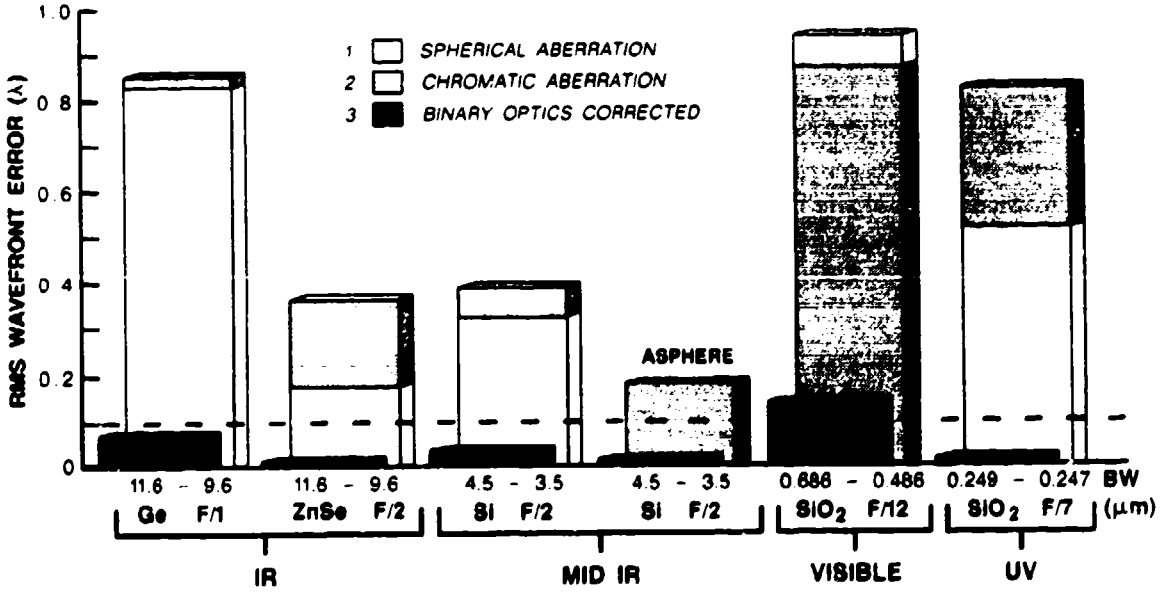
Figure 5-8. (a) The phase aberration and (b) point spread function of a diffractively corrected silicon lens.



124931-19

1. DIFFRACTION LIMITED OPERATION
2. PREDICTED BINARY OPTICS SINGLET
3. MEASURED BINARY OPTICS SINGLET
4. PREDICTED SPHERICAL TRIPLET
5. PREDICTED SPHERICAL SINGLET
6. MEASURED SPHERICAL SINGLET

Figure 5-9. MTF curves for the silicon lens, with and without diffractive correction.



125929-5

Figure 5-10. Examples of the chromatic and spherical aberration reduction possible by using a diffractive corrector.

6. DESIGNING DIFFRACTIVE PHASE PROFILES USING CODE V

In Section 5 we showed analytically the potential usefulness of a diffractive phase profile in reducing aberrations. The analysis was very nonspecific in regard to the exact diffractive phase function needed to optimally reduce the aberrations of a particular refractive lens. The exact determination of the optimum diffractive phase profile for any particular lens requires the assistance of a lens design program which must have the capability to insert diffractive phase profiles into a lens system and optimize the profile.

Inserting and optimizing a diffractive phase profile in a lens system can be accomplished using the commercially available lens design program CODE V; this program is used extensively by lens designers for optimizing and analyzing refractive and reflective systems. Lens designers familiar with the conventional capabilities of CODE V will have little problem learning and using the diffractive surface design capabilities of the program.

In CODE V, as well as other design programs, the lens designer inputs a design that meets the necessary first-order performance specifications of the system. An optimization routine is used that changes the initial first-order design in such a way to achieve maximum optical performance. In the optimization process, the thicknesses, spacings, and radii of curvature of the individual elements are treated as variables. The performance resulting from the optimization routine generally depends on the initial conditions specified by the designer.

