Angular spectrum description of light propagation in planar diffractive optical elements

Rosemarie Hild
HTWK-Leipzig, Fachbereich Informatik, Mathematik und Naturwissenschaften
D-04251 Leipzig, PF 30 11 66
Tel.: 49 341 5804429
Fax:49 341 5808416
e-mail:hild@imn.htwk-leipzig.de

Abstract

The description of the light propagation in or between planar diffractive optical elements is done with the angular spectrum method. That means, the confinement to the paraxial approximation is removed. In addition it is not necessary to make a difference between Fraunhofer and Fresnel diffraction as usually done by the solution of the of the Rayleigh Sommerfeld formula.
The diffraction properties of slits of varying width are analysed first. The transitional phenomena between Fresnel and Fraunhofer diffraction will be discussed from the view point of metrology. The transition phenomena is characterised as a problem of reduced resolution or low pass filtering in free space propagation.
The light distributions produced by a plane screen , including slits, line and space structures or phase varying structures are investigated in detail.

Keywords: Optical Metrology in Production Engineering, Photon Management- diffractive optics, microscopy, nanotechnology

1. Introduction

The knowledge of the light intensity distributions that are created by planar diffractive optical elements is very important in modern technologies like microoptics, the optical lithography or in the combination of both techniques 1.
The theoretical problem consists in the description of the imaging properties of small optical elements or small diffraction apertures up to the region of the wave length. In general the light intensities created in free space or optical media by such elements, are calculated with the solution of the diffraction integral. The limitations arise from the solution of the Rayleigh Sommerfeld formula or the solution of the Kirchhoff diffraction formula in the paraxial case 2, 3.
The parabolic or Fresnel approximation is suitable to analyse the proximity effects of very simple structures in the so called near field region. In the far field region the diffraction images correspond to the Fraunhofer approximation or the spatial Fourier transform of an input object. The field distribution is given by a superposition of plane waves. Far field approximation means one has to consider large distances between the optical element and the plane of observation.
Unfortunately the structures in diffractive optics are made up of different singular structures or complicated phase varying elements. Often their illumination is characterised by oblique incidence.
Ray tracing methods used in diffractive optics describe the influence of oblique illumination but not the proximity effects sufficiently 4.
In the first topic the theoretical fundamentals are explained. We start with the solution of Maxwell equations and the introduction of the angular spectrum.
From the analysis of the response respectively the transfer function of free space, the well known cases of Fresnel and Fraunhofer diffraction can be derived. The Fresnel number is introduced as a tool to decide above Fresnel or Fraunhofer approximation.
The theoretical results of the first topic are discussed for special diffraction apertures.

Considering a the slit of the width "a" at a constant distance "z" illuminated by a plane wave of the wave length λ, the reduction of the Fresnel number is realised by the reduction the slit width. The transition between near and far field region can be interpreted as the "loss" of resolution.
The light distributions in the half space z>0 caused by single slits and extended diffraction gratings combined with optical components like microlenses are calculated.

2. The description of light propagation in free space

The description of light propagation planar diffractive optical elements, through masks in lithographic processes or in the combination of both e.g.1 is connected with some problems caused by the solution of the Maxwell equations itself. In optics the Maxwell equations (1) are considered in non dispersive and non magnetic materials, in homogenous optical media without charges and currents. In this case the Maxwell equations are given by

$\vec{D} = \varepsilon_0 \varepsilon \vec{E}$ and $\vec{B} = \mu_0 \vec{H}$.

$$\mathrm{div}\vec{E} = 0 \qquad \mathrm{div}\vec{B} = 0 \tag{1}$$

$$\mathrm{rot}\vec{E}(\vec{r}) = -\mu_0 \frac{\partial}{\partial t}\vec{H}(\vec{r}) \qquad \mathrm{rot}\vec{H}(\vec{r}) = \varepsilon_0 \varepsilon \frac{\partial}{\partial t}\vec{E}(\vec{r})$$

From Maxwell-equations one gets the wave equation for the electric and the magnetic component the field separated. That's why it is sufficient to consider in the following the electric component only.

