45

Optics Communications 90 (1992) 11–15
North-Holland

OPTICS COMMUNICATIONS

Focal spot estimation for concave diffraction gratings

L. Tsonev and E. Popov
Institute of Solid State Physics, Bulgarian Academy of Sciences, blvd. Trakia 72, 1784 Sofia, Bulgaria

Received 6 November 1991; revised manuscript received 27 January 1992

Two simple formulae are proposed, which estimate the width and the height of the focal images generated by a concave holographic grating under point source illumination. The results are compared with focal spot diagrams evaluated for some flat field spectrograph gratings and monochromator gratings. Good correlation is observed between analytical estimations and ray tracing.

1. Introduction

In the last two decades concave gratings find an impressively increasing application in optics. There are at least two reasons for such a high interest: (i) in many cases concave gratings allow to reduce drastically the weight and the dimensions of spectrographs and monochromators; (ii) holographic recording is a very flexible procedure and makes it possible to produce, if not universal, at least, specialized gratings, well adapted for solving specific classes of problems. Concave gratings perform their dispersion function due to the groove periodicity, and their imaging function due to the concave reflecting substrate, combined with period and groove form variations over the illuminated area. The imaging ability is connected only with the dimensions (width W^{M} and height H^{M}) of the meridional focus spots, while the resolving power depends also on the distance between them.

In practice, spot dimensions are exactly predicted by the ray tracing, i.e. by numerical grating modeling. In such a way the aberration sources are taken into account simultaneously as a complete set. Since in most practical cases the operating conditions are far from the diffraction limit, this geometrical optics approach gives sufficiently correct information about focal spot form and dimensions [1–3].

Together with the ray tracing an analytical approach is developed in the literature [4,5]. It is based on the analysis of the aberration function and on its expansion in power series upon the meridional w and sagittal l grating coordinates. Although approximated (only the first 5–6 expansion terms – i.e. only the strongest aberrations – are considered) this approach has its advantages: (i) it allows to find explicit general connections between substrate and groove characteristics on one hand and focal spot form and dimensions on the other hand; (ii) demonstrating the role of different parameters it assists the intuition; (iii) being relatively simple and explicit, it ensures estimating of the expected grating behavior in a fast and convenient manner.

The aim of the present communication is to propose two simple formulae, describing the dimensions (width and height) of the focal images of a meridional point source, as well as to compare these results with complete numerical ray tracing.

2. Preliminary remarks

The general ideas of the analytical approach to concave grating aberrations are formulated in detail in some excellent papers [4,6]. Unfortunately, the notations are still far from a commonly accepted unification. In order to prevent misunderstandings and give a self-consistent presentation we declare here our version of the basic expressions.

The location of source, grating, image and coordinate systems is shown in fig. 1. The point source A and its ideal image B_0 are assumed to lie in the meridional plane. The aberration function δF is expanded in series with the following coefficients [6]

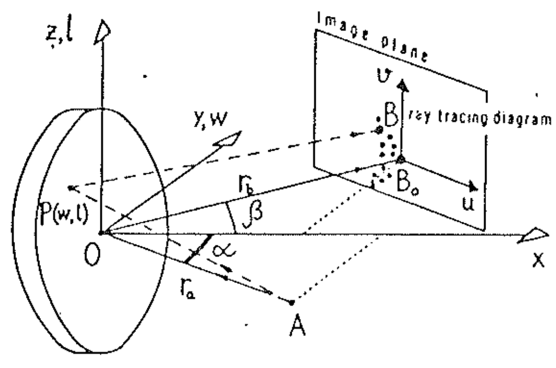

Fig. 1. Operating scheme of a concave grating. A, polychromatic point source in the meridional plane xOy. B_0, ideal image of A. B, one aberrated image point included in the ray tracing real image of A. u, v, image (aberration, deviation) coordinates in the image plane.

$$\delta F = wF_{100} + \tfrac{1}{2}w^2F_{200} + \tfrac{1}{2}l^2F_{020} + \tfrac{1}{2}w^3F_{300} + \tfrac{1}{2}wl^2F_{120} + \dots, \qquad (1)$$

where $F_{100}=0$ determines the diffraction angle, $F_{200}=0$ and $F_{020}=0$ determine the meridional and the sagittal focal distance respectively. The coefficients have a common general structure,

$$F_{ij0} = M_{ij0} - m\lambda N_{ij0},$$

λ being the wavelength and $m=+1$ the diffraction order. The first (mounting) part M_{ij0} describes only the mounting scheme, while the second (groove) part N_{ij0} characterizes the form and the location of the grooves. Moreover, the mounting part can be expressed further as a sum of two terms with identical form, differing only in the coordinates included,