CODE V has the ability to insert one or more diffractive surfaces anywhere into a lens system. These diffractive surfaces are specified by parameters that can be optimized to attain the best system performance. The implementation of diffractive surfaces in CODE V was formulated to emulate the recording of optically generated diffractive surfaces (i.e., holographic optical elements). The recording of a holographic surface is specified by the recording wavelength λ_0 and the location in space of two point sources, as shown in Figure 6-1. The two point sources, located at $R_1(x_1, y_1, z_1)$ and $R_2(x_2, y_2, z_2)$, produce spherical wavefronts. The interference of these two spherical wavefronts results in a diffractive phase profile at the recording plane given by

$$\phi_H(x,y) = \frac{2\pi}{\lambda_0}\left[\sqrt{(x-x_1)^2+(y-y_1)^2+z_1^2}+\sqrt{(x-x_2)^2+(y-y_2)^2+z_2^2}\right]. \qquad (6.1)$$

Note that the diffractive phase profiles that can be generated optically are a small subset of the total possible phase profiles. The spherical nature of the two interfering wavefronts restricts the set of optically generated phase profiles.

Fortunately, CODE V has the ability to analyze and optimize a more general set of diffractive phase profiles than that given by Equation (6.1). An additional diffractive phase term

$$\phi'(x,y) = \frac{2\pi}{\lambda_0}\sum_{k=0}^{10}\sum_{l=0}^{10-k} a_k b_l x^k y^l \qquad (6.2)$$

can be added, in CODE V, to the optically generated phase profile of Equation (6.1). This additional diffractive phase term makes it possible to optimize and analyze a much larger subset of

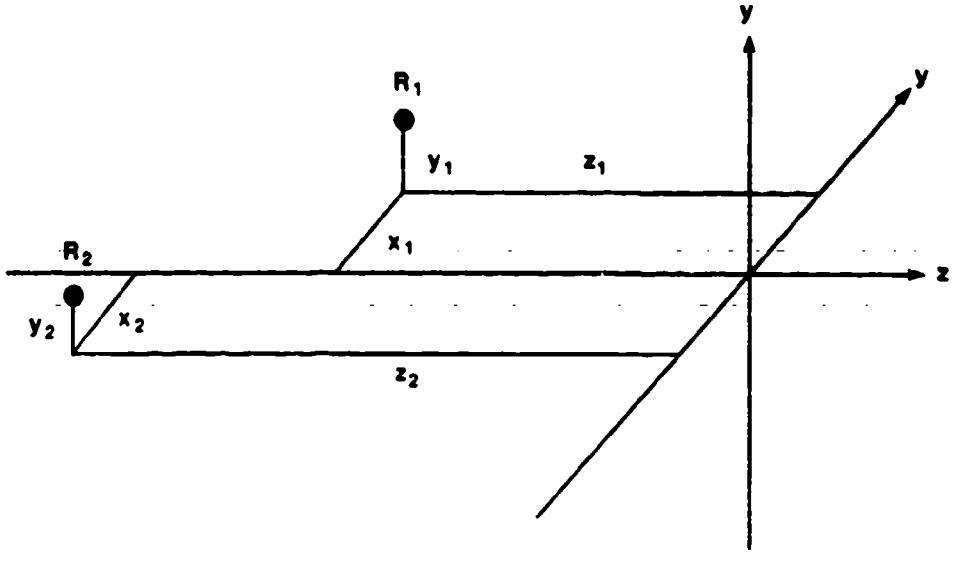


Figure 6-1. Recording setup for producing an optically generated holographic element.

diffractive phase profiles than those of Equation (6.1). Furthermore, it is possible (and preferable) to let Equation (6.2) completely specify the diffractive phase function. This can be accomplished by setting the two point source locations R_1 and R_2 at the same point in space. The resulting interference pattern from two point sources located at the same point is a constant. The resulting phase, given by Equation (6.1), becomes zero. The total diffractive phase is then given by Equation (6.2).

In the screen mode of CODE V, a diffractive surface can be placed in a lens system by entering the surface data screen (Gold S). Choosing the holographic surface option (Number 7) results in a screen that requires numerous input parameters. The first parameter to be entered is simply the surface number in the lens design on which the diffractive profile is to be placed. The second input parameter is the diffraction order of the diffractive phase profile that will be optimized and analyzed by CODE V. The first diffraction order is almost exclusively the order of interest.

