

Birefringence induced by the spatial dispersion in deep UV lithography: theory and advanced compensation strategy

Alexander Serebriakov^{1*}, Evgenii Maksimov², Florian Bociort¹, Joseph Braat¹

¹Optics Research Group, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands;

²P.N.Lebedev Physical Institute, Leninskiy pr. 53, Moscow 119991, Russia

Received

Birefringence induced by spatial dispersion (BISD), also called intrinsic birefringence, can cause a serious deterioration of the optical imaging quality of deep UV lithographic objectives at wavelengths below 193 nm,

Recently the mathematical formalism for analyzing those aspects of the BISD effect that are relevant for optical design has been published. In this paper we give an equivalent but simplified derivation of these results. This mathematical formalism is then applied to optical system design and a compensation strategy is discussed. An example of optical system is given where the phase retardation caused by the BISD effect has been corrected.

KEYWORDS: lithography, birefringence, crystal optics, optical system design

* a.serebriakov@tnw.tudelft.nl

1. INTRODUCTION

The phenomenon of birefringence in certain types of crystals has been known for more than three centuries. It was first discovered by Bartholinus, qualitatively explained by Huygens and marvelously described in a qualitative way by Fresnel. A firm foundation for the phenomenon of birefringence was obtained by applying Maxwell's equations to crystalline media with specific symmetry properties. In this classical description, cubic crystals do not show birefringence and, indeed, for most purposes they can effectively be considered as being isotropic.

It was Lorentz who first suggested the presence of anisotropy introduced by spatial dispersion in cubic crystals as early as 1878 ¹⁾. This observation was made again ²⁾ on the basis of a microscopic investigation of quadripolar transitions in crystals, and on the basis of macroscopic electrodynamics ³⁾. The detailed analysis of this problem has been carried out by Agranovich and Ginzburg ⁴⁾. They showed, for instance, that cubic crystals possess seven optical axes (the three main crystallographic axes and the four body diagonals of the cube). In this respect cubic crystals can be called heptaxial.

A number of experimental and theoretical studies were carried out in the 1970's on the birefringence effect induced by spatial dispersion in semiconductors ⁴⁾. Because of the very small size of the effect and the absence of any practical applications these investigations have not been done in much detail.

Recent publications by Burnett *et al.* ^{5,6)} demonstrate the great practical importance of the phenomenon. The birefringence induced by spatial dispersion has been measured and calculated for CaF₂ and BaF₂ in the ultraviolet spectral range. It was shown that the magnitude of the BISD in these cubic crystals is sufficiently large to cause serious problems when using CaF₂ for precision UV optical systems at wavelengths as short as 157 nm ⁶⁾. The birefringence effect is extensively described and quantified in these references. However some aspects would profit from a further

analysis. In Section 2 of this paper we present a different theoretical approach to the BISD subject which allows us to simplify the derivations and to obtain fully equivalent analytic expressions of the final formulae for cubic crystals (Section 3). In the final fourth Section we present our calculation method for the BISD and the compensation of its detrimental effect on the image quality in a lithographic objective.

2. BIREFRINGENCE IN CRYSTALLINE MEDIA

We will use the macroscopic electrodynamic approach ⁴⁾ to analyze birefringence in crystals. The macroscopic Maxwell equations for the electromagnetic field quantities in a medium are written as

$$\left. \begin{aligned} \text{rot } \mathbf{B} &= \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{j}^{ext} \\ \text{div } \mathbf{D} &= 4\pi \rho^{ext} \\ \text{rot } \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \\ \text{div } \mathbf{B} &= 0. \end{aligned} \right\} \quad (1)$$

Here \mathbf{E} is the electric field strength, \mathbf{D} and \mathbf{B} are the electric and magnetic induction. The quantities \mathbf{j}^{ext} and ρ^{ext} are the external current and charge densities which are sources of the external electromagnetic field. These equations are supplemented by the relation between the electric induction \mathbf{D} and the electric field \mathbf{E} . This material equation can be written in the framework of macroscopic electrodynamics in the following general form

$$D_i(\mathbf{r}, t) = \int_{-\infty}^t dt' \int d\mathbf{r}' \varepsilon_{ij}(t-t', \mathbf{r}-\mathbf{r}') E_j(\mathbf{r}', t') \quad (2)$$

where $\varepsilon_{ij}(t, \mathbf{r})$ is the dielectric tensor of the medium, and the Einstein summation convention has been used.

Here we have used the principle of causality according to which the induction at time t is only determined by the present field and the field at previous times $t' \leq t$.

We Fourier transform eqs. (1) and (2) assuming that

$$E_i(\mathbf{r}, t) = \int E_i(\omega, \mathbf{k}) e^{i(\mathbf{k}\mathbf{r} - \omega t)} d\omega d\mathbf{k} \quad (3)$$

For the other quantities we use the same notation as well. Then we obtain

$$\mathbf{D}(\omega, k) = -\frac{c}{\omega} (\mathbf{k} \times \mathbf{B}(\omega, \mathbf{k})), \quad (4)$$

$$\mathbf{k} \cdot \mathbf{D}(\omega, k) = 0, \quad (5)$$

$$\mathbf{B}(\omega, \mathbf{k}) = \frac{c}{\omega} (\mathbf{k} \times \mathbf{E}(\omega, \mathbf{k})), \quad (6)$$

$$\mathbf{k} \cdot \mathbf{B}(\omega, k) = 0, \quad (7)$$

and

$$D_i(\omega, \mathbf{k}) = \varepsilon_{ij}(\omega, \mathbf{k}) E_j(\omega, \mathbf{k}), \quad (8)$$

where summation over the index j , appearing twice, is assumed.

