FOCAL-PLANE FIELD WHEN LIGHTING DOUBLE-RING PHASE ELEMENTS

A.V. Ustinov 1,2
1 Image Processing Systems Institute of RAS – Branch of the FSRC “Crystallography and Photonics” RAS, Samara, Russia,
2 Samara National Research University, Samara, Russia

Abstract

The focal-plane field amplitude is calculated when lighting double-ring phase elements by flat and Gaussian beams. Emerging conditions in the minimum or maximum centers, including flat-top maxima, are given. For the field amplitude, we obtain equations that define the radius of the first zero-intensity ring based on the deduced expressions. The root values are listed for several parameters of optical elements and incident beams due to the lack of analytical solutions. Numerical simulation results are given for flat incident beams; they are fully consistent with the theoretical calculations.

Keywords: phase optical elements, double-ring phase elements, focal spot size.


Acknowledgements: This work was financially supported by the Russian Foundation for Basic Research (grant No. 16-07-00825).

Introduction

One of the simplest diffractive optical elements (DOE) is a circular plate with thickness step corresponding to phase difference of 180 degrees. Despite its design simplicity, this element happens to have several practical applications.

In particular, such plates are used to adjust intensity of focused Gaussian beams [1]. Moreover, we mean not to focus on a ring/square, but somewhat differently – to obtain a domain with dimensions close to diffraction limits with approximately uniform intensity (not necessarily with sharp edges). DOE proposed in paper [1] in the form of a ring or square step is the simplest element among all considered elements used to solve similar problems [2–4]. In paper [5], two stepped elements are used for one-dimensional dynamic beam correction.

Another application for phase-step plates is sharp focusing. Phase steps are used in accordance with the first Hermite-Gaussian mode to excite an electric-field longitudinal component at the focus center; this was considered in papers [6–8]. It is shown in papers [9, 10] that a focal pattern and, accordingly, contribution of the longitudinal component to the central part of focal domain is influenced by mutual arrangement of polarization axis and a phase step curve for the plate. Besides, in paper [11], the phase-step plate is used to study polarization sensitivity of a near-field microscope.

Radial-step phase plates are also used for sharp focusing and in polarization transformations [12]. In this case, they are capable to perform similar functions like either of multi-ring or multi-level phase plates [13–16]. Iterative or other optimization algorithms are often used to calculate them. The advantage of simple elements (binary one- or two-ring) lies in their ability to analytically evaluate their activity [17, 18].

Sharp focusing of radial-polarized beams is considered in paper [17]. Interrelation between width of a diaphragm ring with the introduced radial phase step and focal spot dimensions and intensity values is also studied there. It is shown that due to destructive interference generated by rings with different phases, it is possible to skip scalar limits related to the first zero of the zero-order Bessel function. The minimum focal spot size (FWHM=0.33λ) is achieved when diaphragm ring width is 20 per cent of radius of a full aperture. In this case, intensity of side lobes does not exceed 30 per cent of a central peak. It is also shown that by introducing phase steps and simultaneously widening a ring aperture, it is possible to generate the focal spot not exceeding the size limit corresponding to a narrow ring aperture while intensity increases almost 6 times. Furthermore, side lobes embrace 35 per cent of the central peak.

Paper [18] cites evidence that two annular zones are sufficient for many problems. Analytical estimates are obtained for parameters of a two-zone element in off-axis scalar approximation; the element provides maximum values of interference maxima on optical axis. Zone radii coincided with zone-plate radii. However, a zone plate with free-space limiting numerical aperture has its central zone \( r_1 = 1.12 \lambda \) in radius, and remaining zones are represented by rings less than \( \lambda/2 \) wide. In this case, decaying waves get mostly through a peripheral part of the optical element; they slightly affect distribution near optical axis. Therefore, it is possible to discuss influence of only one or two central zones. It is shown numerically and analytically that an optical microelement, consisting of only two axially aligned annular zones, can be used for sharp focusing of laser radiation. Moreover, the greatest focusing degree is achieved when the central zone is \( \lambda/2 \) in radius.

