Toward a Consistent and Accurate Approach to Modeling Projection Optics

Danping Peng*, Peter Hu, Vikram Tolani, Thuc Dam

Luminescent Technologies, 2471 East Bayshore Road Suite 600, Palo Alto, CA 94303

Jacek Tyminski, Steve Slonaker

Nikon Precision Inc, 1399 Shoreway Road, Melmont, CA 94002

ABSTRACT

This paper presents a consistent and modularized approach to modeling projection optics. Vector nature of light and polarization effect are considered from the very beginning at source, through mask and projection lens down into film stack. High-NA and immersion effect are also included. Of particular interest is the formulation of mask diffraction model that is pluggable, i.e. it can be a thin-mask model, a empirical approximate mask model, or rigorous mask 3D solver. We demonstrate that under Kirchoff thin-mask assumption our formulation is the same as Smythe formula. A compact film-stack model is formulated. The formulation is first presented in Abbe's source integration approach and then reformulated in Hopkins' TCC approach which allows for a SVD decomposition and are computationally more efficient for fixed optical setting.

Keywords: Abbe, Hopkins, Maxwell equation, vector optics, scalar optics, polarization, partial coherent, lithography simulation

1. INTRODUCTION

Lithography simulation has been indispensable to the research and development in semiconductor industry in its insatiable drive toward printing ever smaller features with increasingly large designs. This trend has be well-captured by Moore's law.

The traditional simulator was built on the assumption of coherent illumination with low NA, and para-axial approximation is used in the modeling of optics and resist. In the past two decades, in anticipation to the introduction of new exposure and inspection tools, lithography simulator has been extended in several directions: from coherent illumination to off-axis illumination with different source shapes; from low NA to high-NA for dry lithography, to hyper-NA for immersion lithography; from scalar optics to vector optics that include the polarization effect at source, mask, lens and resist film stack; from a simple film-stack model that is only valid for normal incident light to one that support the high-incident angle and vector nature of light.

The theoretically development and commercial implementation of these features do not happen over night. It took decades of intensive work in academia and industry; In the process, new features were added over the existing features, not always in a consistent formulation. It seem to the author that there has been no systematic exposition of these developments in a systematic and consistent formulation.

The optical community has benefited tremendously from the classics by Born and Wolf,[?] now in its 7-th edition, and Goodman.[?] But they are mostly focus on scalar imaging theory and uses paraxial approximation, and thus are only valid for low NA optics. In his 1988 paper[?] Yeung proposed a vector imaging and film-stack model that include high-NA effect. In a later paper[?] Yeung simplify his original approach and proposed the Hopkins formulation that include thin-film effect. In his PhD thesis[?] Flagello proposed a thin-film model in matrix format that is easier to compute. Flagello and Rosenbluth[?,?] considerable simplified Yeung's approach for vector optics and film-stack effect. Cole and his co-authors[?,?] formulated the radio-metric correction term for high-NA optics. See also the work of Gallatin[?] on scalar high-NA imaging.

In this paper, we present the above developments in projection optics in a uniform and consistent formulation with some further simplification. The underlying theme of our approach will be using matrix instead of scalars

*Contact: Danping Peng, dpeng@luminescent.com

throughout the whole process to track the polarization state as light propagate through the optical components of projection instrument. We take a critical view on the underlying assumptions of various modeling approaches and comments on their validity for the current exposure tools. In section **??**, we define the polarization vector and coherence matrix at a single point source. In section **??**, we present the vector diffraction model from mask near-field to entrance pupil and from entrance pupil to wafer plane. In section **??**, we present the mask model first in a general setting, and then present some simplifying assumptions, and take a critical view on the validity of Hopkins assumption and thin-mask model. Our imaging formulation allows mask models from the simplest to the full rigorous 3D EMF solution to be readily plugin. In section **??**, we present the projection lens model. We present the ideal lens operation, derive the radio-metric correction term, and present how to add aberration term and Jones pupil. We then present the image formation equation, and briefly mentioned how to include a film-stack model. Compare with the previous approach, our formulation of film-stack model use a 3x2 matrix that is more compact. We then put all the pieces together, thus conclude the Abbe's approach to projection optics. We briefly comment on how to reformulate the Abbe's approach into Hopkins' formulation.