We can also introduce the inverse dielectric function matrix $\varepsilon_{ij}^{-1}(\omega, \mathbf{k})$ and write

$$E_i(\omega, \mathbf{k}) = \varepsilon_{ij}^{-1}(\omega, \mathbf{k}) D_j(\omega, \mathbf{k}). \quad (9)$$

Writing eqs. (4)-(7) we have used the fact that in our case external current and charges are absent. The spatial dispersion is determined by the parameter $a\mathbf{k}$ or by the somewhat more descriptive parameter a/λ , where a is a characteristic dimension (the radius of "the region of influence", radius of molecular action, etc). and λ is the length of the electromagnetic wave. In a condensed non-metallic medium the radius a is about the order of the lattice constant. Therefore, the parameter a/λ is very small, even in the optical or ultraviolet range of the electromagnetic spectrum.

Eliminating the magnetic induction \mathbf{B} from eqs. (4) - (7) we obtain the expression

$$\mathbf{D} = -\frac{c^2}{\omega^2} [\mathbf{k} \times (\mathbf{k} \times \mathbf{E})] = \frac{c^2}{\omega^2} [k^2 \mathbf{E} - \mathbf{k} \cdot (\mathbf{k} \cdot \mathbf{E})], \quad (10)$$

and, substituting eq. (8) into (10), we find

$$\left[\frac{\omega^2}{c^2} \varepsilon_{ij}(\omega, \mathbf{k}) - k^2 \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) \right] E_j(\omega, \mathbf{k}) = 0. \quad (11)$$

If we use the matrix $\varepsilon_{ij}^{-1}(\omega, \mathbf{k})$ we obtain

$$\left[\frac{\omega^2}{c^2} \delta_{ij} - k^2 \left(\delta_{il} - \frac{k_i k_l}{k^2} \right) \varepsilon_{lj}^{-1}(\omega, \mathbf{k}) \right] D_j(\omega, \mathbf{k}) = 0. \quad (12)$$

These homogeneous systems of algebraic equations have nontrivial solutions $\mathbf{E}(\omega, k) \neq 0$ and $\mathbf{D}(\omega, k) \neq 0$, only if the corresponding determinants vanish

$$\left| \frac{\omega^2}{c^2} \varepsilon_{ij}(\omega, \mathbf{k}) - k^2 \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) \right| = 0, \quad (13)$$

and

$$\left| \frac{\omega^2}{c^2} \delta_{ij} - k^2 \left(\delta_{il} - \frac{k_i k_l}{k^2} \right) \varepsilon_{lj}^{-1}(\omega, \mathbf{k}) \right| = 0. \quad (14)$$

The dispersion equations (13) and (14) give the relation between ω and \mathbf{k} for the electromagnetic normal waves (or eigenwaves) in a given medium for $\omega_l = \omega_l(\mathbf{k})$, where the subscript l corresponds to the given normal wave. For these normal waves, we can write the wave vector \mathbf{k} in the form

$$\mathbf{k} = \frac{\omega}{c} n(\omega, \mathbf{s}) \mathbf{s}, \quad (15)$$

where \mathbf{s} is the unit vector in the direction of \mathbf{k} and $n(\omega, \mathbf{s})$ is the corresponding refraction index.

The dispersion equation (13) can be conveniently written in the form

$$\left| n^2(\omega, \mathbf{s}) (\delta_{ij} - s_i s_j) - \varepsilon_{ij} \left(\omega, \frac{\omega}{c} n(\omega, \mathbf{s}) \mathbf{s} \right) \right| = \varepsilon_{ij} s_i s_j n^4 - [(\varepsilon_{ij} s_i s_j) \varepsilon_{il} - \varepsilon_{il} \varepsilon_{lj} s_i s_j] n^2 + |\varepsilon_{ij}| = 0. \quad (16)$$

This is the fundamental equation of crystal optics. In classical crystal optics $\varepsilon_{ij} = \varepsilon_{ij}(\omega)$ and (16)

becomes quadratic with respect to n^2 . This reduced form is frequently called Fresnel's equation.

