High-resolution Fresnel zone plates for x-ray applications by spatial-frequency multiplication

W. B. Yun
Department of Physics, State University of New York at Stony Brook, New York, New York 11794

M. R. Howells
Center for X-Ray Optics, Lawrence Berkeley Laboratory, Berkeley, California 94720

Received May 3, 1986; accepted August 26, 1986

We propose a scheme for deriving a high-resolution zone plate from a coarser one by interfering the light from two positive-order foci. We give a theoretical analysis of the process, which we call spatial-frequency multiplication. The analysis requires calculation of diffracted intensities far off axis, and we derive validity conditions for the application of approximate methods to this calculation. Specifically, we show that the Fresnel approximation is useful in many cases, whereas the method of stationary phase is valid for all cases of technological interest. Some practically useful parameter choices are considered for implementation of the zone-plate fabrication method. An example calculation suggests that a zone plate with smallest zone spacing of 13.8 nm may be possible.

1. INTRODUCTION

In recent years, the technology of microfabrication has advanced rapidly, and this has led to corresponding improvements in diffractive optical devices for soft x rays. There are several different components involved, but the chief one is the soft-x-ray lens or Fresnel zone plate. The spatial resolution of a zone plate is given by

$$\delta = 1.22 \Delta r_k/m,$$

where $\delta$ is the half-width of the Airy disk of the zone plate, $m$ is the diffracted order, and $\Delta r_k = r_k - r_{k-1}$, where $r_k$ is the radius of the outermost ($k$th) zone. According to the Rayleigh criterion, $\delta$ is the shortest distance between two features that can just be distinguished in the zone-plate image.

In view of Eq. (1) the size of the features that can be resolved by the zone plate in the first order is apparently set by the size limits of the microfabrication technology. At the present time, the best fabricators of zone plates achieve zone widths of 50–100 nm. These certainly allow for many interesting experiments, but it is agreed that the goal is zone widths of 10–20 nm. The only way to achieve resolution better than the zone width is to use a higher-order focus. This is generally not practical on efficiency and other grounds, but it provides a clue to a strategy for circumventing the first-order diffraction limit.

In this paper we describe a technique that to some extent circumvents this technological limit. We show that, given a zone plate (parent zone plate) with a suitable central stop, it is possible to derive another zone plate (daughter zone plate) with higher spatial frequency and larger number of rings. The daughter zone plate has a correspondingly higher resolution than the parent one. The method of derivation is a holographic one based on the formation of interference fringes between two diffracted orders of a parent zone plate.

The origins of this basic concept can be traced back almost a century, whereas the application of microfabrication is relatively recent. We analyze the possibilities of the application of the technique to zone-plate fabrication mathematically and discuss its practical implementation.

2. GEOMETRICAL-OPTICS ANALYSIS

The arrangement for making the daughter zone plate is shown in Fig. 1. The parent zone plate is on the left. It is assumed to be illuminated by a plane wave so light can be delivered to all orders for which foci exist. For illustration, Fig. 1 shows only the positive $m$th and $n$th orders. Consider two rays, both originating at the $k$th zone, one at the top at $A_k$, and other at the bottom at $B_k$. Suppose that the $A_k$ ray, for example, is directed toward the $m$th-order focus and the $B_k$ one toward the $n$th-order focus. The two rays intersect at a distance

$$z_{m,n} = \frac{2f}{f_m + f_n} = \frac{2f}{m + n}$$

from the zone plate, where $f = f_t$ and $f_m = f/m$. Since $z_{m,n}$ is independent of $k$, all the pairs of rays associated with the $m$th- and $n$th-order foci and originating from the same zone must intersect in the same plane. By a similar argument the two rays from the $k$th zone (whose radius is $r_k$) meet at a distance $R_k$ from the axis, where

$$R_k = r_k \frac{m - n}{m + n}.$$  

(As a convention in this paper, we always assume that $m$ is greater than $n$.) Thus we see that the pattern of ray intersections at the $z_{m,n}$ plane is a linear transform of the parent zone plate, and so we assume that it must also be a zone plate.

0740-3232/87/010034-07$02.00 © 1987 Optical Society of America
Fig. 1. Geometrical illustration of interference of the mth-order and nth-order diffracted beams. One may consider the points F_m and F_n to be coherent point sources for hologram formation. The recording region then contains the hologram (zone plate) whose focal properties are evident from the diagram.