Here, as in paper [18], influence of the total element (instead of the ring element) is regarded in far-field diffraction. Though a more accurate off-axis model is used in paper [18], analytical calculation was carried out only for the optical axis field. The transverse in-plane field corresponding to the lens focal plane is analyzed in this paper.

1. Theoretical study of double-ring DOEs

Let’s consider a phase DOE with radius \( r_2 \); its structure is shown in Fig. 1. The phase is equal to zero in the internal circle of radius \( r_1 \) and to 180 degrees in a ring between the radii \( r_1 \) and \( r_2 \). Let’s consider the field being generated in the lens focal plane when lighting this element.

In this paper, we examine two types of incident beams: flat-top and Gaussian beams.
The field amplitude in the lens focus plane is determined by the radially symmetrical Fourier transform as follows:

$$A(\rho) = \int_0^\rho A_0(\rho) J_0 \left( \frac{k \rho r}{f} \right) r \, dr,$$

where $A_0(\rho)$ is the field amplitude in the input plane, $J_0(x)$ is the Bessel function of the zero order, $k = 2\pi \lambda / \lambda$ is the wavenumber, $\lambda$ is the wavelength of an illuminating beam, $f$ is the focal length of a lens, which is combined with the a double-ring phase DOE. In our case, we have the following:

$$A_0(\rho) = \begin{cases} 
1, & 0 \leq \rho \leq r_1 \\
-1, & r_1 < \rho \leq r_2 
\end{cases}.$$

(2)

Substituting Eq. (2) in Eq. (1), we obtain the following expression for the amplitude:

$$A(\rho) = \frac{f}{k \rho} \left[ 2r_1 J_1 \left( \frac{k \rho r}{f} \right) - r_2 J_1 \left( \frac{k \rho r}{f} \right) \right].$$

(3)

The zero-point amplitude is equal to the following:

$$A(\rho = 0) = r_1^2 - r_2^2.$$

(4)

$A(\rho)$ is equal to zero, when $r_1^2 = r_2^2 / 2$; provided $r_1^2 > r_2^2 / 2 \ A(0) > 0$, otherwise $A(0) < 0$.

Since

$$A'(\rho) = -\frac{k^2}{f} A_0(\rho) J_0 \left( \frac{k \rho r}{f} \right) r^2 \, dr,$$

we see that a critical point will always be zero. The second derivative is equal to the following:

$$A''(\rho) = -\left( \frac{k^2}{f} \right)^2 \int_0^\rho A_0(\rho) J_0 \left( \frac{k \rho r}{f} \right) r^2 \, dr,$$

$$A''(\rho = 0) = -\left( \frac{k^2}{f} \right)^2 \left( \frac{r_1^4}{4} - \frac{r_2^4}{8} \right).$$

wherefrom we obtain $A''(\rho = 0) = 0$, when $r_1^4 = r_2^4 / 2$; provided $r_1^4 > r_2^4 / 2 \ A''(0) < 0$ (maximum), otherwise $A''(0) > 0$ (minimum). For $r_1^4 = r_2^4 / 2$ an extremum shall be flat.

Thus, using the Taylor series expansion, we can prove that this will be the maximum.

Let us introduce the following notation:

$$\mu = \frac{r_1}{r_2} (\leq 1).$$

(5)

Thus, the following three configurations are possible (Fig. 2): the negative minimum (provided $0 \leq \mu < 1 / \sqrt{2}$); the positive minimum (provided $1 / \sqrt{2} < \mu < 1 / \sqrt{2}$), and the positive maximum (provided $1 / \sqrt{2} < \mu \leq 1$). The fourth configuration, i.e. the negative maximum, is impossible.

$$\rho_0 = \frac{x_0 \cdot f}{kr_2}.$$

(7)

It is directly shown in Eq. (6) that at $\mu = 0$ and $\mu = 1$, $x_0 = j_{1.1}$ is the first root of the function $J_1(x)$. Thus solving Eq. (6), we obtain the following values.