All results in the paper has been presented in published sources in one form or another by various authors. We simply compile the results scattered in many sources by dozens of authors in a consistent formulation.

A few comments on our choice of notation. The variables on the source side will have either a subscript or sup-script s, as in (α_s, β_s); mask side variables does not use subscript or sup-script, as in (α, β); image side variables use sup-script $'$, as in (α', β').

We use two coordinate systems, the right-handed (x, y, z) system and the intrinsic local (s, p) system which are dependent on unit wave-vector $\hat{k}$. Whenever possible, we use the (s, p) system as the formula is simplest in local intrinsic system.

We will follow the Matlab convention in denoting $\left[\alpha; \beta; \gamma\right]$ as column-vector while $\left[\alpha, \beta, \gamma\right]$ as row-vector.

2. POLARIZATION VECTOR AND COHERENT MATRIX

Consider a monochromatic plane wave propagating in the direction $\hat{k}$, where $\hat{k} = \left[\alpha; \beta; \gamma\right]$ is the direction cosine with $\gamma = \sqrt{1 - \alpha^2 - \beta^2}$. We define the TE- or s-direction, and TM- or p-direction, as

$$\hat{e}_\perp = \frac{\hat{z} \times \hat{k}}{|\hat{z} \times \hat{k}|} = \begin{bmatrix} \frac{-\beta}{\rho} \\ \frac{\alpha}{\rho} \\ 0 \end{bmatrix}, \hat{e}_\parallel = \hat{k} \times \hat{e}_\perp = \begin{bmatrix} \frac{-\alpha\gamma}{\rho} \\ \frac{-\beta\gamma}{\rho} \\ \rho \end{bmatrix}, \tag{1}$$

where $\rho = \sqrt{\alpha^2 + \beta^2}$, such that $[\hat{e}_\perp; \hat{e}_\parallel; \hat{k}]$ is right-handed. In the special case when the plane wave propagate along the z-axis, that is when $\alpha = \beta = 0$, we agree that $\hat{e}_\perp = \left[1; 0; 0\right]$, and $\hat{e}_\parallel = \left[0; 1; 0\right]$.

Since electro-magnetic wave is transverse, an arbitrary E-field can be represented as

$$\vec{E} = E_\perp \hat{e}_\perp + E_\parallel \hat{e}_\parallel = \left[\hat{e}_\perp, \hat{e}_\parallel\right] \begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix}, \tag{2}$$

where $E_\perp$ is the electric field in s-state, while $E_\parallel$ in p-state. The vector $\left[E_\perp; E_\parallel\right]$ describe the polarization state of the plane wave. We will call it polarization vector. Over the cause of propagation, the polarization vector will be transformed by different components of the optical system; The modeling of the optical system is accomplished by tracking the changes of polarization vector though the optical apparatus.

At any instance of time, the magnitude and phase of $E_\perp$ and $E_\parallel$ are rapidly changing, and the changes between the two state may be co-related. It is not the field $E_\perp$ and $E_\parallel$ but the average of their intensity and the correlation between them over a microscopically long enough period of time that are detectable by physical devices. So it only makes sense to talk about the average of their intensity and correlation. The coherence matrix is defined as

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \begin{bmatrix} \langle \bar{E}_\perp E_\perp \rangle & \langle \bar{E}_\perp E_\parallel \rangle \\ \langle \bar{E}_\parallel E_\perp \rangle & \langle \bar{E}_\parallel E_\parallel \rangle \end{bmatrix} \tag{3}$$

where $\langle \cdot \rangle$ denote the time average. Coherent matrix completely describe the polarization state of the plane wave.