For our purpose, it is more convenient to investigate the dispersion equation (14) because we can use the property that the electric induction \mathbf{D} is always transverse for normal waves. It means that we can choose a coordinate system whose z-axis is directed along \mathbf{s} and then the vector \mathbf{D} will have only two components D_x and D_y . By setting $s_1 = s_2 = 0$ and $s_3 = 1$ the wave equation and the dispersion equation have the following form

$$m^2 D_\alpha = \varepsilon_{\alpha\beta}^{-1} D_\beta, \quad (17)$$

$$\left| m^2 \delta_{\alpha\beta} - \varepsilon_{\alpha\beta}^{-1} \right| = m^4 - (\varepsilon_{11}^{-1} + \varepsilon_{22}^{-1}) m^2 + \varepsilon_{11}^{-1} \varepsilon_{22}^{-1} - (\varepsilon_{12}^{-1})^2 = 0 \quad (18)$$

where we introduced the notation

$$\frac{1}{n^2} = m^2; \quad \alpha, \beta = 1, 2. \quad (19)$$

The dispersion equation (18) has two roots for the quantity m^2

$$m_{1,2}^2 = \frac{\varepsilon_{11}^{-1} + \varepsilon_{22}^{-1}}{2} \pm \frac{1}{2} \sqrt{(\varepsilon_{11}^{-1} - \varepsilon_{22}^{-1})^2 - 4(\varepsilon_{12}^{-1})^2}, \quad (20)$$

and, consequently, we are led to two mutually orthogonal vectors \mathbf{D}_1 and \mathbf{D}_2 . As it is well-known^{4,7)}, the existence of two values of the refractive index n for a given direction of the wave vector \mathbf{k} is

the origin of the birefringence effect. The dispersion equation (18) has a multiple root $m_1^2 = m_2^2$ if

$$(\varepsilon_{11}^{-1} - \varepsilon_{22}^{-1})^2 - 4(\varepsilon_{12}^{-1})^2 = 0 \quad (21)$$

and the birefringence is absent in this case. Also it is well-known^{4,7)} that in the case of traditional crystal optics, i.e. for $\varepsilon_{ij} = \varepsilon_{ij}(\omega)$, the multiple root for m exists for every direction of \mathbf{k} only in the case of cubic crystals. For all other crystals with a lower symmetry, birefringence is absent only for

the wave with the wave vector \mathbf{k} oriented along the principal optical axis. For hexagonal, tetragonal or trigonal crystals there is one such an axis and these crystals are called uniaxial. For the three remaining crystal systems, namely the orthorhombic, monoclinic and triclinic ones, there are two privileged normal wave directions for which there is no birefringence and the crystals are called biaxial. As shown in what follows, the situation is more complicated when the spatial dispersion is taken into account, i.e. if the dielectric function matrix depends on both the frequency ω and the wave vector \mathbf{k} .

3. SPATIAL DISPERSION OF CUBIC CRYSTALS

We first write the inverse dielectric matrix for cubic crystals in the case of small spatial dispersion in the form

$$\varepsilon_{ij}^{-1}(\omega, \mathbf{k}) = \varepsilon_{ij}^{-1}(\omega)\delta_{ij} + \beta_{ijlm}(\omega)k_l k_m, \quad (22)$$

or as

$$\varepsilon_{ij}^{-1}(\omega, \mathbf{k}) = \varepsilon_{ij}^{-1}(\omega)\delta_{ij} + \beta_{ijlm}(\omega)\frac{\omega^2}{c^2}n^2 s_l s_m. \quad (23)$$

The fourth-rank tensor β_{ijlm} has only three independent and non-zero components for cubic crystals with symmetry classes O , T_d and O_h ^{4, 8)}. These are

$$\beta_1 = \beta_{xxxx} = \beta_{yyyy} = \beta_{zzzz}, \quad (24)$$

$$\beta_2 = \beta_{xxzz} = \beta_{yyxx} = \beta_{zzyy} = \beta_{zzxx} = \beta_{xxyy} = \beta_{yyzz}, \quad (25)$$

$$\beta_3 = \beta_{xyxy} = \beta_{yzyz} = \beta_{zxzx}. \quad (26)$$

Using these expressions the matrix elements of ε_{ij}^{-1} can be written as

$$\varepsilon_{xx}^{-1} = \varepsilon^{-1}(\omega) + \left(\frac{\omega}{c}n\right)^2 [\beta_1 s_x^2 + \beta_2 (s_y^2 + s_z^2)], \quad (27)$$

$$\epsilon_{xy}^{-1} = 2 \left(\frac{\omega}{c} n \right)^2 \beta_3 s_x s_y, \quad (28)$$

$$\epsilon_{yy}^{-1} = \epsilon^{-1}(\omega) + \left(\frac{\omega}{c} n \right)^2 \left[\beta_2 (s_x^2 + s_z^2) + \beta_1 s_y^2 \right], \quad (29)$$

$$\epsilon_{xz}^{-1} = 2 \left(\frac{\omega}{c} \tilde{n} \right)^2 \beta_3 s_x s_z, \quad (30)$$

$$\epsilon_{zz}^{-1} = \epsilon^{-1}(\omega) + \left(\frac{\omega}{c} n \right)^2 \left[\beta_2 (s_x^2 + s_y^2) + \beta_1 s_z^2 \right], \quad (31)$$

$$\epsilon_{yz}^{-1} = 2 \left(\frac{\omega}{c} n \right)^2 \beta_3 s_y s_z. \quad (32)$$