The next parameter to be entered is the holographic recording wavelength. Typically, the wavelength of the laser used to optically record a holographic element would be entered. In our case, the value entered is irrelevant, since the diffractive phase profile is computer generated instead of optically generated. We prefer to set the wavelength equal to the center wavelength of the operating bandwidth only for consistency.

CODE V assumes that the diffraction order chosen will have 100-percent diffraction efficiency unless the next three input parameters are entered. These three parameters, used to model the diffraction efficiency of volume holograms, are: the volume thickness, the volume index of refraction,

and the index of refraction modulation. Since the surface relief profiles described in this report are not volume elements, it is best to leave these three parameters set to their default value of zero. The actual diffraction efficiency of an element will have to be determined outside of CODE V by using the theory developed in previous sections of this report.

The next set of eight input parameters specifies the two point source locations and whether the point sources are real or virtual. The spatial coordinates of both point sources are set to the same location, as mentioned above. It is then irrelevant whether the point sources are real or virtual. We set both point sources to be real for no particular reason other than consistency.

The last entry on the holographic surface screen is the number of aspheric diffractive phase terms, of Equation (6.2), to be entered. This entry is misleading in its wording. It is not the number of terms that should be entered, rather the number corresponding to the maximum term number in the polynomial expansion. The CODE V terminology for the aspheric phase polynomial is

$$o\prime(x, y) = \frac{2\pi}{\lambda_0} \sum_k \sum_l a_{kl} x^k y^l. \tag{6.3}$$

The number N, representing a particular term in the expansion, is determined by the expression

$$N = \frac{1}{2}[(k + l)^2 + 3l + k]. \tag{6.4}$$

The polynomial expansion of Equation (6.3) is truncated to values of $(k+l)$ less than or equal to 10. The total number of possible terms is 65. The term N = 65 for example, as given by Equation (6.4), represents the coefficient $a_{0,10}$ of the y^{10} term.

Once the maximum desired term number is entered on the screen, a final input screen consisting of a two-column table will appear. The term numbers N desired in the expansion are entered in the left-hand column; the values of the corresponding coefficients are entered in the right-hand column, directly opposite the appropriate term number. Any particular coefficient value entered can be set to a variable by pressing the Gold V key after entering the coefficient value. In many cases, there is little *a priori* knowledge as to what the coefficient values should be. For these cases, it is best to enter initial values of zero for all the desired coefficients and let the optimization routine determine their optimum values.

A complete description of entering a diffractive phase profile on a surface in a lens system has been given. The optimum diffractive phase profile is attained by using the CODE V automatic design feature. A deficiency of the CODE V program is that the diffractive aspheric phase terms are neglected in determining the first-order parameters of a lens system. These first-order parameters (i.e., effective focal length, F/#, etc.) are often used as constraints in the automatic design routine. If a diffractive element is optimized in CODE V using first-order constraints, the result can be erroneous. Only exact ray trace parameters can be used as constraints when optimizing a diffractive phase profile in CODE V.

The vast majority of optical systems are designed to operate over a field of view that is radially symmetric. If the elements in a lens system are constrained to be radially symmetric, it is only necessary to optimize the performance over a radial slice of the field of view (i.e., y-axis). The lens system is then guaranteed to have the same performance over any radial slice of the field of view. The advantages to optimizing over a radial slice as compared with the full field of view are speed and cost. Each additional field point used in the automatic design routine increases the computation time and, therefore, the expense.

The diffractive phase profile, described in Equation (6.3), is not radially symmetric. If this diffractive profile is to be optimized for use in an optical system that is to operate over a radially symmetric field of view, the field points used in the optimization routine would have to cover the whole field of view. If only field points lying on the y-axis were used in the optimization, the resulting profile would perform well for y-axis field points. Field points lying on the x-axis, or any radial axis other than the y-axis, would not be guaranteed suitable performance.