When tracking the propagation of light through optical systems, we found it easier to use (s,p) systems. But when we need to compute the actual field or intensity, it is convenient to use (x,y,z) coordinate systems. The mapping between the two systems are:

$$\begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} \hat{e_\perp}, \hat{e_\parallel} \end{bmatrix} \begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix} = \begin{bmatrix} \frac{-\beta}{\rho} & \frac{-\alpha\gamma}{\rho} \\ \frac{\alpha}{\rho} & \frac{-\beta\gamma}{\rho} \\ 0 & \rho \end{bmatrix} \begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix} \tag{4}$$

Note that since electromagnetic wave is transverse, we only need to specify the two components E_x and E_y,

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = T \begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix} = \begin{bmatrix} \frac{-\beta}{\rho} & \frac{-\alpha\gamma}{\rho} \\ \frac{\alpha}{\rho} & \frac{-\beta\gamma}{\rho} \end{bmatrix} \begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix}. \tag{5}$$

The third component E_z is determined from

$$E_z = -\frac{1}{\gamma}(\alpha E_x + \beta E_y). \tag{6}$$

Conversely, given E_x and E_y, the (s,p) components can be determined by

$$\begin{bmatrix} E_\perp \\ E_\parallel \end{bmatrix} = T^{-1} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} -\frac{\beta\gamma}{\rho} & \frac{\alpha\gamma}{\rho} \\ -\frac{\alpha}{\rho} & -\frac{\beta}{\rho} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix} \tag{7}$$

We will use the mapping T and T^{-1} frequently below to change between (x,y) and (s,p) coordinate systems. The following representation of T and T^{-1} are quite handy

$$T = \begin{bmatrix} \hat{e}_{\perp,x} & \hat{e}_{\parallel,x} \\ \hat{e}_{\perp,y} & \hat{e}_{\parallel,y} \end{bmatrix}, T^{-1} = \begin{bmatrix} \hat{e}_{\parallel,y} & -\hat{e}_{\parallel,x} \\ -\hat{e}_{\perp,y} & \hat{e}_{\perp,x} \end{bmatrix} \tag{8}$$

For scanner and AIMS tool, the angle of incident and diffracted light with optical axis on mask-side is small, hence $\gamma \approx 1$. Some implementations ignore γ in the above mapping and use the following approximation to T and T^{-1}

$$\tilde{T} = \begin{bmatrix} \frac{-\beta}{\rho} & \frac{-\alpha}{\rho} \\ \frac{\alpha}{\rho} & \frac{-\beta}{\rho} \end{bmatrix}, \tilde{T}^{-1} = \begin{bmatrix} -\frac{\beta}{\rho} & \frac{\alpha}{\rho} \\ -\frac{\alpha}{\rho} & -\frac{\beta}{\rho}. \end{bmatrix} \tag{9}$$

We will see later that keeping the factor γ will lead to the celebrated Smythe-Kirchoff formula, and is more consistent with Maxwell's equation even under the this-mask approximation.

3. DIFFRACTION MODEL

In modeling projection optics, there are two kinds of diffraction that we encounter. One is the diffraction from a flat screen to a spherical surface, the other configuration reverse the former, namely the diffraction from a spherical surface to a flat surface near the focus. The former describes the diffraction from mask near field to entrance pupil of projectors in far-field, and the latter describe the focusing effect of projectors from exit pupil to wafer plane near the focus of projectors. In both set up we assume the radius of the spherical surface is much larger than the wavelength so that we are in Fraunhofer domain. We first present the scalar diffraction formula and then extend them to vector case.

It is well known that the propagation of light in free-space is governed by Helmholtz equation

$$(\Delta + k^2)U = 0, \tag{10}$$

where U is any of the components of the vector of electromagnetic field. The distribution of U on a spherical surface in the far field from a flat screen Σ can be shown to be?

$$U_{\mathrm{diffr}}(\hat{k}) = U_{\mathrm{diffr}}(\alpha, \beta) = \frac{\cos\theta}{j\lambda}\frac{e^{jkr}}{r}\int_{\Sigma} U(x,y)e^{-j2\pi(fx+gy)}\,\mathrm{d}x\mathrm{d}y, \tag{11}$$

where $\hat{k} = [\alpha; \beta; \gamma]$ is the unit vector in the direction of propagation, $\gamma = \cos\theta$ represent the direction cosine of the propagation direction with optical axis, $k = \frac{2\pi n}{\lambda}$ the wave number, r is the radius of the spherical surface, $U(x,y)$ the field distribution over flat screen Σ, and

$$f = \frac{n}{\lambda}\alpha, g = \frac{n}{\lambda}\beta, h = \frac{n}{\lambda}\gamma, \tag{12}$$

represent the spatial frequency.