The factor two in the expressions for ϵ_{xy}^{-1} , ϵ_{xz}^{-1} and ϵ_{yz}^{-1} is due to the summation in (22). The equations (22) – (32) allow us to write the expression for the inverse dielectric matrix in the form

$$\epsilon_{ij}^{-1}(\omega, \frac{\omega}{c} n \mathbf{s}) = \left(\epsilon^{-1}(\omega) + \frac{\omega^2}{c^2} n^2 \beta_2 \right) \delta_{ij} + \frac{\omega^2}{c^2} n^2 \tilde{\beta} s_i^2 \delta_{ij} + 2 \frac{\omega^2}{c^2} n^2 \beta_3 s_i s_j, \quad (33)$$

where

$$\tilde{\beta} = \beta_1 - \beta_2 - 2\beta_3. \quad (34)$$

The first term in the expression (33) is the isotropic contribution, the second one is anisotropic, but it is expressed in terms of a diagonal matrix; the last term is purely longitudinal. If we rewrite the wave equation (14)

$$\left[\frac{\delta_{ij}}{n^2(\omega, \mathbf{s})} - (\delta_{il} - s_l s_i) \epsilon_{ij}^{-1}(\omega, \frac{\omega}{c} n \mathbf{s}) \right] D_j(\omega, \frac{\omega}{c} n \mathbf{s}) = 0, \quad (35)$$

it follows after some simple algebra that the longitudinal part of ϵ_{ij}^{-1} disappears from this equation due to the prefactor $(\delta_{il} - s_l s_i)$, because the multiplication of this prefactor with the longitudinal part yields zero.

We will solve eq. (35) with the aid of perturbation theory and rewrite this equation as

$$(L_0 + L_1)_{ij} D_j = \rho(\omega, \mathbf{s}) \delta_{ij} D_j \quad (36)$$

and

$$\rho(\omega, \mathbf{s}) = \rho_0(\omega) + \rho_1(\omega, \mathbf{s}), \quad (37)$$

where

$$\rho(\omega, \mathbf{s}) = \frac{1}{n^2(\omega, \mathbf{s})} \quad (38)$$

is the eigenvalue of this equation which should be calculated up to the first order perturbations described by the perturbation operator L_1 . L_1 has the form

$$(L_1)_{ij} = \frac{\omega^2}{c^2} n^2(\omega, \mathbf{s}) \tilde{\beta} s_i^2 \delta_{ij} - \frac{\omega^2}{c^2} n^2(\omega, \mathbf{s}) \tilde{\beta} s_i s_j^3. \quad (39)$$

In the framework of perturbation theory, we can change the value $n^2(\omega, \mathbf{s})$ in eq. (39) into $n_0^2(\omega)$ defined by the zero-order approximation. This zero-order perturbation is defined by the equation

$$(L_0)_{ij} D_j^0 = \rho_0(\omega) D_i^0, \text{ where} \quad (40)$$

$$(L_0)_{ij} = \left(\varepsilon^{-1}(\omega) + \frac{\omega^2}{c^2} n_0^2 \beta_2 \right) (\delta_{ij} - s_i s_j) \quad (41)$$

and, consequently, eq. (40) for the zero-order approximation can be written as

$$\left[\left(\varepsilon^{-1}(\omega) + \frac{\omega^2}{c^2} n_0^2 \beta_2 \right) (\delta_{ij} - s_i s_j) \right] D_j^0 = \rho_0(\omega) D_i^0. \quad (42)$$

By using eq. (5) this equation reduces to the wave equation for isotropic media, which has the multiple root for $\rho_0(\omega)$ equal to

$$\rho_0(\omega) = \frac{1}{n_0^2(\omega)} = \varepsilon^{-1}(\omega) + \frac{\omega^2}{c^2} n_0^2(\omega) \beta_2. \quad (43)$$

Moreover, in the framework of macroscopic electrodynamics we can neglect the second term in the right part of (43) because there are no possible experiments which can help us to distinguish the terms $\varepsilon^{-1}(\omega)$ and $\frac{\omega^2}{c^2}n_0^2(\omega)\beta_2$. It means that for the refraction index $n_0(\omega)$ in the zero-order approximation we can write

$$n_0^2(\omega) = \varepsilon(\omega), \quad (44)$$

where

$$\varepsilon(\omega) = \left(\varepsilon^{-1}(\omega)\right)^{-1}. \quad (45)$$

The existence of the multiple root of $\rho_0(\omega)$ in the zero-order approximation tells us that the system of equations (36) and (37) has a degenerate kernel⁹⁾ and we should use for the calculation of the first order correction to $\rho_1(\omega, \mathbf{s})$ a so-called secular equation. This equation will give also the two correct values of $\tilde{D}_i^0(\omega, n(\omega, \mathbf{s}))$ resulting from the removing of the degeneracy due to the perturbation operator L_1 . If we write the first order correction $\rho_1(\omega, \mathbf{s})$ as

$$\rho_1(\omega, \mathbf{s}) = \frac{\omega^2}{c^2} n_0^2(\omega) \tilde{\beta} \tilde{\rho}_1(\mathbf{s}), \quad (46)$$

the secular equation will have the form

$$\left(s_i^2 \delta_{ij} - s_i s_j^3\right) e_j(\mathbf{s}) = \tilde{\rho}_1(\mathbf{s}) e_i(\mathbf{s}), \quad (47)$$

where we introduced the unit vector \mathbf{e} in the direction of $\tilde{\mathbf{D}}^0(\mathbf{s})$. The pair of eigenmodes $(e_1(\mathbf{s}), e_2(\mathbf{s}))$ has been chosen in such a way that in this basis the 2×2 matrix associated to the operator L_1 becomes diagonal. Because the vector $\tilde{\mathbf{D}}^0(\mathbf{s})$ is orthogonal to the vector \mathbf{s} , when the matrix is computed, the contribution of the terms $s_i s_j^3$ vanishes and eq. (47) can be replaced by

$$s_i^2 e_i(\mathbf{s}) = \tilde{\rho}_1(\mathbf{s}) e_i(\mathbf{s}). \quad (48)$$