Table 1. Zero amplitude position versus ring radius ratio

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$</td>
<td>3.832</td>
<td>3.742</td>
<td>3.513</td>
<td>3.212</td>
<td>2.866</td>
<td>2.462</td>
<td>1.912</td>
</tr>
</tbody>
</table>

The first nonzero root of this equation is considered. The equation is solved numerically by using Bessel function tables, and then, if necessary, the root curve $x_0(\mu)$ can be approached to the analytic curve. If we know $x_0$, we can determine the radius of the dark ring as follows:

$$\rho_0 = \frac{x_0 \cdot f}{kr_2}.$$

(8)

In this case, the field amplitude in the input plane is as follows:

$$A_0(\rho) = e^{-i r^2 / 2 \sigma^2} \left\{ \begin{array}{ll}
1, & 0 \leq r \leq r_1 \\
-1, & r_1 < r \leq r_2
\end{array} \right.$$
if the pupil’s radius meets the condition $r_2 > \sim (2\pm 2.5)\sigma$ when an upper limit could be replaced by an infinite limit. The following formula is given in tables presented in paper [19] (p. 186).

$$
\int_0 e^{-\nu^2} J_0 (\nu x) \, dx = \frac{\sqrt{\pi}}{8\nu^2} e^{\frac{\nu^2}{4}} I_{(v-1)/2} \left( \frac{\nu^2}{8\nu} \right) - I_{(v+1)/2} \left( \frac{\nu^2}{8\nu} \right),
$$

(10)

When $\nu = 0$, the formula is simply as follows:

$$
\int_0 e^{-\nu^2} J_0 (\nu x) \, dx = \frac{1}{2\nu} e^{-\nu^2}.
$$

(10a)

Using Eq. (10a), we obtain the following:

$$
B_1 + B_2 = \frac{\sqrt{\pi}}{8\nu^2} e^{\frac{\nu^2}{4}} \int_0 J_0 \left( \frac{k\nu}{\nu} \right) e^{-\nu^2} r \, dr = \sigma^2 e^{-\frac{1}{2} \nu^2}.
$$

(11)

We can calculate the value of $B_2$ with the specified degree of accuracy replacing a Gaussian function bell with a parabolic cone, when $r_1 \leq \sigma$ (a knee of the Gaussian curve). The case when this condition isn’t well fulfilled is uninteresting: an exterior ring has no meaningful effect. So, the result is limited to one circular hole to be lighted.

Suppose the condition $r_1 \leq \sigma$ has been satisfied. Then the following approximation should be made with a small accuracy over the range of $r \leq r_1$:

$$
e^{-\frac{\nu^2}{2}} = \alpha_1 - \alpha_2 r^2.
$$

(12)

The values of $\alpha_1$ and $\alpha_2$ depend on used approximation and they are given below. Using Eq. (12), we receive the following:

$$
2B_2 = 2 \int_0 (\alpha_1 - \alpha_2 r^2) J_0 \left( \frac{k\nu}{\nu} \right) e^{-\nu^2} r \, dr = 2 \int_0 J_1 \left( \frac{k\nu}{\nu} \right) \left( \alpha_1 - \alpha_2 r^2 \right) + 4 \int_0 J_2 \left( \frac{k\nu}{\nu} \right) \alpha_2.
$$

(13)

The overall amplitude is equal to the following:

$$
A = 2 \int_0 J_1 \left( \frac{k\nu}{\nu} \right) \left( \alpha_1 - \alpha_2 r^2 \right) + 4 \int_0 J_2 \left( \frac{k\nu}{\nu} \right) \alpha_2 = \sigma^2 e^{-\frac{1}{2} \nu^2}.
$$

(14)

This shows that for large values of $\nu$ decreasing is determined with the first summand and the rate of decay is identical to that one valid for lighting with flat-top beams. When $r_1 = 0$ (no phase step), we get a trivial response:

$$
A = \sigma^2 e^{-\frac{1}{2} \nu^2},
$$

which is the Gaussian beam.