$$\Delta\vec{E}(\vec{r},t) - \frac{n^2}{c_0^2}\frac{\partial^2}{\partial t^2}\vec{E}(\vec{r},t) = 0 . \tag{2}$$

The term $c_0 = \sqrt{\frac{1}{\varepsilon_0 \mu_0}}$ is equal to the velocity of light in vacuum and $n = \sqrt{\varepsilon}$ corresponds to the refractive index of the medium. Considering the illumination with monocromatic light of frequency ν, the wave function is given by

$$\vec{E}(\vec{r},t) = \vec{E}(\vec{r}) \cdot \exp\{-i2\pi\nu t\}$$

The spatial component of the wave function is determined as the solution of the Helmholtz equation.

$$\Delta\vec{E}(\vec{r}) + k^2\vec{E}(\vec{r}) = 0 \qquad \text{with } k = k_0 n^2 \text{ and } k_0 = \frac{2\pi}{\lambda} \;; \quad \lambda\text{-vacuum wave length.} \tag{3}$$

Now we consider the propagation of the electric field between two planes. The sources of the field are situated in the half space $z \le 0$. At $z = 0$ we have the field distribution $\vec{E}_0(\vec{r})$. The field $\vec{E}(\vec{r})$ in the half space $z \ge 0$ with refractive index n has to accomplish the Maxwell equation $\mathrm{div}\vec{E}(\vec{r}) = 0$; the so called condition of transversality.

The field in the half space $z \ge 0$ can be developed with the so called angular spectrum method e.g. 5. Considering the electric field, composed of an amount of partial waves with the wave vector $\vec{m}$ and the amplitude $\tilde{\vec{E}}_0(\xi,\eta)$:

$$\vec{E}(\vec{r})\Big|_{z\ge 0} = \iint_\infty \tilde{\vec{E}}_0(\xi,\eta)\exp\{i2\pi\vec{m}\vec{r}\}d\xi d\eta \quad \text{with } \vec{m} = (\xi,\eta,m) \text{ and } m = \sqrt{\frac{1}{\lambda^2} - \xi^2 - \eta^2} . \tag{4}$$

The m-th. component of the wave vector $\vec{m}$ is given from the solution of the Helmholtz equation. Regarding the condition of transversality one gets:.

$$\tilde{\vec{E}}(\xi,\eta) = \begin{pmatrix} \tilde{\vec{E}}_{0x}(\xi,\eta) \\ \tilde{\vec{E}}_{0y}(\xi,\eta) \\ \frac{1}{m}\left(\xi\tilde{\vec{E}}_{0x}(\xi,\eta) + \eta\tilde{\vec{E}}_{0y}(\xi,\eta)\right) \end{pmatrix} \tag{5}$$

The Fourier-transform $\tilde{\vec{E}}_{0x}(\xi,\eta),\tilde{\vec{E}}_{0y}(\xi,\eta)$ are called mode spectra.

$$\tilde{\vec{E}}_{0x,y} = \frac{n^2}{\lambda^2}\iint_{\infty} \vec{E}_{0x,y}(x,y)\exp\{-i2\pi(x\xi + y\eta)\}dxdy\,. \tag{6}$$

From eq. (4) one sees real partial waves are possible if m is real. That means details $\left(\xi = \frac{1}{\Delta x}, \eta = \frac{1}{\Delta y}\right)$ in the field $\vec{E}_0$ up to the dimension of the wave length are transmitted. That is why the free space acts like a low pass filter.

One solution of eq. (4) is easy calculated with the method of stationary phases possible in the case of $kr >> 1$. $\xi = \frac{x}{\lambda r}$, $\eta = \frac{y}{\lambda r}$, $m = \frac{z}{\lambda r}$, $|\vec{r}| = r$

$$E_{x,y}(\vec{r})\Big|_{z\geq 0} \cong \frac{z}{\lambda r}\tilde{E}_{0x,y}\left(\frac{x}{\lambda r},\frac{y}{\lambda r}\right)\frac{\exp\{ikr\}}{r}$$

$$E_z(\vec{r})\Big|_{z\geq 0} \cong \left[\frac{x}{\lambda r}\tilde{E}_{0x}\left(\frac{x}{\lambda r},\frac{y}{\lambda r}\right) + \frac{y}{\lambda r}\tilde{E}_{0y}\left(\frac{x}{\lambda r},\frac{y}{\lambda r}\right)\right]\frac{\exp\{ikr\}}{r}\,. \tag{7}$$

In the far field approximation $\frac{x}{r},\frac{y}{r} \to 0$ und $\frac{z}{r} \to 1$ the z-component of the field vanishes. That means the field components E_x and E_y are independent and the fundamentals of scalar diffraction theory are accomplished. In the following we want to consider the scalar wave field that corresponds to one independent component of the wave field only, or one has to realise the far field conditions 5, 6.