$$M_{ij0} = M_{ij0}(\mathrm{A}) + M_{ij0}(\mathrm{B}_0).$$

Here A means the coordinates (r_a, α) of the point source, and B_0 the coordinates (r_b, β) of its ideal image. In the simplest case of a spherical substrate with a radius of curvature R we have [4]

$$M_{100}(\mathrm{A}) = -\sin\alpha,$$

$$M_{200}(\mathrm{A}) = \frac{\cos^2\alpha}{r_a} - \frac{\cos\alpha}{R},$$

$$M_{020}(\mathrm{A}) = \frac{1}{r_a} - \frac{\cos\alpha}{R},$$

$$M_{300}(\mathrm{A}) = \frac{\sin\alpha}{r_a}\left(\frac{\cos^2\alpha}{r_a} - \frac{\cos\alpha}{R}\right),$$

$$M_{120}(\mathrm{A}) = \frac{\sin\alpha}{r_a}\left(\frac{1}{r_a} - \frac{\cos\alpha}{R}\right).$$

The groove part N_{ij0} can be expressed in a similar manner for the case of holographic recording as a difference between two terms with identical form, the first of them including the coordinates of the recording source C (r_c, γ) and the second one the coordinates of the recording source D (r_d, δ):

$$N_{ij0} = N_{ij0}(\mathrm{C}) - N_{ij0}(\mathrm{D}).$$

The quantities $N_{ij0}(\mathrm{Q})$ are simply related to $M_{ij0}(\mathrm{Q})$, namely

$$N_{ij0}(\mathrm{Q}) = M_{ij0}(\mathrm{Q})/\lambda_0,$$

where λ_0 is the recording wavelength and Q is any point represented by its polar coordinates.

In the image plane a new coordinate system is introduced (fig. 1); its meridional and sagittal axes are denoted by u and v respectively. The real focal spot is considered as a set of points $\{B(u, v)\}$ each of them originating from the intersection between the ray APB (diffracted at the grating point $P(w, l)$) and the image plane. A sufficiently exact connection between the image coordinates and the aberration function is given by the expressions [7]

$$\frac{u}{r_b} = \frac{1}{\cos\beta}\frac{\partial}{\partial w}(\delta F), \quad \frac{v}{r_b} = \frac{\partial}{\partial l}(\delta F).$$

If we consider only the cases $\cos\beta \approx 1$ we obtain the following equations,

$$u/r_b = F_{100} + wF_{200} + \tfrac{3}{2}w^2F_{300} + \tfrac{1}{2}l^2F_{120},$$

$$v/r_b = lF_{020} + wlF_{120}.$$

F_{200} and F_{020} are the meridional and the sagittal astigmatism, F_{300} and F_{120} are the classical and the mixed coma.

The free term F_{100} is always set equal to zero by locating the origin of the image coordinate system exactly in the ideal image B_0. The respective ray AOB_0 diffracts at the grating center and satisfies the dispersion equation there: $\sin\alpha + \sin\beta = m\lambda/d_0$, the period for holographically recorded gratings being calculated from the formula $d_0 = \lambda_0/(\sin\gamma - \sin\delta)$. In the common case the most significant from the remaining terms are the astigmatic ones. For improvement of resolution minimization of spot width is required. It can be achieved by zeroing the coefficient F_{200}, i.e. by placing the image plane and the detector exactly on the meridional focal curve, at a distance r_b^M where $F_{200}(r_b^M)=0$. Further on all the quantities satisfying $F_{100}=F_{200}=0$ get an upper index M (meridional focus). Now the image coordinates are described by the system

$$u^M/r_b^M=\tfrac{3}{2}w^2F_{300}^M+\tfrac{1}{2}l^2F_{120}^M, \tag{2}$$

$$v^M/r_b^M=lF_{020}^M+wlF_{120}^M. \tag{3}$$

3. Spot dimensions estimation

The spot dimensions are determined by the limits of the interval, where u^M and v^M vary, when w and l take all possible values between $-\Phi/2$ and $\Phi/2$ (Φ, grating diameter).