The diffraction from a spherical surface S to the plane that is vertical to the optical axis and through its focus is described by

$$U_{\mathrm{diffr}}(x,y) = \frac{1}{j\lambda}\frac{e^{-jkr}}{r}\int_{S} U(\alpha,\beta)e^{j2\pi(fx+gy)}\,\mathrm{d}\sigma, \tag{13}$$

where $U(\alpha,\beta)$ is the field distribution on the sphere, $\mathrm{d}\sigma$ is the spherical areal element, which can be expressed using direction cosine

$$\mathrm{d}\sigma = r^2\frac{\mathrm{d}\alpha\mathrm{d}\beta}{\gamma} \tag{14}$$

Since the three components of a vector field $\vec{E}$ in a homogeneous and isotropic media must individually satisfy Helmholtz equation, the above formula is also valid for vector optics. But he three components are not independent, $\vec{E}$ must satisfy Maxwell's equation, which translates to the divergence-free condition $\nabla\cdot\vec{E}(x,y,z) = 0$ in spatial domain, or $\hat{k}\cdot\vec{E}(\alpha,\beta)$ in frequency domain. Because of this, it is only necessary to specify the behavior of two components of E under consideration.

Now we apply the above consideration to vector diffraction, with the assumption that only two of the components are governed by the formula, and the third components are implicitly defined by equation (**??**) if we need it. For diffraction from mask near- field to entrance pupil with radius r,

$$\vec{E}^{\mathrm{ent}}(\alpha,\beta) = \frac{\cos\theta}{j\lambda}\frac{e^{jkr}}{r}\int_{\Sigma}\vec{E}^{\mathrm{mask}}(x,y)e^{-j2\pi(fx+gy)}\,\mathrm{d}x\mathrm{d}y. \tag{15}$$

For diffraction from exit pupil to wafer plane with radius r',

$$\vec{E}^{\mathrm{wafer}}(x',y') = \frac{n'}{j\lambda}r'e^{jk'r'}\iint_{S}\vec{E}^{\mathrm{ext}}(\alpha',\beta')e^{j2\pi(f'x'+g'y')}\,\frac{\mathrm{d}\alpha'\mathrm{d}\beta'}{\gamma'}. \tag{16}$$

Since $\hat{k}\cdot\vec{E}^{\mathrm{ent}}(\alpha,\beta) = 0$, we obtain the full-vectorial diffraction equation

$$\vec{E}^{\mathrm{ent}}(\alpha,\beta) = \frac{\cos\theta}{j\lambda}\frac{e^{jkr}}{r}\iint_{\Sigma}\left[\vec{E}^{\mathrm{mask}} - \frac{(\hat{k}\cdot\vec{E}^{\mathrm{mask}})}{(\hat{k}\cdot\hat{z})}\hat{z}\right]e^{-j2\pi(fx+gy)}\,\mathrm{d}x\mathrm{d}y \tag{17}$$

from the observation the the integration of the term $\frac{(\hat{k}\cdot\vec{E}^{\mathrm{mask}})}{(\hat{k}\cdot\hat{z})}\hat{z}$ is zero. The introduction of this term into the mask near-field guarantee the far-field is transverse. This equation has an particularly elegant vector representation

$$\vec{E}^{\mathrm{ent}}(\alpha,\beta) = \frac{1}{j\lambda}\frac{e^{jkr}}{r}\hat{k}\times\iint_{\Sigma}(\hat{z}\times\vec{E}^{\mathrm{mask}}(x,y))e^{-j2\pi(fx+gy)}\,\mathrm{d}x\mathrm{d}y. \tag{18}$$

This is the Smythe formula.? We will derive this equation in a more restricted setting and comment on it later when we discuss mask models.