Before proceeding any further, it’s worthwhile to get at least one technique for finding the values $\alpha_1$ and $\alpha_2$. The simplest method, though probably not the best one in terms of its accuracy, is a collocation method. We equal zero and edge values as follows:

$$
1 = \alpha_1,
$$

(15)

$$
e^{-\left(\nu^2/2\sigma^2\right)} = \alpha_2 = (1 - e^{-\left(\nu^2/2\sigma^2\right)}) \cdot \frac{1}{r_1^2}.
$$

Substituting Eq. (15) in Eq. (14), we obtain the following:

$$
A = 2 \int_0 J_1 \left( \frac{k\nu}{\nu} \right) e^{-\frac{1}{2} \nu^2} + 4 \int_0 J_2 \left( \frac{k\nu}{\nu} \right) \left( 1 - e^{-\frac{1}{2} \nu^2} \right) - \sigma^2 e^{-\frac{1}{2} \nu^2}.
$$

(16)

The value of $A(\nu = 0)$ can be calculated exactly without any approximations made. It equals to the following:

$$
A(\nu = 0) = \sigma^2 \left( 1 - 2 e^{-\left(\nu^2/2\sigma^2\right)} \right).
$$

(17)

Providing however that $r_2$ is larger (see the line before Eq. (10)), the latter summand shall be neglected:

$$
A(\nu = 0) = \sigma^2 \left( 1 - 2 e^{-\left(\nu^2/2\sigma^2\right)} \right).
$$

(17a)

Configurations will be determined according to Eq. (17a) being similar for derivatives. We write out accurate values, just as in Eq. (17), but they are less convenient for analysis due to their complexity.

$$
A(0) \neq 0, \quad \text{when } r_1 = \sigma \sqrt{2\ln 2} = 1.1774\sigma.
$$

(18)

At less values of $r_1$, we have $A(0) < 0$, otherwise we have $A(0) > 0$.

The first derivative $A'(\nu = 0)$ equals to zero and the second one is as follows:

$$
A''(\nu = 0) = -\left( \frac{k^2}{f} \right) \left( \sigma^2 - \sigma^4 (r_1^2 + 2\sigma^2) e^{-\frac{1}{2} \nu^2} + 0.5\sigma^4 (r_1^2 + 2\sigma^2) e^{-\frac{1}{2} \nu^2} \right).
$$

(19)

With disregard for the latter summand, we get the following:

$$
A''(\nu = 0) = -\left( \frac{k^2}{f} \right) \sigma^2 \left( \sigma^2 - (r_1^2 + 2\sigma^2) e^{-\frac{1}{2} \nu^2} \right).
$$

(19a)

$A''(0) = 0, \quad \text{when } r_1 = 1.832\sigma.
$$

(20)

where 1.832 is the root of the equation $1 = (x^2 + 2)e^{-x^2/2}$. This corresponds to the flat maximum at a smaller vale of $A''(0) > 0$ (minimum), otherwise we have $A''(0) < 0$ (maximum). Configuration change takes place similarly as in case with flat-top beams (Fig. 2).

Configuration I: $r_1 < 1.1774\sigma$;

Configuration II: $1.1774\sigma < r_1 < 1.832\sigma$;

Configuration III: $r_1 > 1.832\sigma$. 
It is difficult to compare edges obtained from fully exact expressions because of the fact that there are two radii, not one, in exact expression. However, we can compare them with those ones that will be used in Eq. (16). Using two summands of the Taylor series \((J_1(x) \approx \pi/2 - (x^2/16)\), \(J_2(x) \approx (x^2/8) - (x^4/96)\)), the near-axis amplitude can be conceived of as follows:

\[
A = \left[\frac{\pi}{2} - \left(1 + e^{-x^2/2\sigma^2}\right) - \sigma^2\right] + \left(\frac{k}{f}\right)^2 \left[-\frac{x}{12} \left(1 + 2e^{-x^2/2\sigma^2}\right) + \frac{\sigma^2}{2}\right].
\]