Fig. 2. Pictorial illustration that two beam interferences can be obtained from a zone plate by use of central and outer stops. In the figure, an axial plane wave illumination and a/b = 1 are assumed. The notation n&m indicates part of an annular area with the nth-order and mth-order interference is recorded. For example, 1&3 indicates part of an annular area of first-order and third-order interference.

The idea of combining only two of the many diffracted beams associated with each focus can be realized by utilization of central and outer stops as shown in Fig. 1. This is made clear pictorially in Fig. 2. The use of such stops is not a disadvantage in practice. In fact, real zone plates for x-ray applications are usually designed with just this configuration to allow unwanted orders to be blocked out.3

The spatial frequencies \( f_A \) and \( f_B \) of the \( A_k \) and \( B_k \) rays are of the form \( \sin \theta/\lambda \), where \( \theta \) is the angle between the ray and the axis, so
where $\theta_{mk}$ and $\theta_{nk}$ are the corresponding angles of the $A_k$ and $B_k$ rays with the axis and $2\Delta r_k$ is the spatial wavelength at the $k$th zone of the parent zone plate. The spatial frequency of the interference fringes at the intersecting point of the two rays $A_k$ and $B_k$ is therefore given by

$$f_A - f_B = \frac{m + n}{2\Delta r_k} = (m + n)f_0,$$

where $f_0 = 1/(2\Delta r_k)$. Thus we have an $(m + n)$-fold enhancement of the spatial frequency of the parent zone plate and corresponding improvement of the numerical aperture.

It is not hard to prove by using Eqs. (3) and (5) that

$$k' = (m - n)k,$$

where $k'$ is the zone number of the daughter zone plate.

By using Eqs. (3) and (6) we can transform the zone-plate equation of the parent zone plate to get the equivalent relation for the derived daughter zone plate,

$$R_k^2 = k'f',$$

where

$$f' = \frac{m - n}{(m + n)^2}f.$$

Equation (8) shows that the daughter zone plate has a focal length $f'$ that is a function of $m$ and $n$ and is proportional to $f$. The physical justification for considering only rays originating from points that are opposite each other becomes clear with the insights of the method of stationary phase.

The interference intensity-distribution pattern and the focal length of the corresponding zone plate (daughter zone plate) can be also derived by treating the two involved foci as point sources: an object point and a reference point. The derived zone plate is then the hologram that results from the interference of the beams from the two sources. This can be visualized by examination of Fig. 1.

3. ANALYSIS UNDER FRESNEL APPROXIMATION

In the last section we obtained some important geometrical parameters, which will be confirmed analytically in this section. In the analysis, the Fresnel approximation is used and its range of validity is discussed. The analysis offers a clear physical justification for considering only rays originating from points that are opposite each other becomes clear with the insights of the method of stationary phase.

The function $\psi(x, y; g)$ is of importance in optical data processing since it is the impulse response of the shift-invariant linear system consisting of free-space propagation over a distance $1/g$, treated in the Fresnel approximation.

In Eq. (12), $F$ is the inverse of the primary focal length of the zone plate considered and $C_n$ is the expansion coefficient whose value is proportional to the diffraction amplitude of the $n$th-order beam when the zone plate is illuminated by a plane wave. It is important here to recognize that $\psi^*(x, y, nF) = \psi(x, y, -nF)$ represents the spherical wave associated with the $n$th-order focus under the Fresnel approximation.

Thus we see that at the parent zone-plate plane those $\psi$'s with positive $n$ in Eq. (12) represent the converging beams (positive order), while those with negative $n$ represent the diverging beams (negative order).

For simplicity and without loss of the essence of the analysis, we restrict our discussion here to an amplitude zone plate of which the transmittance is unity and zero for open and opaque zones, respectively. Furthermore, plane-wave illumination is assumed.