4. MASK MODEL

Consider a typical point source s that emanates a plane wave in the direction $\hat{k}_s = [\alpha_s; \beta_s; \gamma_s]$ with wave vector $\vec{k}_s = \frac{2\pi n}{\lambda}[\alpha_s; \beta_s; \gamma_s]$, and illuminate the mask.

The incident plane wave are modulated by mask pattern and generate a field distribution just below mask called mask near-field. Let the mask near-field from an incident plane wave in s-mode as $[E^x_\perp(x, y; \alpha_s, \beta_s); E^y_\perp(x, y; \alpha_s, \beta_s)]$, and that from a p-mode as $[E^x_\parallel(x, y; \alpha_s, \beta_s); E^y_\parallel(x, y; \alpha_s, \beta_s)]$. For an incident plane wave with polarization vector $\left[E^s_\perp; E^s_\parallel\right]$, the mask near-field can be represented as

$$\vec{E}^{\text{mask}}(x, y; \alpha_s, \beta_s) = M(x, y; \alpha_s, \beta_s) \begin{bmatrix} E^s_\perp \\ E^s_\parallel \end{bmatrix} \tag{19}$$

where

$$M(x, y; \alpha_s, \beta_s) = \begin{bmatrix} E^x_\perp(x, y; \alpha_s, \beta_s) & E^x_\parallel(x, y; \alpha_s, \beta_s) \\ E^y_\perp(x, y; \alpha_s, \beta_s) & E^y_\parallel(x, y; \alpha_s, \beta_s) \end{bmatrix} \tag{20}$$

is called the mask transfer matrix for point source s. It is a 2 by 2 matrix that has the mask induced polarization effect built-in.

Once the mask near-field is known, the mask diffraction vector in the (s, p) coordinates can be represented by

$$\vec{E}^{\text{mask}} = \begin{bmatrix} E^{\text{mask}}_\perp \\ E^{\text{mask}}_\parallel \end{bmatrix} = T(\alpha, \beta) M(\alpha, \beta; \alpha_s, \beta_s) \begin{bmatrix} E^s_\perp \\ E^s_\parallel \end{bmatrix}, \tag{21}$$

where $T(\alpha, \beta)$ is the mapping from (x, y) coordinates to (s, p) system, and $M(\alpha, \beta; \alpha_s, \beta_s)$ is the Fourier transform of the 2 by 2 mask transfer matrix called mask diffraction matrix.

In general, mask near-field has to be computed with a rigorous Maxwell-equation solver,using for example RCWA or FDTD approaches. For an extended source distribution, it is impractical to rigorously solve Maxwell equation for the mask near-field for each source points. At most the rigorous solvers are invoked to compute rigorously the mask near-field for a few reference source points judiciously chosen, and approximate the mask near-field for nearby source points from the information of those of reference source points. For ease of presentation, we pick the on-axis point source as reference, and assume the mask transfer matrix for this reference is obtained through either a rigorous solver or some approximation.

In the case of scalar optics, the mask diffraction matrix collapse into scalar mask diffraction order. In this case, we normally assume that the diffraction order of an off-axis point source is the same in amplitude but with a shift in frequency space by the off-axis amount. This is the so-called Hopkins assumption for scalar optics.

In vector optics, extension will be performed in two steps. We first relate the polarization vector at an arbitrary point source s with that of reference by first map $\left[E^s_\perp; E^s_\parallel\right]$ to (x, y) coordinates using transfer matrix T_s, and then map the (x, y) components to (s, p) components using the transfer matrix T_r^{-1} for reference point-source. We them assume the diffract matrix is a shifted version of that of the reference. When the reference is ona-axis, the transfer matrix $T_r = I$. We assume

$$M(\alpha, \beta; \alpha_s, \beta_s) = M(\alpha - \alpha_s, \beta - \beta_s) T(\alpha_s, \beta_s) \tag{22}$$

where $M(\alpha - \alpha_s, \beta - \beta_s) = M(\alpha - \alpha_s, \beta - \beta_s; 0, 0)$ is the shifted diffraction matrix of reference. This is the Hopkins assumption for vector optics. It is a crucial assumption that makes the TCC formulation possible. For scanner and AIMS modeling, this is not a serious constraints[?] . For inspection tool, the incident angle on mask side can be over 70 degrees. Some modification of this assumption has to be made.