2.1. The scalar diffraction theory

$V(\vec{r})$ is the component of the wave field. The solution of the Helmholtz-equation (3) $V_1(x_1,y_1,z_1)$ can be described as convolution between the impulse response $h(x_0 - x_1, y_0 - y_1, z_0 - z_1)$ and the function $V_0(x_0,y_0,z_0)$.

$$V_1(x_1,y_1,z_1) = \int V_0(x_0,y_0,z_0)h(x_0 - x_1, y_0 - y_1, z_0 - z_1)dx_0dy_0\,. \tag{8}$$

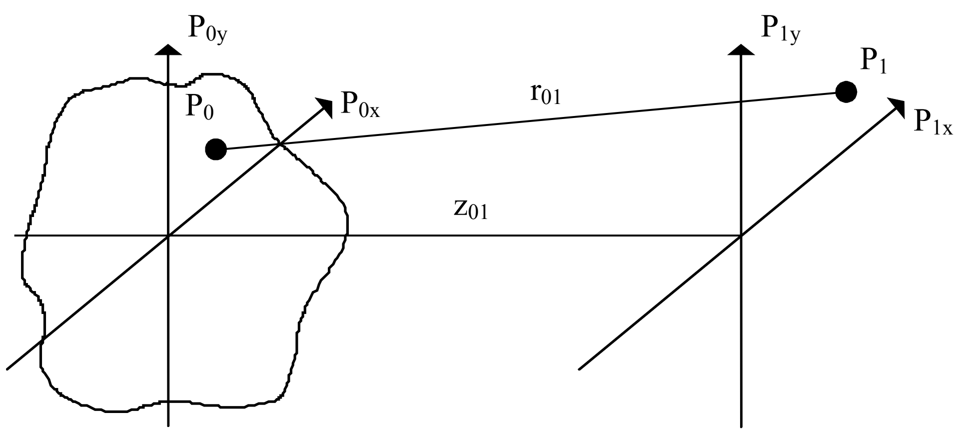

Fig. 1. Propagation of light between two planes

The Fourier-Transform of the impulse response h is called the transfer function H of free space (n=1). From comparison between eq. (4) and eq. (6) one gets

$$H(\xi,\eta,z)=\frac{\tilde{V}_1(\xi,\eta,z)}{\tilde{V}_0(\xi,\eta,z)}=\exp\left\{i2\pi z\sqrt{\frac{1}{\lambda^2}-(\xi^2+\eta^2)}\right\}\ . \qquad (9)$$

The paraxial approximation means $\xi^2+\eta^2 << \frac{1}{\lambda^2}$, or that the square root in the exponent can be developed. The condition $\theta^2=\lambda^2\xi^2+\lambda^2\eta^2<<1$ is realised.

In this case the function is given by an exponential row $\sqrt{1-\theta^2}=1-\frac{\theta^2}{2}+\frac{\theta^4}{8}-.......$.

If the third term is small, e.g. it is possible to break up the development

$$\frac{2\pi}{8\lambda}z\theta^4<<\pi\ , \qquad (10)$$

The transfer function for free space in paraxial approximation is given by

$$H(\xi,\eta,z)=\exp\{ikz\}\exp\{-i\pi z\lambda(\xi^2+\eta^2)\}. \qquad (11)$$

Considering eq. (10) one can introduce the so called Fresnel number N_F 2, 3, a relation between the maximal extension a^2 of the opening and the distance between the opening and the considered plane at the distance z.

$$\frac{1}{4}\frac{a^2}{\lambda z}\frac{a^2}{z^2}\ = N_F\ \frac{a^2}{4z^2}\ <<1 \quad \text{with} \qquad N_F=\frac{a^2}{\lambda z} \qquad (12)$$

The introduced approximation is called Fresnel-approximation if the Fresnel-number N_F is greater than 1. The resulting wave field is an interference of parabolic waves.