Using the notation in Fig. 3 and the Vander Lugt function, the amplitude $a(x', y'; z)$ at the recording plane can be expressed, still in the Fresnel approximation and neglecting the constant phase factor $e^{jhz}$, as

$$a(x', y'; z) = \frac{1}{f\lambda z} \int_{-\infty}^{\infty} t(x, y) \times \exp \left\{ \frac{jk}{2z} [(x - x')^2 + (y - y')^2] \right\} dx dy = \frac{1}{f\lambda z} t(r) \ast \psi(r, S),$$

where $\ast$ denotes the convolution operation, $r$ denotes a vector in the zone-plate plane, and

$$S = 1/z.$$  

Using the expansion of $t(r)$ in Eq. (12) and the convolution property of the Vander Lugt function, we get

$$a(x', z) = \frac{1}{f\lambda z} \sum_{n} C_n \psi(r; -nF) \ast \psi(r; S) = \sum_{n} C_n' \psi(r; nF S/y = nF - S),$$

where

$$C_n' = \frac{1}{f\lambda z} \frac{S}{S - nF} C_n.$$
It is legitimate to assume (see Fig. 1) that the central and outer stops of the zone plate are placed so that in certain annular regions at the recording plane (see Fig. 2) there are only two intersecting beams, which are evolved from the two beams $\psi(x, y; -nF)$ and $\psi(x, y; -mF)$ at the parent zone-plate plane. For the cases that $m$ and $n$ are positive, one of the two beams at the recording plane is diverging and the other is converging. For example, for the case that $m = 3, n = 1$, the beam $\psi[x', (SE/F - S)]$ is converging, while the beam $\psi[x', (3FS/3F - S)]$ is diverging because the recording plane is before the first-order focus but behind the third-order focus. The amplitude in the region of interference can be expressed as

$$a(r') = C_n\psi\left(r', \frac{nFS}{nF - S}\right) + C_m\psi\left(r', \frac{mFS}{mF - S}\right),$$

with intensity

$$I(r') = |a(r')|^2 = |C'_n|^2 + |C'_m|^2 + 2C'_nC'_m \cos \phi_{m,n},$$

where

$$\phi_{m,n} = \frac{\pi FSr'^2}{\lambda} \left(\frac{n}{nF - S} - \frac{m}{mF - S}\right).$$

We have used $C'_m = C_m$ and $C'_n = C_n$. This is true for amplitude zone plates but not true in general. From Eqs. (19) and (20) one can see that the intensity $I(r')$ is a periodic function of $r'^2$. This is the property of zone plates whose equation can be approximated as $r'^2 = k'f\lambda$, as we pointed out earlier. The intensity distribution $I(r')$ can be used to manufacture a zone plate that we called the daughter zone plate. In what follows, we deduce those parameters of the daughter zone plate that were obtained in Section 2.

Let

$$\phi_{m,n} = \pm k'\pi,$$

where $k'$ is an integer and the $\pm$ is not important so we choose it so that the $f'$ in Eq. (22) is positive.

Substituting Eq. (20) into Eq. (21) after some algebra leads to

$$R_{k}^2 = k'f'\lambda,$$

where we have replaced $r_k'$ by $R_k$ for the purpose of easy comparison with the results of Section 2 and

$$\frac{1}{f'} = FS\left(\frac{n}{nF - S} - \frac{m}{mF - S}\right) = FS^2\frac{n - m}{(nF - S)(mF - S)}.$$

Considering the interference of the positive $n$th- and $m$th-order beams and setting $S = 1/2n_m = (m + n)/(2f)$ and using $F = 1/f$, we get

$$F' = \frac{m - n}{(m + n)^2}f,$$

which is exactly Eq. (8). Equation (22) is also of the same form as Eq. (7), so that we have arrived at the same results as in Section 2.

The visibility of the high-spatial-frequency fringes is given by

$$V_{m,n} = \frac{|C'_nC'_m|}{|C'_n|^2 + |C'_m|^2}.$$

Consider, for example, an infinitely large amplitude zone plate, i.e., an infinite square wave in $r^2$ space. The coefficients $C_n$ are well known and are given by

$$C_n = \left\{ \begin{array}{ll} 1, & n = 0 \\ (-1)^{(n-1)/2}, & n = \text{odd} \\ 0, & \text{otherwise} \end{array} \right.$$

Setting $S = (m + n)/(2f)$ and substituting it and the $C_n$ from Eq. (25) into Eq. (17) to obtain $C'_m$ and $C'_n$ finally leads to

$$V_{m,n} = \frac{2mn}{m^2 + n^2}.$$

For the case of interference of first- and third-order beams, namely, $m = 3$ and $n = 1$, we get

$$V_{3,1} = 60\%,$$

which is quite good. It is useful to notice that the smaller the ratio of $m$ to $n$, the greater the visibility. If the ratio of the width of the transparent zone to the width of the opaque zone can be properly chosen, then the visibility can be made equal to unity. This is because by changing the ratio one can effectively change the relative amplitude of the two interfering beams.