For many application, the speed of a rigorous solver for even a few reference points is not fast enough. Various approximations have to be used.

The simplest mask model is the thin mask assumption. The mask diffraction matrix for reference is assumed to be

$$M(\alpha,\beta) = \begin{bmatrix} m(\alpha,\beta) & 0 \\ 0 & m(\alpha,\beta) \end{bmatrix} \tag{23}$$

where $\hat{m}(\alpha,\beta)$ is the Fourier transform of $m(x,y)$, a scalar mask transmission function that is dependent on the mask pattern.

Although in wide use, the thin-mask assumption has been demonstrated inaccurate for feature width much smaller than wavelength. The boundary-layer model (BLM) in an attempt to correct some of the inaccuracies. It postulate that the transmission function $m(x,y)$ in a thin layer around the mask edge is a constant different from the bulk regions. We derived and implemented a generalization of the basic BLM to make the transmission function dependent on the polarization state and mask edge normal and tangential direction, and get very good match to the far-field from a rigorous solver.

Putting what we obtained so far together, we obtained

$$\begin{bmatrix} E_\perp^{\text{mask}} \\ E_\parallel^{\text{mask}} \end{bmatrix} = T^{-1}(\alpha,\beta) M(\alpha-\alpha_s,\beta-\beta_s) T(\alpha_s,\beta_s) \begin{bmatrix} E_\perp^s \\ E_\parallel^s \end{bmatrix}, \tag{24}$$

Under the thin-mask assumption, the above formulation becomes.

$$\begin{bmatrix} E_\perp^{\text{mask}} \\ E_\parallel^{\text{mask}} \end{bmatrix} = \hat{m}(\alpha-\alpha_s,\beta-\beta_s) T^{-1}(\alpha,\beta) T(\alpha_s,\beta_s) \begin{bmatrix} E_\perp^s \\ E_\parallel^s \end{bmatrix}. \tag{25}$$

Simple algebra shows that this is exactly the celebrated Smythe-Kirchoff equation:[?]

$$\vec{E}(\hat{k}) = \hat{k} \times \iint (\hat{z} \times \vec{E}_s) m(x,y) e^{-j\frac{2\pi}{\lambda}(\alpha x+\beta y)}\, \mathrm{d}x\mathrm{d}y \tag{26}$$

We comment that we can replace the transformation matrix T with $\tilde{T}$ in the above formula, and obtained a slightly different model that is used by some implementation. We feel that keeping the γ factor, especially on the image-side, is more consistent with vector formulation and is important to correctly characterize high-NA effect. Our formulation here is a generalization to arbitrary patterns to the formula proposed in Chen[?] for line-and-space patterns. Chen and his co-authors showed that results obtained by keeping the γ factor match with experimental data better.

5. PROJECTOR MODEL AND THE FORMATION OF IMAGES

An ideal projector will transform a diverging spherical wave at the entrance pupil into converging spherical wave at the exit pupil. Let us denote $\hat{k} = [\alpha;\beta;\gamma]$ as the unit wave-vector on the mask side, and $\hat{k}' = [\alpha';\beta';\gamma']$ as the unit wave vector of the corresponding wave-vector at the exit pupil. The unit vectors $\hat{e}_\perp$ is unchanged from entrance to exit pupil, but $\hat{e}_\parallel$ undergoes a rotation and becomes $\hat{e}'_\parallel$.

For a projector with demagnification factor M, the sine-condition says

$$Mn\alpha = -n'\alpha' \tag{27}$$

$$Mn\beta = -n'\beta' \tag{28}$$

where n and n' are the refractive index of the media on the mask and image side, respectively.