The impulse responce h is given by the Fourier-Transform of the transfer function H eq. (11)

$$h(x_0-x_1,y_0-y_1)=\frac{\exp\{ikz_{01}\}}{i\lambda z_{01}}\exp\left\{\frac{ik}{2z_{01}}\left((x_0-x_1)^2+(y_0-y_1)^2\right)\right\}. \qquad (13)$$

An interference of plane waves can be described if the Fresnel number is smaller than 1, that means in large distances z or the so called far field approximation. In this case the impulse response is realized as

$$h(x_0-x_1,y_0-y_1)=\frac{\exp\{ikz_{01}\}}{i\lambda z_{01}}\exp\left\{-\frac{ik}{2z_{01}}(x_0x_1+y_0y_1)\right\}\ . \qquad (14)$$

The wave field V_1 describes the far field diffraction image or Fraunhofer diffraction image that is equal to the Fourier-transform of the input opening.

Since it is not possible to measure the field amplitudes, it is sufficient to calculate the image intensity. For the scalar fields, the intensity is given by

$$I(x_1,y_1)=V(x_1,y_1)\cdot V^*(x_1,y_1) \qquad (15)$$

3. Light distribution caused by a slit aperture

First let us discus the light intensity distribution, that is produced behind a plane screen with a slit opening of varying width, considered in wave length units . A plane wave illumination is supposed.
The intensity distribution is calculated with angular spectrum following eq.'s (9; 4; 15).
First we consider the intensity distribution of a slit of a width of 10 in Fig. 2. We can see the intensity in the near field or in the Fresnel region. The edge position at 5 is connected with the intensity level of 0.25 as expected for normalised intensity with coherent illumination.
The Fresnel number N_F eq. (12) depends on the width of the slit and the distance z. One can obtain a reduction of the Fresnel number by increasing the distance z between the screen and the image plane or by an reduction of the width of the slit. The reduction of slit width should be our first point of view.
We want to combine the width of 10 with a Fresnel number equal to 100. Analysing the intensities, that are calculated for smaller openings of 4 and 6 one can see the same effect; the changing intensity structure inside of the opening. The slit width is found as described before. It corresponds to an intensity level of 0.25. If the opening has a width of 2 only, the structure inside of the slit is lost. One gets one intensity peak without structures. The peak minimum corresponds to the limits of the opening. The same behaviour is obtained for the opening of 1 as well, following our notation this corresponds to a Fresnel number of 1. We reach the limit from condition (4), that means the diffraction is taken place in openings in the region of the wave length. We can not find any similarity between the input structure and the intensity. The information of the edges corresponding to high spatial frequencies is lost.
Fig. 3 shows the intensity distributions of an openings smaller than 1 , or Fresnel numbers smaller than 1. The information of the opening seams to be lost . The position of the intensity minimum does not have any relation to the opening width. What happens?
We are leaving the range of near field and move to the range of far field region. In the far field or Fraunhofer region, the image intensity corresponds to the Fraunhofer diffraction image. The case is reached for an opening of 0.5 , or a Fresnel number of 0.25. The position of the minimum is found by:

$$x = \frac{\lambda z}{a} \tag{17}$$

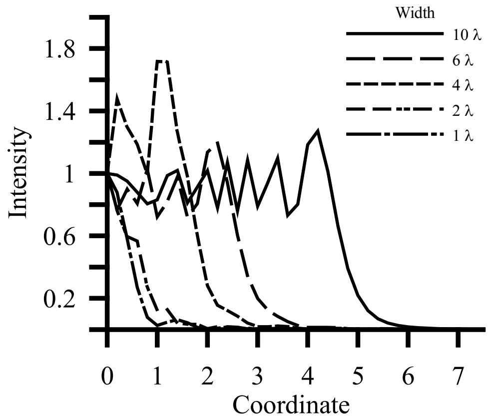

Fig. 2. Normalised intensity distributions behind slit openings in the Fresnel region