Let us analyze the width W^M of the meridional focal spot, using expression (2) for the meridional deviation u^M. In order to find an upper estimate we shall take the absolute values of the aberration coefficients in eq. (2) and form the modified coordinate $(u^M)^*$ as follows

$$(u^M)^*/r_b=\tfrac{3}{2}w^2|F_{300}^M|+\tfrac{1}{2}l^2|F_{120}^M|.$$

Since w and l are included here only to their second power, the quantity $(u^M)^*$ is not negative. When $w=l=0$, it reaches its minimum which is equal to zero: $(u^M)^*_{min}=0$. When $w=w_{max}=l=l_{max}=\Phi/2$, the quantity $(u^M)^*$ reaches evidently its maximum, which is positive: $(u^M)^*_{max}>0$. Therefore the spot width estimate $\overline{W^M}$ can be expressed as

$$\overline{W^M}=(u^M)^*_{max}-(u^M)^*_{min}=(u^M)^*_{max},$$

i.e.

$$\overline{W^M}=r_b^M\tfrac{1}{2}(\Phi/2)^2(3|F_{300}^M|+|F_{120}^M|). \tag{4}$$

The spot height is determined by the sagittal image deviation v. We take again the absolute values of the coefficients in eq. (3), forming in such a manner the modified coordinate $(v^M)^*$,

$$(v^M)^*/r_b=l|F_{020}^M|+wl|F_{120}^M|.$$

According to this formula when $w=l=0$ we have $(v^M)^*_i=0$. When w and l vary between $-\Phi/2$ and $\Phi/2$, $(v^M)^*$ takes positive as well as negative values between its maximum $(v^M)^*_{max}>0$ and its minimum $(v^M)^*_{min}<0$. The spot height can be evaluated as a sum of these limiting values taking them as positively defined quantities

$$\overline{H^M}=(v^M)^*_{max}+|(v^M)^*_{min}|.$$

We only overstate this estimation, but in the same time we significantly simplify the final expression substituting both addends by a new quantity $\overline{(v^M)^*}$ chosen to satisfy simultaneously the conditions

$$\overline{(v^M)^*}\geqslant(v^M)^*_{max} \quad\text{and}\quad \overline{(v^M)^*}\geqslant|(v^M)^*_{min}|.$$

The quantity just introduced can be evaluated from the general formula for $(v^M)^*$ assuming that $w=w_{max}=l=l_{max}=\Phi/2$. Therefore

$$\overline{H^M}=2\overline{(v^M)^*}$$

$$=2r_b^M[(\Phi/2)|F_{020}^M|+(\Phi/2)^2|F_{120}^M|]. \tag{5}$$

4. Results and discussion

Both estimation formulae (4), (5) need some comments. Let us go back to the series expansion of the aberration function (1). In the common case its terms decrease with increasing indices and coordinate power. Therefore we have to consider one specific term only when all the preceding terms tend to zero.

F_{100} is made always equal to zero by a proper *angular* location of the detector (at the dispersion polar angle β in respect to the grating normal, $\beta=\arcsin(\sin\alpha-m\lambda/d_0)$). The meridional astigmatism F_{200} is also made usually exactly or nearly equal to zero by a proper *linear* location of the detector (at the meridional focal distance r_b^M or in its vicinity). The remaining aberrations are sagittal astigmatism F_{020}^M and coma F_{300}^M, F_{120}^M.

The spot width expression (4) includes only classical and mixed coma terms, they are of one and the same order of magnitude and have to be considered simultaneously. However, the spot height expression $\overline{H^M}$ includes two addends of different order. In the common case the astigmatic term exceeds the comatic one. Practically all the concave gratings are designed to have zero sagittal astigmatism for one or two points in the operating spectral region. Therefore the following recommendation can be made concerning H^{M}: (i) at all the wavelengths where $F^{M}_{020} \neq 0$ one has to consider only the first, the astigmatic term; (ii) at one or two wavelengths λ^* where $F^{M}_{020}(\lambda^*)=0$ and in their vicinity one has to consider only the second, the comatic term; (iii) when $F^{M}_{020} \neq 0$ consideration of the comatic addend can lead to unwanted overstating of the spot height estimation.

Looking for specific examples we decided to apply the proposed formulae to three spherical holographic gratings denoted further as I, II and III and described in detail in ref. [8] (case I) and in ref. [9] (cases II and III). They are designed for the two most popular modern spectral devices, flat-field spectrograph (cases I and II) and Seya-Namioka monochromator (case III). The recording is performed at the wavelength 457.9 nm. All the samples operate in first order in the interval 350–900 nm.

Grating I has a blank diameter $\Phi=92$ mm and radius of curvature $R=210.8$ mm, the groove period in the center being $d_0=1.26135$ µm. The astigmatic coefficient F^{M}_{020} vanishes only one time, at $\lambda=\lambda^*=$ 570 nm.

Gratings II and III have blanks with diameter $\Phi=75$ mm and radius of curvature $R=500$ mm, as well as groove period in the center $d_0=1.611$ µm. In the case II, F^{M}_{020} decreases to zero two times, at $\lambda_1^*=425$ nm and $\lambda_2^*=710$ nm, while in the case III this happens only once, at $\lambda^*=625$ nm.