In the deduction above, we have used the Fresnel approximation for propagation from the parent zone plate to the daughter zone plate.
recording plane (Fig. 3) and have neglected the pupil effect that is due to finite aperture of the parent zone plate. These approximations are of practical importance to the technique described here, so it is necessary to discuss them in more detail.

Referring to Fig. 3, we see that the optical path length from point Q to P is

$$\Delta = \left[ z^2 + |r - r'|^2 \right]^{1/2} = z + \frac{1}{2} |r - r'|^2 - \frac{1}{8z^2} |r - r'|^4 + \ldots$$

In the Fresnel approximation, only the first two terms are included. Let a and b be the radii of the annular aperture (Fig. 1), and as an example assume that a = b/2. Then one can see that the Fresnel approximation can be satisfied by setting

$$\frac{2\pi}{\lambda} \left( \frac{3b}{2} \right)^4 \ll 2\pi.$$  \hspace{1cm} (27a)

If we take π/2 as the limit, which is the Rayleigh criterion, and use $z = 2f/(m + n)$ and $b^2 = kbf\lambda$, the validity condition for the Fresnel approximation can be written as

$$k_b \leq \left( \frac{2}{3} \right)^2 \left( \frac{2}{m + n} \right)^{3/2} \left( \frac{2f}{\lambda} \right)^{1/2}.$$

For given $f$, $\lambda$, and $m + n$, the $k_b$ in expression (28) gives the maximum number of zones in the parent zone plate for which the Fresnel approximation is valid for the preceding analytical deduction. Failure of the Fresnel approximation is a rough indication that the daughter zone plate is aberrated, i.e., nonparabolic.

In order to get useful reduction of the zone spacing of the parent zone plate, one wants [see Eq. (6)] $m + n \geq 2$. This leads to $k_b < k_{\text{max}}$, where $k_{\text{max}}$ is defined in expression (10). If a parent zone plate is made so that the zone index of the outermost open zone is determined by the right-hand side of expression (29), then the smallest zone spacing of the parent zone plate would be

$$\Delta r_{k_b} = \frac{1}{2} \left( \frac{f}{k_b} \right)^{1/2} = \frac{3}{8} (m + n)^{3/4} \left( \frac{f}{\lambda} \right)^{1/4} \lambda.$$  \hspace{1cm} (29)

Thus the smallest zone spacing in the daughter zone plate from such a parent would be, according to Eq. (5),

$$\Delta r_{k_b} = \frac{\Delta r_{k_b}}{m + n}.$$  \hspace{1cm} (30)

Consider a case when $\lambda = 1.5 \text{ mm}$, $f = 2 \text{ mm}$, and $n = 1$, $m = 3$. We get $\Delta r_{k_b'} = 13.8 \text{ nm}$. This is an attractive number for zone-plate applications. The smallest spacing of the parent zone plate in this case would be $4 \times 13.8 \text{ nm} = 55.2 \text{ nm}$. At the present time, a zone plate with such a spacing indeed exists.\(^5\)

Now the pupil effect that is due to the stops of the parent zone plate should be considered. The central and outer stops are necessary to get just two beams to interfere. It can be shown that analytical results including the pupil effect can be obtained by replacing $\phi(r)$ (nFS/nF - S) by $\phi(r)$ (nFS/nF - S) + $H(f)$, where $H(f)$ is the Fourier transform of the function that characterizes the aperture of the parent zone plate defined by the central and outer stops, and the coordinates in frequency space are related by $f^2 = f_2^2 + f_3^2$. A good example is the function

$$h(r) = \begin{cases} 1 & \text{if } a \leq r \leq b \\ 0 & \text{otherwise} \end{cases},$$

and $H(f_3)$ is then given by\(^11\)

$$H(f_3) = 2\pi \left[ b^2 J_1(k_b a b) - a^2 J_1(k_b a o) \right].$$  \hspace{1cm} (31)

where $J_1(x)$ is the first-order Bessel function and $\omega = (f \lambda)^{-1}$.