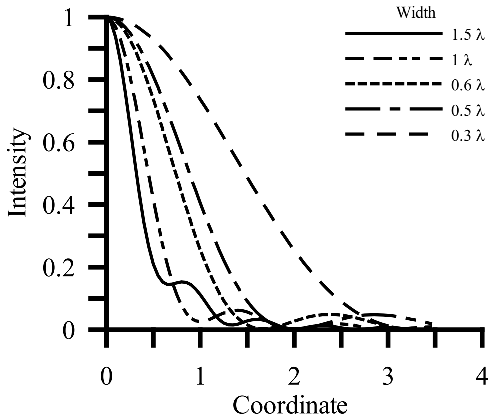

Fig. 3. Normalised intensity distributions behind slit openings in the Fraunhofer region

In fig.3 one can find the position of the minimum for a = 0.5 ; z = 1 at x = 2 and at last for a = 0.3 at x = 3.33 .
Fig. 4 shows the intensity distribution behind a slit of the width of 5 in dependence on the distance z from the opening in wave length units. From table 1 one can see the classification of the Fresnel and Fraunhofer regions along the z-axis. The position of the first minimum in the Fraunhofer regions follows eq. (17). Up to the Fresnel number 1 at z = 25 the slit is very close. Leaving the Fresnel region the light beam goes wider and wider. In Fig. 4b one can realise the Fraunhofer diffraction image up to infinity. The position of the first minimum is given by the relation (17) at x = 6.4 , x = 14.2 and at x = 20 in Fig. 4c in dependence on slit width and distance.

Table 1. Slit width and Fresnel number in dependence on the distance

a	$N_F = 1$	$N_F = 0.5$	$N_F = 0.25$
5λ	z = 25λ	z = 50λ	z = 100λ
10λ	z = 100λ	z = 200λ	z = 400λ
20λ	z = 400λ	z = 800λ	z = 1600λ

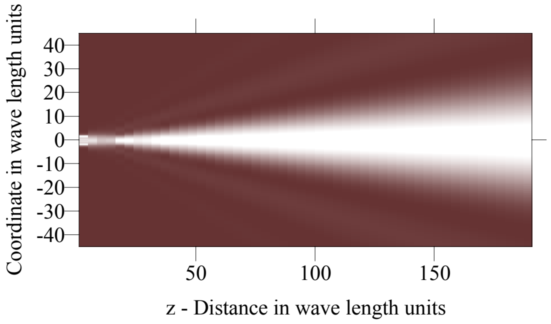

Fig. 4 a. Light intensity distribution created by a 5 wide slit opening

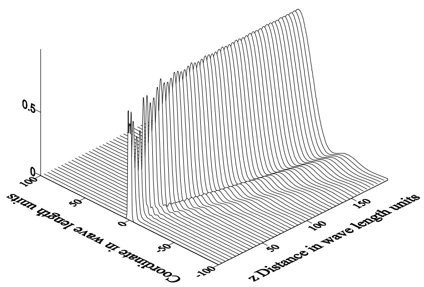

Fig. 4b. Light intensity distribution created by a 5 wide slit opening

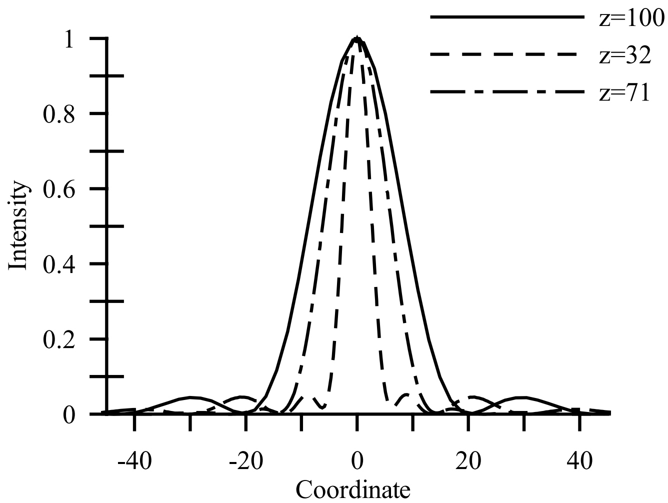

Fig. 4c. Axial cut through the light intensity distribution created by a 5 wide slit opening