The results are demonstrated in fig. 2 (case I), fig. 3 (case II) and fig. 4 (case III). The ray-tracing diagrams in fig. 2 are entirely evaluated by us, the diagrams in fig. 4 are taken from ref. [9] and coincide with our tests, the diagrams in fig. 3 are partially taken from ref. [9] (for 350, 625 and 900 nm), partially included by us (for 425 and 710 nm).

The images of a polychromatic point source generated by spectrograph gratings (cases I and II) have typically small astigmatic extension (moderate height to width ratio) and therefore can be presented in one and the same scale in vertical and horizontal directions, fig. 2 and fig. 3. This is not convenient, however, for the astigmatically dominated images given

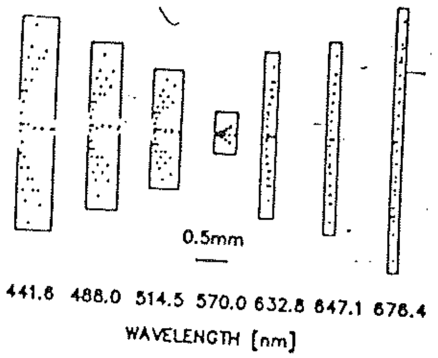

Fig. 2. Comparison between ray tracing diagrams (points) and analytical dimension estimations (rectangles) in the spectrograph grating case I for different wavelengths. Scale is valid only for the spot dimensions, but not for the separating distances.

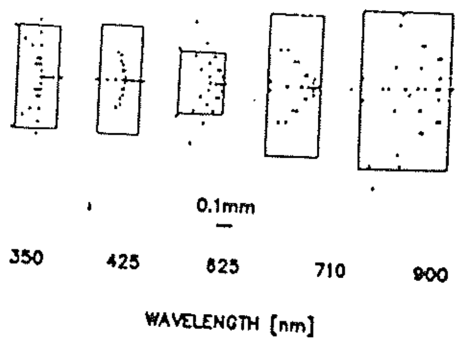

Fig. 3. As fig. 2, but in the spectrograph grating case II.

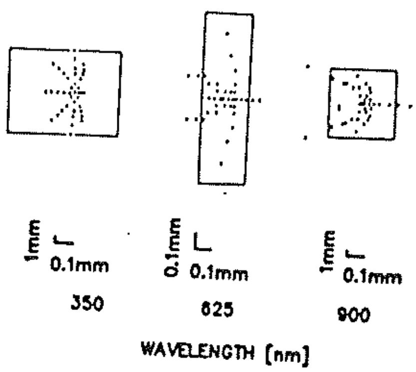

Fig. 4. As fig. 2, but in the monochromator grating case III.

by the monochromator grating III, requiring different scales for each spot in fig. 4.

Each focal spot is combined with an estimating rectangle according to formulae (4) and (5).

Although not exhausting all existing cases, the examples presented and examined here permit to conclude that the formulae (4) and (5) are adapted for easy, fast and realistic estimation of focal spot dimensions, at least far from grazing incidence ($\cos\beta \approx 1$ is always assumed). Sometimes the ray-tracing spots are slightly greater than the estimation. Two comments can be made for this situation: (i) Such small deviation does not play a significant role, only a few rays with small density hit the image plane out of the rectangular frames, i.e. a very weak light intensity can be expected there. (ii) In the cases, where such fine effects are essential, the method proposed here is obviously insufficient and one must use the full aberration function, (i.e. the ray-tracing) as an exact modeling of the imaging properties. The approximation in the formulae proposed here is the price for their simplicity and convenience.

References

[1] H. Noda, T. Namioka and M. Seya, J. Opt. Soc. Am. 64 (1974) 1037.

[2] R. Iwanaga and T. Oshio, J. Opt. Soc. Am. 69 (1979) 1538.

[3] B.J. Brown and I.J. Wilson, Optica Acta 28 (1981) 1587.

[4] T. Namioka, M. Seya and H. Noda, Jpn. J. Appl. Phys. 15 (1976) 1181.

[5] E. Ishiguro, R. Iwanaga and T. Oshio, J. Opt. Soc. Am. 69 (1979) 1530.

[6] H. Noda, T. Namioka and M. Seya, J. Opt. Soc. Am. 64 (1974) 1031.

[7] M. Pouey, J. Optics (Paris) 14 (1983) 235.

[8] E.G. Loewen, E.K. Popov, L.V. Tsonev and J. Hoose, J. Opt. Soc. Am. A 7 (1990) 1764.

[9] W.R. McKinney and C. Palmer, Appl. Optics 26 (1987) 3108.