Therefore one sees that the two interfering beams are not spherical waves but rather waves of convolution of the spherical wave with a function characterizing the aperture of a parent zone plate. Without the stops each spherical wave corresponding to each term in Eq. (12) is represented by a single sine wave in $r^2$ space and a single frequency in $1/r^2$ space. The frequencies are multiples of the fundamental frequency $(2\pi f)^{-1}$. The spectrum in $f^2$ space is thus a set of equally spaced $\delta$ functions. The effect of the stops is to convolve this spectrum with the spectrum $H(f_3)$ of the aperture function. This follows from the convolution theorem because $t(r)$ in direct $r^2$ space becomes $t(r)h(r)$ on introduction of the stops. This does not complicate matters much because the width of $H(f_3)$ is small compared with $(2\pi f)^{-1}$ so the convolution of the spherical wave with $H(f_3)$ simply has the effect of slightly broadening each $\delta$ function to the form given by Eq. (31). In other words, each frequency gets some sidebands. The interference fringes that we are interested in are formed by combining two waves with different center frequencies, but now in addition we expect to observe another set of fringes formed by the center frequency of one wave and a sideband of the other. The second set has a frequency close to the first, so we might expect to observe slowly varying beats between them. In fact various combinations are possible since there are two main peaks and several sidebands. Thus the intensity distribution in general consists of a set of high-spatial-frequency fringes, mainly arising from the central frequencies, plus a low-frequency background arising from the other interferences, which will generally be of low magnitude.

4. METHOD OF STATIONARY PHASE

In the preceding sections we have discussed the technique when the Fresnel approximation is valid for which the validity condition is given by expression (28). This gives good insight into the physics of the problem, and for many cases this is an adequate treatment. However, when the zone number of a parent zone plate is larger than the $k_b$ given in expression (28), the analysis under the Fresnel approximation begins to fail. One expects this to happen for cases when the daughter zone-plate pattern that we are calculating extends far off axis. It may happen even in cases when expression (10) is still satisfied and thus be of practical importance. In these cases the method of stationary phase can be applied and is a good basis for numerical calculation.\(^6\)

We are interested in the amplitude $a(r')$ at point $r'$ in the recording plane (see Fig. 3). From the Rayleigh–Sommerfeld expression of Huygens's principle\(^12\) we have

$$a(r') = \frac{a}{J\lambda} \int_0^\pi t(r)f(r', r')dr,$$

where $J\lambda$ is the wavelength.
where
\[ f(r, r') = \frac{a}{\sqrt{2\pi}} \int_0^{2\pi} \frac{\exp[ik\rho(\theta)]}{\rho(\theta)} \, d\theta, \]
\[ \rho(\theta) = |r|^2 = \left[z^2 + p^2 + r'^2 - \cos(\theta - \theta') \right]^{1/2}, \]
and \( \theta \) is the obliquity factor. This is exact.

The method of stationary phase is useful when the argument of the exponential in Eq. (33) oscillates rapidly over most of the range of the variable and gives canceling contributions to the integral. The regions over which the argument does not oscillate rapidly (near its stationary points) give the only important contributions to the integral.

The method of stationary phase gives an asymptotic approximation to integrals of the form
\[ I = \int g(Z) \exp[ih f(Z)] \, dz \]
for large \( k \) and is treated in many texts.\(^{11,13}\) The leading term of the series that approximates the integral is a sum of contributions from all the stationary points \( z_0 \) of \( f(Z) \) in the \( Z \) domain in question. For each stationary point the contribution is
\[ \left( \frac{\pi}{2|g'(Z_0)|} \right)^{1/2} g(Z_0) \exp\left( \frac{j\pi}{4} \right) \exp[ih f(Z_0)]. \]

For our case \( g(Z) = 1/\rho(\theta) \) and \( f(Z) = \rho(\theta) \) with stationary points \( f(Z) = 0 \) at \( \theta - \theta' = 0 \) and \( \theta - \theta' = \pi (\theta' \text{ being considered constant}); \) therefore, from expression (34) we get
\[ f(r, r') = 2 \left( \exp \left( \frac{j\pi}{4} \right) \left( \frac{\pi B}{k\sqrt{A + B}} \right)^{1/2} \exp\left(ikh/A + B \right) \right. \]
\[ \left. + j\left( \frac{\pi B}{k\sqrt{A - B}} \right)^{1/2} \exp\left(ikh/A - B \right) \right] \]
where \( A = z^2 + p^2 + r'^2 \) and \( B = -2rr' \).