4. Line and space structures

The diffraction behaviour of single screens, slits, squares, circular openings, holes etc. has been discussed widely. Next we want to consider the intensity distributions that are created behind line and space structures. Line and space structures of 3, 5, and 11 lines are analysed. The imaging properties of such periodical repeating objects tend to behave like optical gratings the greater the number of lines is. The structures used in the Fig. 5 resp. 6 have slit widths of 5λ and grating periods of p = 10 λ.

Fig.'s 5a-c show the expected behaviour for the light intensity distributions from a periodical repeating objects. Such structures give reasons to the so called self imaging effect, or Talbot effect 6. This self image is realized at $z = n\frac{2p^2}{\lambda}$ with a period of p = 10 λ at z = 200 λ and 400 . A doubled frequency one can observe at z = 50 , 150 , 250 and 350 . The contrast reversal is obtained at z = 100 and 300 .

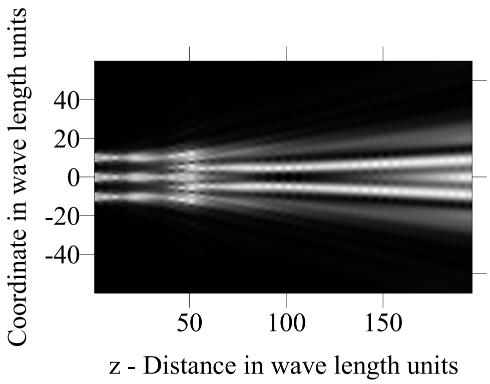

Fig. 5 a.. Diffraction image of a 3-line object; self image at z = 200λ

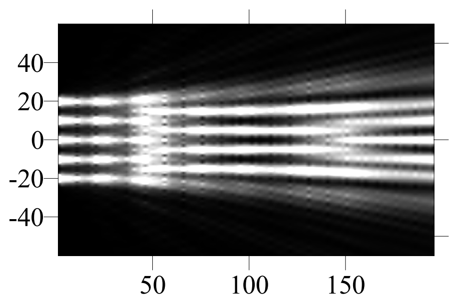

Fig. 5 b. Diffraction image of a 5-line object; self image at z = 200λ

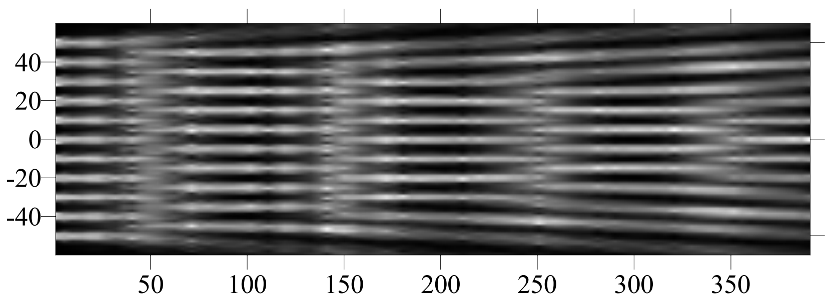

Fig. 5c. Diffraction image of a 11-slit object; Self image at z = 200λ and z = 400λ

In Fig. 6 one can see axial cuts at different planes. Fig. 6 a shows an axial cut at z = 200λ, so one can realise that the Fourier transform or the Fraunhofer diffraction image of the 3 line object is given. The grey line corresponds to the intensity of one single slit of the width 5 . We can repeat our conclusion for the 5 slit object in Fig. 6b. Here we can see the Fraunhofer image for an 5 slit object compared with that for the single slit.

The Fraunhofer diffraction image, or the far field diffraction for the 11 line is drawn in Fig. 6c for z = 400λ. Because of the larger extension of this object, the far field region is reached just in a larger distance. In the Fig.'s the diffraction image of a single line is shown in a grey line. Especially in Fig. 6c the grey line is superposed with a lot of computer noise.
The images underline two aspects. First one gets the expected diffraction images and second "self imaging" means, that one can see a diffraction image and not the image of the grating.