Because of its complexity, the computation has been done by using computer algebra software. The two solutions for $\tilde{\rho}_1(\mathbf{s})$ i.e. the diagonal element of the matrix mentioned above turn out to be the roots of the quadratic equation

$$\tilde{\rho}_1^2(\mathbf{s}) - 2(s_x^2 s_y^2 + s_x^2 s_z^2 + s_y^2 s_z^2) \tilde{\rho}_1(\mathbf{s}) + 3s_x^2 s_y^2 s_z^2 = 0. \quad (49)$$

These solutions are

$$\tilde{\rho}_1^{(1,2)}(\mathbf{s}) = (s_x^2 s_y^2 + s_x^2 s_z^2 + s_y^2 s_z^2) \pm \sqrt{(s_x^2 s_y^2 + s_x^2 s_z^2 + s_y^2 s_z^2)^2 - 3s_x^2 s_y^2 s_z^2}. \quad (50)$$

Note that the equation for $\tilde{\rho}_1(\mathbf{s})$ has only one solution for the seven directions of the propagation vector \mathbf{s} , mentioned above i. e. the three main crystallographic axes and the four body diagonals of the cube.

We can also obtain the expression for the eigenmodes of eq. (48). The components of the eigenmode $e_2(\mathbf{s})$, which correspond to the eigenvalue $\tilde{\rho}_1^{(2)}(\mathbf{s})$ the one with minus sign in front of square roote have the following form

$$e_{2z}(\mathbf{s}) = \frac{1}{\sqrt{2}} \sqrt{s_x^2 + s_y^2 + \frac{s_x^2 s_y^2 (s_x^2 + s_y^2) - (s_x^4 + s_y^4) s_z^2}{r}}, \quad (51)$$

$$e_{2x}(\mathbf{s}) = e_{2z}(\mathbf{s}) \frac{s_y^2 (s_z^2 - s_x^2) + r}{s_x s_z (s_x^2 - s_y^2)}, \quad (52)$$

$$e_{2y}(\mathbf{s}) = e_{2z}(\mathbf{s}) \frac{s_x^2 (s_z^2 - s_y^2) + r}{s_y s_z (s_y^2 - s_x^2)}, \quad (53)$$

where r is given by

$$r = \sqrt{(s_x^2 s_y^2 + s_x^2 s_z^2 + s_y^2 s_z^2)^2 - 3s_x^2 s_y^2 s_z^2}. \quad (54)$$

For certain directions \mathbf{s} zero denominators appear and then these expressions cannot be used directly. When eqs. (51)-(54) must be used for a direction $\tilde{\rho}_1$ which leads to zero denominators using a

non-singular direction very close to the singular one gives sufficient accuracy for practical purposes. Analytical expressions can also be derived for these special cases. For instance for $s_x > 0$ and $s_y > 0$, but s_x very small we obtain

$$e_{2x}(\mathbf{s}) = -\frac{s_z}{|s_z|}, \quad e_{2y}(\mathbf{s}) = \frac{s_x s_z}{2|s_z|\sqrt{1-s_z^2}}, \quad e_{2z}(\mathbf{s}) = \frac{s_x}{2|s_z|}. \quad (55)$$

The components of the eigenmode $e_1(\mathbf{s})$ corresponding to the eigenvalue $\tilde{\rho}_1^{(1)}(\mathbf{s})$ can be obtained from the equation

$$e_1(\mathbf{s}) = e_2(\mathbf{s}) \times \mathbf{s}. \quad (56)$$

Note that our first order correction to the eigenvalue $\tilde{\rho}_1(\mathbf{s})$ coincides with that obtained by Burnett et al ⁶⁾. An alternative approach that leads to relations that are equivalent to eqs. (51-56) is discussed in ref. 10.

If we define the variation of the birefringence effect with propagation direction according to eq. (22) from ref. 6 as

$$\Delta\tilde{\rho}_1(\mathbf{s}) = \tilde{\rho}_1^{(1)}(\mathbf{s}) - \tilde{\rho}_1^{(2)}(\mathbf{s}) = 2\sqrt{(s_x^2 s_y^2 + s_x^2 s_z^2 + s_y^2 s_z^2)^2 - 3s_x^2 s_y^2 s_z^2}, \quad (57)$$

this value will have the same sign for every direction. The origin for a possible change of the sign of the birefringence variation shown in Fig. 6 of ref. 6 is related to the behavior of the eigenmodes $e_{1,2}(\mathbf{s})$. This fact can be easily understood from Fig. 1.