The spatial frequency on image and mask side are related by

$$f' = \frac{n'\alpha'}{\lambda} = -M\frac{n\alpha}{\lambda} = -Mf \tag{29}$$

$$g' = \frac{n'\beta'}{\lambda} = -M\frac{n\beta}{\lambda} = -Mg, \tag{30}$$

For an ideal projector without aberration, projector operation is a identity in (s,p) coordinate system

$$\begin{bmatrix} E_{\perp}^{ext} \\ E_{\parallel}^{ext} \end{bmatrix} = \begin{bmatrix} E_{\perp}^{ent} \\ E_{\parallel}^{ent} \end{bmatrix} \tag{31}$$

Our goal next is to first express the field in wafer plane in terms of the diffraction of mask near-field, and eventually in the polarization vector. For the purpose, let $\vec{E}^{\text{mask}}(\alpha,\beta)$ be the Fourier transform of $\vec{E}^{\text{mask}}(x,y)$, and $\vec{E}^{\text{wafer}}(\alpha',\beta')$ be the Fourier transform of $\vec{E}^{\text{wafer}}(x',y')$. We can rewrite the diffraction from mask near-field to entrance pupil as

$$\vec{E}^{\text{ent}}(\alpha,\beta) = \frac{\gamma}{j\lambda}\frac{e^{jkr}}{r}\vec{E}^{\text{mask}}(\alpha,\beta), \tag{32}$$

and represent the field at wafer plane in terms of the field at exit pupil as

$$\vec{E}^{\text{wafer}}(\alpha',\beta') = \frac{1}{\gamma'}\frac{1}{j\lambda}r'e^{jk'r'}\vec{E}^{\text{ext}}(\alpha',\beta'). \tag{33}$$

We next derive the radio metric correction term from consideration of energy conservation. Represent the field at exit pupil in terms of wafer-plane field as

$$\vec{E}^{\text{ext}}(\alpha',\beta') = \gamma'\frac{j\lambda}{n'}\frac{e^{-jk'r'}}{r'}\vec{E}^{\text{wafer}}(\alpha',\beta'). \tag{34}$$

Consider the pencil of rays entering entrance pupil at the direction $[\alpha;\beta;\gamma]$ and exit the pupil at the direction $[\alpha';\beta';\gamma']$. For an ideal lens, the conservation of energy requires that

$$n|\vec{E}^{\text{ent}}(\alpha,\beta)|^2 r^2\frac{\mathrm{d}\alpha\mathrm{d}\beta}{\gamma} = n'|\vec{E}^{\text{ext}}(\alpha',\beta')|^2 r'^2\frac{\mathrm{d}\alpha'\mathrm{d}\beta'}{\gamma'}. \tag{35}$$

This translates to

$$\frac{|\vec{E}^{\text{wafer}}(\alpha',\beta')|}{|\vec{E}^{\text{mask}}(\alpha,\beta)|} = c\sqrt{\frac{\gamma}{\gamma'}} \tag{36}$$

where c is a constant independent of diffraction order. The factor

$$R(\gamma,\gamma') = \sqrt{\frac{\gamma}{\gamma'}} \tag{37}$$

is called the radio-metric correction term.

In general, the projector has aberration, the wafer plane may not be at focus. All these effect can be represented by multiply a scalar or 2 by 2 matrix (Jones matrix). In(s,p) components the wafer field can be expressed as

$$\begin{bmatrix} E_{\perp}^{\text{wafer}} \\ E_{\parallel}^{\text{wafer}} \end{bmatrix} = R(\gamma,\gamma')A(\alpha',\beta')J(\alpha',\beta')C(\alpha',\beta')\begin{bmatrix} E_{\perp}^{\text{mask}} \\ E_{\parallel}^{\text{mask}} \end{bmatrix}, \tag{38}$$

where

$$C(\alpha',\beta') = \begin{cases} 1, & \sqrt{\alpha'^2+\beta'^2} \leq NA \\ 0, & \text{otherwise} \end{cases} \tag{39}$$

and $R(\gamma,\gamma')$ is the radio-metric correction term, $J(\alpha',\beta')$ a 2 by 2 Jones matrix, $A(\alpha',\beta')$ the aberration term.