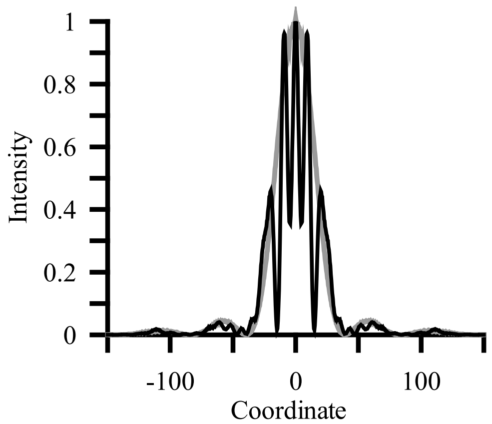

Fig. 6a. Axial cut for the 3 slit object at a distance z = 200λ

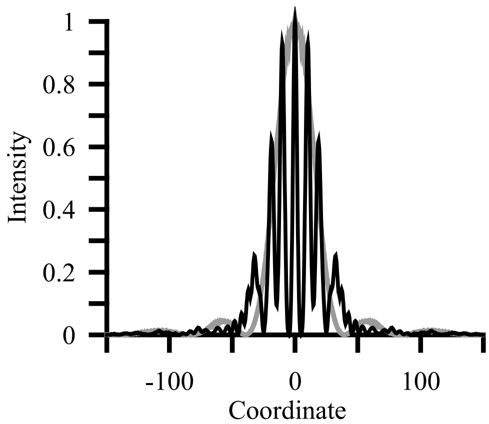

Fig. 6b. Axial cut for the 5 slit object at a distance z = 200λ

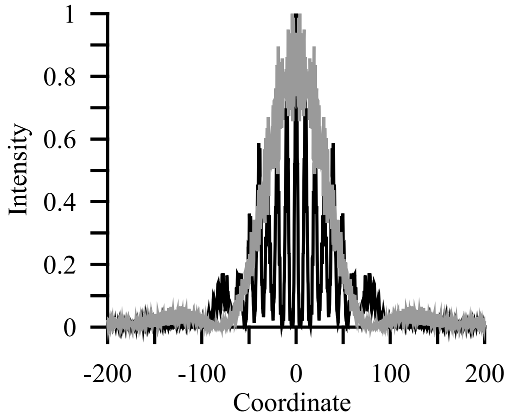

Fig. 6c. Axial cut for the 11 slits object at a distance z = 400λ

5. Diffraction behaviour of a phase structure

To simulate the transfer behaviour of optical beam shaping elements, it seams to be useful, to consider the intensity distribution in dependence on z of phase objects. To simulate an optical lens in the opening a phase function is assumed.

$$V_0(x_0, y_0) = \mathrm{rect}\left(\frac{x_0}{2a}\right) \cdot \exp\left(i\frac{k}{2f}x_0^2\right)$$

The Fig. 7a shows the light distribution behind a screen with a slit of the width of 20λ.The intensity distribution caused by an object of the same width with the phase function is shown in Fig. 7b. The phase function acts as an optical lens. At the distance z = f = 71λ e.g. the distance of the focal length, one obtains a very narrow intensity line. This line corresponds to the Fraunhofer diffraction image of the input line. To underline this, we consider the axial cuts in the focal plane at the distance z = f = 71λ in Fig. 7c. The dashed line describes the image intensity without lens. One can see the Fresnel image of the rectangular opening. The opening with phase structure realises the Fraunhofer image or the Fourier transform in the focal plane. The position of the minimum is given by eq. (17).