Here we show the behavior of the eigenmodes $e_{1,2}(\mathbf{s})$ for two propagation directions in the diagonal plane $[\bar{1}10]$ of the cubic cell containing the $[001]$, $[111]$ and $[110]$ directions. We can see that the eigenmode $e_1(\mathbf{s})$ corresponding to the larger value of $\tilde{\rho}_1(\mathbf{s})$ lies in the plane $[\bar{1}10]$ for the directions above the axis $[111]$. The eigenmode $e_2(\mathbf{s})$ is normal to this plane. The situation is

reversed for the directions below the axis $[111]$. Here the eigenmode $e'_2(\mathbf{s}')$ lies in the plane $[\bar{1}10]$. It means that, if we consider the difference of the refraction indices between one wave polarized in the $[\bar{1}10]$ plane and on other normal to this plane, this difference will have opposite sign for directions above and below the axis $[111]$.

4. CONSEQUENCES FOR OPTICAL SYSTEM DESIGN

In optical system design, BIRD leads to the appearance of multiple polarized rays during refraction. There are two basic effects depending on crystal orientation, wavelength, thickness and shape of the lens. The first BIRD consequence is the optical path-length difference between the two orthogonally polarized components corresponding to a ray. This path-length difference, resulting in a phase retardation, can be visualized in pupil maps for arbitrary field points. An example is shown in Table I for a one-lens case.

The second consequence of BIRD is an angular difference in ray paths. In an optical system with N components, for each incident ray on the first surface there are 2^N outgoing rays at the image plane. The bifurcation of the rays causes an angular difference in ray path between ordinary and extraordinary rays after each refraction, which results in a ray deviation at the image plane. Thus instead of one ray there is a full cone of outgoing rays and the opening angle of this cone can be called the maximal angular deviation. The compensation of this effect requires additional consideration and will not be considered in this paper.

The first issue in optical system design with BIRD is the calculation of the effect itself. Only one commercial optical design program (Code V, © Optical Research Associates) supports the calculation of the effect within a given approximation and allows to analyze the image quality

taking into account the phase retardation caused by BISD. (In the present version of this program the ray bifurcation and its consequences on energy transmission are not included).

The second important design issue is the calculation speed. For an adequate estimation of BISD it is necessary to trace through the system at least 50 pupil rays taking into account their polarization properties (see Table II). For a lithographic lens with about 50 surfaces this requirement drastically increases the time of image quality estimation.

The third issue is to identify a relevant characteristic measure of the image quality for BISD in DUV systems. In optical design it is very convenient to define the influence of BISD on the image quality by one number. We found that the standard deviation of the retardation (phase difference) over the pupil is a good indicator for the BISD influence. This quantity shows a very good correlation with the image quality loss due to the phase retardation. This is confirmed by the calculation of BISD in optical systems with and without BISD compensation. An example for a system with $NA=0.75$ is given in Table III. The value of the Strehl Ratio is computed by taking into account the polarization of light and in the case of $NA=0.75$ this value cannot exceed 0.84844 even for an ideal optical system. By computing the Strehl Ratio values with and without BISD we observe that the standard deviation of the retardation adequately indicates the loss of image quality caused by BISD. Therefore by taking the standard deviation of the retardation into the error function we can optimize optical systems for improving the imaging quality.

A major obstacle for the BISD compensation is the possible effect of asymmetry of the pupil map around center of the pupil. For preserving the symmetry only the three directions $\{001\}$, $\{110\}$ and $\{111\}$ can be selected as an optical axis. The directions $\{001\}$ and $\{110\}$ have an advantage as compared to $\{111\}$ because around them the BISD effect hardly changes when ray parameters change, but along $\{110\}$ we have a maximum of the effect. From the point of view of crystal

manufacturing the direction $\{111\}$ is preferable because for this direction the residual stresses are minimal. Generally it is possible to select an arbitrary direction $\{abc\}$ as an optical axis but in that case both the BIRD compensation and the technological issues may be much more difficult.

The general strategy of the effect compensation is the following. It can be observed from Table I that, if chosen as an optical axis, the direction $\{100\}$ has a 4-fold (90°) angular symmetry and $\{111\}$ has 3-fold (120°) angular symmetry. It is thus possible to tune the individual components to achieve an almost circular distribution of the retardation over the pupil. For this purpose, a combination of lenses with orientations $\{001\}$ - 0° and $\{001\}$ - 45° or $\{111\}$ - 0° and $\{111\}$ - 60° can be used. The angles indicate a rotation around the axis with respect to the drawings in Table I. The BIRD compensation for off-axis points is more difficult, because the distribution of the birefringence is not symmetric, but in general the behavior for such a field point is comparable to that of the axial point. In all cases studied by us the compensation for the axial point helps to the compensation for the off-axis points. For off-axis points it would be useful to have software support for the arbitrary choice of the crystal orientation during optimization.

The known approach to BIRD compensation consist of using combinations of optical components with the optical axes along the crystal orientations $\{001\}$ and $\{111\}$, which can be oriented at different angles around the optical axis. A certain choice of the crystal orientations of the components in the optical system allows to compensate significantly the total phase retardation, which can then be further adjusted or nulled out with the available optimization tools.