Though Eq. (21) is approximate, the zero value is exact; it equals to the following:

\[
A(0) = \frac{\pi}{2} - \left(1 + e^{-0}\right) - \sigma^2,
\]

wherefrom

\[
A(0) = \frac{\pi}{2} - 1.15\sigma,
\]

where 1.15 is the root of the equation \(x^2(1+e^{-x^2/2\sigma^2}) = 2\); it is somewhat less than the factor given in Eq. (18). The second derivative is equal to the following:

\[
A''(0) = \frac{k}{f} \left[-\frac{x}{12} \left(1 + 2e^{-x^2/2\sigma^2}\right) + \frac{\sigma^2}{2}\right],
\]

wherefrom

\[
A''(0) = 0 \quad \text{at} \quad r_1 = 1.687\sigma,
\]

where 1.687 is the root of the equation \(x^2(1+2e^{-x^2/2\sigma^2}) = 12\); it is less than the factor given in Eq. (20). Overall, the area of configuration II is narrower than that one obtained by exact calculation. It should be pointed out that the factor given in Eq. (22) and especially in Eq. (23) is larger than permitted; this is necessary to make the approximation mentioned in Eq. (12), on the basis of which Eq. (16) has been obtained, be good.

Zero values of Eq. (16) shall be determined using the below tables. With the use of

\[
x = (kr_1/f);
\]

we have \(A(p) = 0\) be written as follows:

\[
2\mu^2 \cdot \frac{J_1(x)}{x} - e^{-\mu^2/2} + 4\mu^4 \cdot \frac{J_2(x)}{x} \cdot \left(1 - e^{-\mu^2/2}\right) = e^{-x^2/(2\sigma^2)}.
\]

The equation is solved for \(x\) at given \(\mu\). If we know the root of \(x_0\), the radius of the dark ring is determined as follows:

\[
\rho_0 = \frac{x_0}{\mu} = \frac{f}{k\sigma}.
\]

In contrast to Eq. (7), a real factor is not \(x_0\) itself, but \(x_0/\mu\). That’s why we added a proper line in the table of roots of equation (25) (implication of one more line will be clarified below).

<table>
<thead>
<tr>
<th>(\mu)</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
<th>1.1</th>
<th>1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_0)</td>
<td>0</td>
<td>0.3043</td>
<td>0.51</td>
<td>0.669</td>
<td>0.79</td>
<td>0.879</td>
<td>0.936</td>
<td>0.961</td>
<td>0.948</td>
<td>0.888</td>
<td>0.755</td>
<td>0.464</td>
<td>0</td>
</tr>
<tr>
<td>(x_0/\mu)</td>
<td>(\infty)</td>
<td>3.043</td>
<td>2.55</td>
<td>2.23</td>
<td>1.975</td>
<td>1.757</td>
<td>1.56</td>
<td>1.372</td>
<td>1.185</td>
<td>0.986</td>
<td>0.755</td>
<td>0.422</td>
<td>0</td>
</tr>
<tr>
<td>(2\mu\sqrt{-\ln \mu})</td>
<td>0</td>
<td>0.3035</td>
<td>0.5075</td>
<td>0.658</td>
<td>0.766</td>
<td>0.833</td>
<td>0.858</td>
<td>0.836</td>
<td>0.756</td>
<td>0.584</td>
<td>0</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Then the left-hand side of the equation will be averagely equal to the following:

\[
2\mu^2 \cdot \frac{1}{2} + \frac{1}{8} + \frac{1}{4} \mu^2 = \mu^2 + \frac{1}{4} \mu^2 = \mu^2,
\]

so, we get the following equation:

\[
\mu^2 = e^{-x^2/(2\sigma^2)} \rightarrow x_0 = 2\mu \sqrt{-\ln \mu}.
\]

The ratio \(x_0/\mu = 2\sqrt{-\ln \mu}\) goes up with no limit when \(\mu \to 0\).

2. Numerical simulation results

Numerical simulation was implemented under the flat incident beam with the following parameter: the outer radius is \(r_2 = 100\lambda\) (\(\lambda\) is the wavelength of incident radiation equalized to 1 μm); the lens focal length is \(f = 400\lambda\). Let \(S\) denote a ratio between the side lobe intensity and the centered intensity. The results are given in Fig. 3-6.