The two terms in Eq. (35) correspond to the contributions from the two stationary points: one at \( \theta - \theta' = 0 \) and the other at \( \theta - \theta' = \pi \). Once \( f(r, r') \) is known, it is straightforward to evaluate the integral in Eq. (32) numerically, which otherwise would have been intractable as a double integral.

However, we are also interested in evaluating the size of the lowest neglected term in order to arrive at a validity condition for our case. Using a general method of Erdelyi\(^{14}\) and setting the ratio of the lowest neglected term to the leading term less than or equal to 1%, we arrive at
\[ k_{\text{min}} \simeq \frac{12}{(m + n)^3/2}, \]
as the validity condition for expression (34).

This expression gives a minimum value of the zone number of the parent zone plate for which the method of stationary phase can still be applied. For \( m + n \geq 2 \) we see that \( k_{\text{min}} \) is less than 2. In the derivation of Eq. (36), the obliquity factor was neglected, but it can be shown that Eq. (36) holds in general. One can see that the method of stationary phase is essentially always applicable to the situations that we are considering because we are interested in cases where \( m + n \geq 2 \) and because a parent zone plate with a central stop is always used. The method is even applicable to parent zone plates that cannot be approximated as parabolic zone plates, although it is questionable whether this has any practical interest.

It is useful to see that there are cases in which the calculation of a daughter zone plate pattern can be carried out either under the Fresnel approximation or by using the method of stationary phase with adequate accuracy since results obtained can be compared and checked.

The numerical computation in our earlier publication\(^{6}\) showed that in cases when Fresnel approximation and the method of stationary phase are valid, the results obtained with the two methods are in good agreement, while in large-angle cases, for which the Fresnel approximation is not valid, only the method of stationarity gives the correct result.

5. TECHNOLOGICAL REQUIREMENTS
The practical conditions required for the above scheme to work are listed below. For convenience, the case of first- and third-order interference is considered.

In view of the desired value of the minimum zone width it will be necessary to use a soft-x-ray source for the exposure. This source must have high coherent power to make the exposure time reasonable. We believe that even a storage-ring bending magnet is marginal and that a soft-x-ray undulator or a soft-x-ray laser may be needed. Specifically, if a zone plate like those currently in use in soft-x-ray imaging is used as a parent zone plate and if the soft-x-ray undulator to be installed next year at Brookhaven National Laboratory is used as the source, then the estimated exposure time is of the order of minutes.

- The illuminating beam must be sufficiently monochromatic for the third-order focus to be diffraction limited; i.e., \( \lambda/\Delta \lambda > 3k_{\text{max}}/0.61 \).
- The parent zone plate must be sufficiently well made that the diffraction-limited third-order spot size is achieved. The Rayleigh quarter-wave criterion would require placement of each ring with an error less than \( \lambda/60 \) of a zone width. Although this sounds challenging it is probably achievable by using holographic methods.
- A technology must exist for transforming the recording into a zone plate with suitable fidelity. This would in fact be possible by using state-of-the-art microfabrication technology. Structures with line spacing as small as 10 nm have been demonstrated.\(^{15}\)

6. SUMMARY
The interference pattern of two diffracted orders of a parent zone plate is another zone-plate pattern. By using the high-brilliance soft-x-ray sources being developed at present and with state-of-the-art microfabrication technology, the pattern may in principle be transformed into a practical zone plate. Spatial resolution of the daughter zone plate might be as small as 10 nm.

ACKNOWLEDGMENTS
We have benefited from discussions of this work with J. Kirz, D. Sayre, and B. J. Thompson. W. B. Yun wishes to acknowledge receipt of an IBM Fellowship.
REFERENCES