The novelty of our approach is the compensation of the phase retardation with crystal orientation $\{011\}$ only. In this case we use the opposite orientation of the phase retardation for components with crystal orientations $\{011\}$ - 0° and $\{011\}$ - 90° . Such an arrangement allows to compensate the effect without or with minor additional optimization and does not break the geometrical aberration

correction. An optical system compensated according to this approach is shown in Fig 2. The values of Strehl Ratio for the system with and without compensation are given in Table III.

5. CONCLUSIONS

A simplified mathematical description of BIRD has been given in a form that is suitable for optical design. Analytic expressions for the BIRD eigenmodes have been obtained.

The consequences of BIRD in the optical design for DUV lithography corresponding have been considered. A relevant indicator of the effect of BIRD on image quality and an improved BIRD correction approach has been discussed. An example of a high NA lithographic projection system with a high degree of BIRD correction has been presented.

ACKNOWLEDGEMENTS

The first author gratefully acknowledges the support of his research by ASM Lithography. We also sincerely thank Vitaly Ginzburg and John Burnett for helpful discussions.

REFERENCES

- 1) H.A. Lorentz: *Collected papers* (Nijhoff, Den Haag, 1936) Vol. 2-3.
- 2) K.H. Hellwege: *Z. Phys.* **129** (1951) 626.
- 3) V.L. Ginzburg: *Sov. Phys.-JETP* **7** (1958) 1096.
- 4) V. M. Agranovich, V. L. Ginzburg: *Crystal Optics with Spatial Dispersion, and Excitons* (Springer-Verlag, Berlin, 1984).
- 5) J. H. Burnett, Z. H. Levine and E. L. Shirley: *Phys. Review B* **64** (2000) 241102;
- 6) J. H. Burnett, Z. H. Levine, E. L. Shirley and J. H. Bruning: *Journal of Microlithography, Microfabrication and Microsystems* **1** (2002) 213.
- 7) L.D. Landau: *Electrodynamics of continuous media* (Pergamon Press, Oxford, 1984).
- 8) J.F. Nye: *Physical properties of crystals* (Clarendon Press, Oxford, 1985).
- 9) R. Courant: *Methods of mathematical physics* (Interscience, New York, 1989) Vol. 1.
- 10) A. E. Rosenbluth, G. Gallatin, N. Seong, O. Dittmann, M. Totzeck: *Image Formation In A Lens With Spatial Dispersion (Intrinsic Birefringence)* presented at 4th International Symposium on 157 nm Lithography, Yokohama, Japan (2003).
- 11) T. Takahashi, J. Nishikawa, Y. Omura: U.S. Patent Application 20020044260

Figure captions

Fig. 1. The behavior of the eigenvectors in the $[\bar{1}\bar{1}0]$ -diagonal plane of the elementary cubic cell.

Fig 2. BIRD compensation in a DUV lithographic system. This optical system has a working wavelength of 157 nm, NA=0.75, image size 22×22 mm¹¹⁾

TABLES

Table I. Directional dependence of the BISD of a single lens for the axial object point. The symmetry axis of the lens is oriented along the crystal direction given in the first column. The dark tone indicates regions with low phase retardation and light tone indicates areas with high phase retardation