As we can see, when the size of the central light spot decreases, the energy goes out of it into a peripheral ring.
Focal-plane field when lighting double-ring phase elements

A.V. Ustinov


Fig. 3. Focusing results for the element with phase step radius \( r_1 = 0.6r_2 \): the element’s interior-part field (dashed line, \( \text{FWHM} = 3.43\lambda, S = 0.015 \)), the element’s front-end field (solid line, \( \text{FWHM} = 1.74\lambda, S = 0.123 \)) and their superposition (dotted line, \( \text{FWHM} = 1.2\lambda, S = 1.6 \)). Amplitudes (a) and intensities (b) are shown.

Fig. 4. Focusing results for: \( r_1 = 0 \) (FWHM = 2.05\( \lambda \), dotted line), \( r_1 = 0.25r_2 \) (FWHM = 1.91\( \lambda \), dashed line), and \( r_1 = 0.5r_2 \) (FWHM = 1.51\( \lambda \), solid line).

Fig. 5. Focusing results for: \( r_1 = 0.6r_2 \) (FWHM = 1.22\( \lambda \), dashed line), \( r_1 = 0.65r_2 \) (FWHM = 0.96\( \lambda \), solid line), and \( r_1 = 0.7r_2 \) (FWHM = 0.35\( \lambda \), dotted line) (a); central part for \( r_1 = 0.7r_2 \) (b).

Fig. 6. Focusing results for: \( r_1 = 0.75r_2 \) (dashed line), \( r_1 = 0.8r_2 \) (solid line), \( r_1 = 0.85r_2 \) (FWHM = 3.9\( \lambda \), dotted line).

By parameter values from Eq. (7) used in simulation, we get the following:

\[
\rho_0 = (0.637x_0) - \lambda. \tag{27}
\]

Values of the radius of the first dark ring calculated by Eq. (27) and determined according to certain intensity graphs obtained from simulation are shown in Table 3. Values of \( x_0 \) are taken from Table 1 or obtained by interpolation.

Table 3. Comparison of zero points discovered theoretically and by simulation

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>0</th>
<th>0.25</th>
<th>0.5</th>
<th>0.6</th>
<th>0.65</th>
<th>0.7</th>
<th>0.75</th>
<th>0.8</th>
<th>0.85</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_0 ) as in (27)</td>
<td>2.44</td>
<td>( \approx 2.148 )</td>
<td>1.568</td>
<td>1.218</td>
<td>( \approx 0.785 )</td>
<td>0.352</td>
<td>( \approx 3.726 )</td>
<td>3.528</td>
<td>( \approx 3.289 )</td>
</tr>
<tr>
<td>( \rho_0 ) as in graph</td>
<td>( \approx 2.5 )</td>
<td>( \approx 2.2 )</td>
<td>( \approx 1.5 )</td>
<td>( \approx 1.2 )</td>
<td>( \approx 0.85 )</td>
<td>( \approx 0.32 )</td>
<td>( \approx 3.75 )</td>
<td>( \approx 3.5 )</td>
<td>( \approx 3.3 )</td>
</tr>
</tbody>
</table>

Values of the radius are given in units of wavelength. Designated values mean the use of linear interpolation. Note that for \( \mu = 0.25 \) and 0.85 interpolation produces adequate accuracy; for \( \mu = 0.75 \) – mean accuracy, and for \( \mu = 0.65 \) – low accuracy. This is due to the fact that interpolation accuracy increases when a derivative of the interpolated function approximates a constant.

It may be concluded that compatibility between theoretical calculations and numerical simulation results is very high.

Conclusion

Analysis of operation of double-ring optical elements allowed us to come to the following conclusions:

1. Three significant configurations for amplitude (two – for intensity) are possible near optical axis. Edge configurations are supposed to be the centered zero value and the largest flat-topped value similar to the configuration mentioned in paper [1]. A type of configuration is determined by the ratio between inner and outer radii at flat beam incidence or by the ratio between inner radius and width of the Gaussian beam.

2. Equations are derived for the first zero intensity position. Tables of their roots are given.

3. Numerical simulation implemented for the case of incident beams has fully confirmed theoretical predictions.

References