Within CODE V, a way exists to constrain the diffractive phase profile of Equation (6.3) to be radially symmetric. Constraining the diffractive profile to be radially symmetric allows for the optimization over the complete field of view, using only the y-axis field points. A radially symmetric diffractive phase profile can be expressed in (x,y) coordinates as

$$\phi_r'(x,y) = \frac{2\pi}{\lambda_0}[a_1(x^2+y^2)+a_2(x^4+2x^2y^2+y^4)+a_3(x^6+3x^4y^2+3x^2y^4+y^6)+\ldots]. \qquad (6.5)$$

By entering only the (x,y) terms of this equation in the diffractive phase expression [Equation (6.3)] and constraining the coefficient values to conform to the proportions of Equation (6.5), the diffractive phase can be made radially symmetric.

When optimizing a diffractive profile, the coefficients of Equation (6.3) can be constrained to conform to the ratios of Equation (6.5) by introducing a sequence file in the automatic design routine. This sequence file acts as a user-defined constraint in the optimization process. The introduction of the proper sequence file in the automatic design routine allows for the optimization over the total field of view from only field points lying on the y-axis.

The generation of sequence files is explained in the CODE V manual. A sequence file is basically a file type .SEQ in the VMS directory. An example of a user-defined constraint sequence file that forces the diffractive phase profile to be radially symmetric is given below. For this example, the sequence file is given the name HOE2.SEQ;1. It constrains the coefficients of Equation (6.5), on surface Number 2 of the lens system, to be radially symmetric.

Filename: HOE2.SEQ;1

```
@H21:=(HCO S2 C3)-(HCO S2 C5)
@H21=0
@H22:=(HCO S2 C10)-(HCO S2 C14)
@H22=0
@H23:=(HCO S2 C12)-2*(HCO S2 C10)
```

```
@H23=0
@H24:=(HCO S2 C21)-(HCO S2 C27)
@H24=0
@H25:=(HCO S2 C23)-(HCO S2 C25)
@H25=0
@H26:=(HCO S2 C23)-3*(HCO S2 C21)
@H26=0
```

The file HOE2.SEQ;1 will be read into CODE V as a user-defined constraint by entering IN HOE2 while in the command mode version of CODE V's automatic design.

The first line of this sequence file defines a variable, H21, that is the difference between the N = 3 and N = 5 terms of Equation (6.3). The second line of the file constrains H21 to be zero. In other words, the first two lines constrain the coefficients of the x^2 term and y^2 term to be equal. The diffractive phase profile will therefore be radially symmetric in the r^2 term. In a similar fashion, the next four lines constrain the profile to be radially symmetric in r^4, while the last six lines constrain the profile to be radially symmetric in r^6. This sequence file could easily be extended to constrain the profile to be radially symmetric up to the r^{10} term if desired.

The sequence file example given above constrains the diffractive profile on surface 2 of the lens system to be radially symmetric. Similiar sequence files can be generated and stored in the user's directory that constrain the diffractive phase profile to be radially symmetric on any surface of the lens system. More elaborate sequence files can also be generated that constrain the diffractive profile in any way desired by the designer.

7. SUMMARY

In the past, optical designers have avoided considering diffractive elements as practical alternatives to refractive and reflective elements. The neglect had been justified based on the fact that no reliable and cost-effective fabrication capability existed.

Hopefully, this report has provided the reader some insight into the potential usefulness of multi-level diffractive phase profiles. These profiles can be easily designed and evaluated by using standard lens design programs along with the procedures detailed in this report. The fabrication of these elements has been shown to be reliable and straightforward. The fabrication tools and equipment necessary to produce these elements are not inexpensive. However, it is standard equipment used in the fabrication of integrated circuits and available for use at many places.

Multi-level diffractive elements are in no way the solution to all optical design problems. However, there are many systems where a diffractive element can be used to gain an advantage over a conventional design. The applications section of this report (Section 5) attempted to elucidate some of the distinct capabilities, as well as the limitations, of diffractive elements.

It is our hope that an optical designer, after reading this report, will begin to seriously consider diffractive surfaces as potential solutions to some of his/her lens design problems. The use of these surfaces is in its infancy. The larger the number of designers considering these structures, the faster diffractive elements will begin to appear in real optical systems.