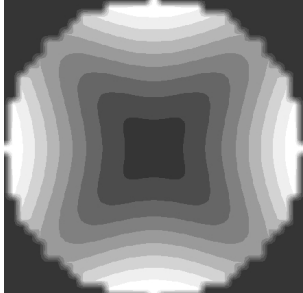
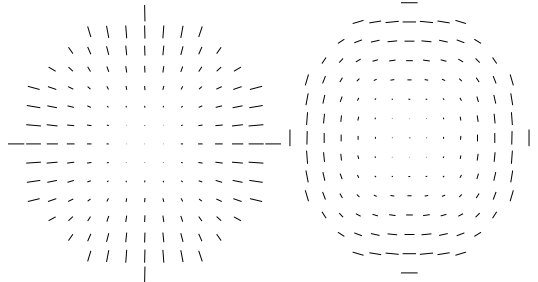
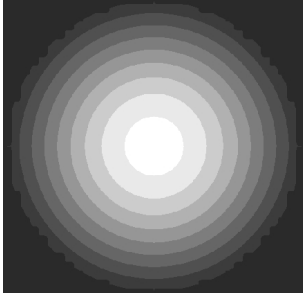
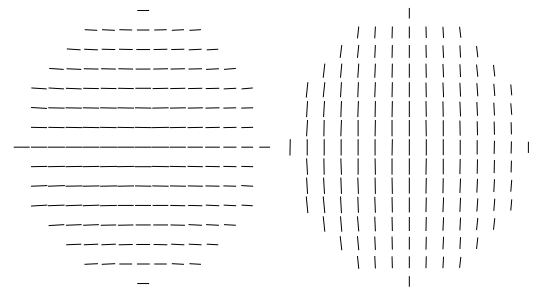
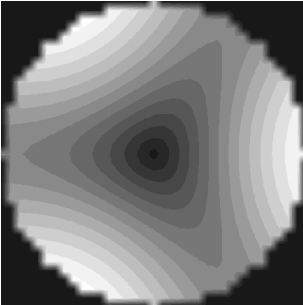
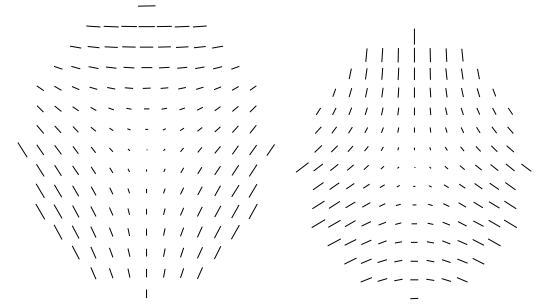
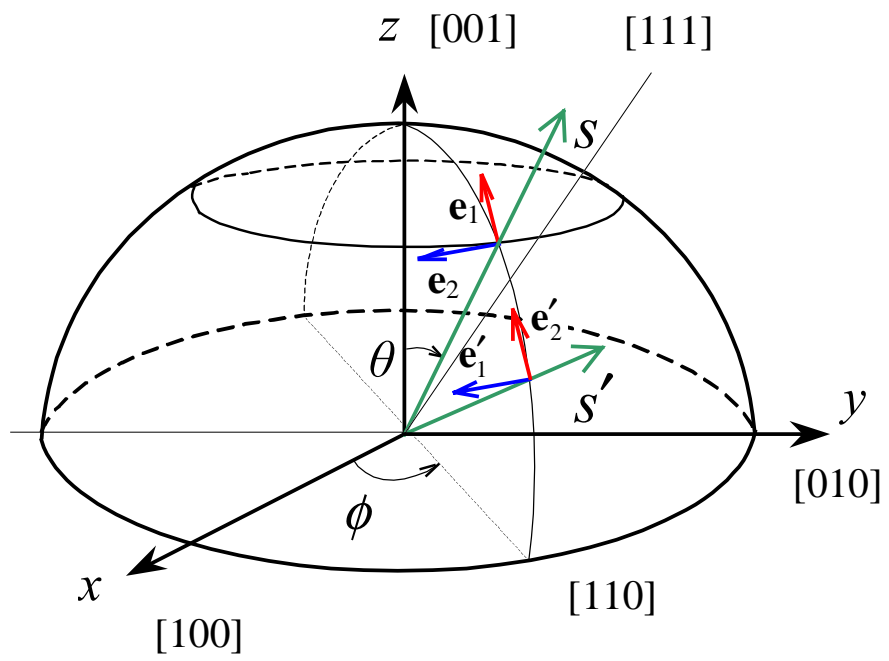
Chosen optical axis	The phase retardation dependence over the pupil	The projections of the two eigenvector states on the pupil plane
{100}		
{110}		
{111}		

Table II. Example shows the dependence of the accuracy of BISD calculation on the number of rays

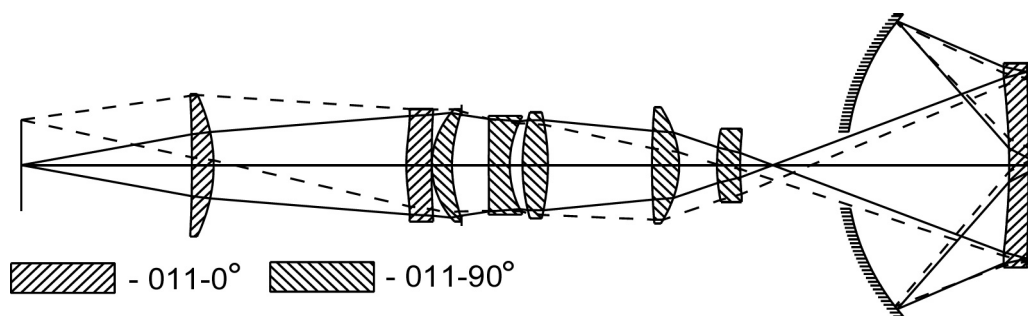
	Number of rays	5	13	24	49	89	481	973
Low BISD	Mean retardation	154.32	149.68	151.17	148.95	151.54	151.04	151.11
	standard deviation of retardation	11.174	10.717	12.075	12.235	13.546	14.136	14.103
High BISD	Mean retardation	28.176	86.253	76.489	88.437	82.043	85.119	84.729
	standard deviation of retardation	35.734	59.455	41.456	49.023	51.595	49.944	49.791

Table III. Image quality in an optical system with and without BISD compensation

Field, mm	Strehl Ratio (in the absence of BISD)	BISD uncompensated			BISD compensation		
		Strehl Ratio	mean retardation, degrees	standard deviation of retardation	Strehl Ratio	mean retardation, degrees	standard deviation of retardation
0.0	0.811952	0.25678	108.241	40.2502	0.79837	46.2033	4.20398
5.0	0.841904	0.26579	108.028	40.3193	0.80797	44.2758	6.10463
10.0	0.839376	0.27746	107.105	40.5278	0.80280	38.4220	10.7555
14.0	0.836742	0.28396	105.950	40.6184	0.78340	31.5285	15.2787
16.4	0.811973	0.31566	105.552	40.5935	0.78448	27.3252	18.1214



Serebriakov, Maksimov, Bociort, Braat Fig. 1.



Serebriakov, Maksimov, Bociort, Braat Fig. 2.