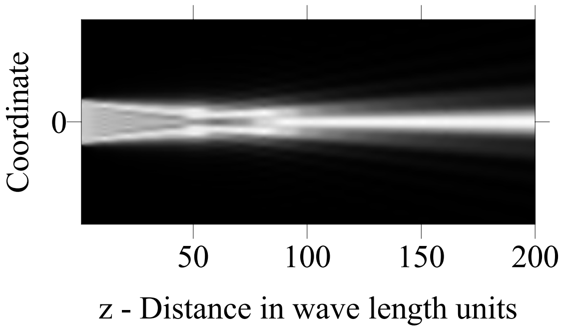

Fig. 7a. Diffraction image of a single slit of width 20

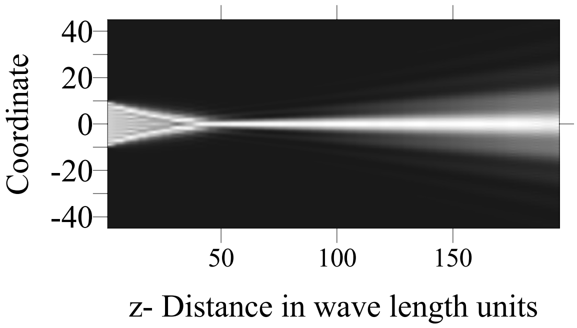

Fig. 7b. Diffraction image of a slit of width 20 with a square phase modulation (lens) of focal length 71

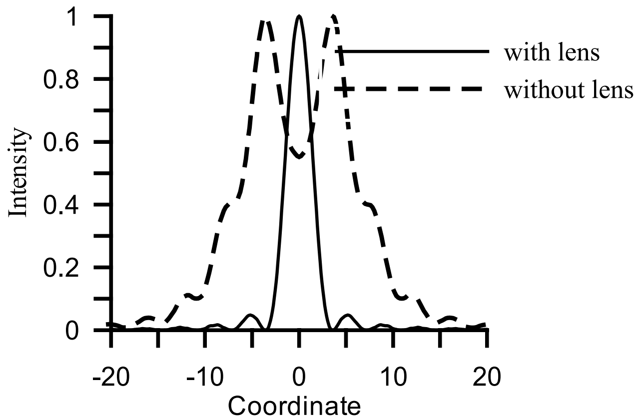

Fig. 7c. Axial cut through the light intensity distribution created by a 20 wide slit opening with and without lens at z = 71

6. Summary

The angular spectrum method was used to calculate light intensity distributions behind different optical elements. The advantage of the angular spectrum method is the compact description. So it is not necessary to distinguish between Fresnel and Fraunhofer approximation and the limitation to the paraxial region is arisen.
The transition between Fresnel and Fraunhofer diffraction was discussed in detail. This transitional region is characterised by Fresnel numbers between 1 and 0.25. For applications in optical metrology or lithography one has to find a strong correlation between the slit/ line width and the intensity relation in dependence on the distance behind the screen. In the Fresnel region the edge position is given at an intensity level of 0.25 respectively at the position of the minimum if the slit is not resolved for Fresnel numbers between 2 and 1. In Fraunhofer region the information about the slit width can be obtained by analysing eq. (17). In the transitional region between Fresnel and Fraunhofer diffraction it seams impossible to find a strong correlation to define a line width.
The diffraction images of periodical repeating structures were investigated with the angular spectrum method as well. The calculated intensities show the effect of self imaging (Talbot effect), contrast reversal and frequency doubling. Analysing the pictures in detail one can see that the so called "Talbot-image " or self image is equal to the corresponding diffraction images. The decision if Fresnel or Fraunhofer image is shown depends on the Fresnel number as well. One has to notice that in the Fresnel number the width of the whole object is included.
As an example for beam shaping elements one simple lens is discussed. The investigations in that direction could be extended in the next time.
The case of oblique illumination has to be analysed also.

References

1. R.J.Blaikie, S.J.Nab, Simulation study of "perfect lenses" for near-field optical nanolithography, Microelectronic Engineering 61-62 (2002),97-103
2. J. D. Gaskil, Linear Systems, Fourier Transforms and Optics, John Wiley&Sons, New York 1978
3. Eugene Hecht, Optik, Addison-Wesley, Bonn, Paris...1994
4. St. Sinzinger, J. Jahns, Microoptics ,Wiley-VCH, Weinheim, 2002
5. B. E. A. Saleh, M. C. Teich, Fundamentals of Photonics, John Wiley & Sons Inc., New York, 1999
6. M. Born, E. Wolf, Principles of Optics, Pergamon Press, Oxford, 1985
7. J. Jahns, ElMeedhi ElJoudi, D. Hagedorn, S. Kinne, Optik 112, (2001), 295-298