When we compute the total field from different diffraction orders at the imaging plane, we can not simply add the s- and p-componets for different plane waves directly, since the unit s-vector and p-vectors are different for different diffraction orders. We must change from (s,p) coordinate system to the fixed (x,y,z) system. We also need to include the z-component that has been suppressed up to now to account for the high-NA effect.

$$\begin{bmatrix} E_x' \\ E_y' \\ E_z' \end{bmatrix}(\alpha',\beta') = \frac{1}{\sqrt{1-\gamma'^2}}\begin{bmatrix} -\beta' & -\alpha'\gamma' \\ \alpha' & -\beta'\gamma' \\ 0 & 1-\gamma'^2 \end{bmatrix}\begin{bmatrix} E_{\perp}^{\text{wafer}} \\ E_{\parallel}^{\text{wafer}} \end{bmatrix} = B(\alpha',\beta';\alpha_s,\beta_s)\begin{bmatrix} E_{\perp}^{s} \\ E_{\parallel}^{s} \end{bmatrix} \tag{40}$$

where

$$B(\alpha', \beta'; \alpha_s, \beta_s) = \begin{bmatrix} \frac{-\beta'}{\rho'} & \frac{-\alpha'\gamma'}{\rho'} \\ \frac{\alpha}{\rho'} & \frac{-\beta'\gamma'}{\rho'} \\ 0 & \rho' \end{bmatrix} W(\alpha', \beta'; \alpha_s, \beta_s) \tag{41}$$

and $W(\alpha', \beta')$ is a 2 by 2 matrix defined by

$$W(\alpha', \beta'; \alpha_s, \beta_s) = A(\alpha', \beta')J(\alpha', \beta')R(\gamma, \gamma')C(\alpha', \beta')T^{-1}(\alpha, \beta)M(\alpha - \alpha_s, \beta - \beta_s)T(\alpha_s, \beta_s). \tag{42}$$

The interference of different plane waves at the image plane is mathematically a Fourier inversion transformation. Hence the total E-field from points source s can be represented as

$$\vec{E}_s'(x', y') = \int\int \vec{E}(\alpha', \beta')e^{j2\pi(f'x'+g'y')}\, \mathrm{d}\alpha' \mathrm{d}\beta' = B(x', y') \begin{bmatrix} E_\perp^s \\ E_\parallel^s \end{bmatrix} \tag{43}$$

where $B(x', y') = \mathcal{F}^{-1}B(x', y')$ is the component-wise Fourier transform of the 3 by 2 matrix $B(x', y')$. Its column represent the E-field in physical domain for a source point in pure s-state or p-state.

In the presence of film-stack, we need to modified the 3 by 2 matrix in equation (**??**) to

$$\begin{bmatrix} \frac{-\beta'}{\rho'} A_\perp^q(z) & \frac{-\alpha'\gamma'}{\rho'} A_\parallel^q(z) \\ \frac{\alpha}{\rho'} A_\perp^q(z) & \frac{-\beta'\gamma'}{\rho'} A_\parallel^q(z) \\ 0 & \rho' \frac{n'\gamma'}{n_q\gamma_q} B_\parallel^q(z) \end{bmatrix} \tag{44}$$

where $A_\perp^q(z)$and $A_\parallel^q(z)$ are the horizontal components of field within layer q of the film-stack, for a unit incident wave to the film-stack in s-state and p-state, respectively, $B_\parallel^q(z)$ is the z-component in p-state. The details will be presented in a future work.

The intensity from point source s can be computed as

$$I_s = |E_s'|^2 = \left[\bar{E}_\perp^s, \bar{E}_\parallel^s\right] B^* B \left[E_\perp^s E_\parallel^s\right]. \tag{45}$$

The total intensity from an extended source can be computed by

$$I = \iint e(\alpha_s, \beta_s) \left[\bar{E}_\perp^s, \bar{E}_\parallel^s\right] B^* B \begin{bmatrix} E_\perp^s \\ E_\parallel^s \end{bmatrix} \mathrm{d}\alpha^s \mathrm{d}\beta^s \tag{46}$$

where $e(\alpha_s, \beta_s)$ represent the source strength at point source s.

The above formulation has follow the Abbe's source integration approach. Under the Hopkins' assumption, equation (**??**) can be reformulated by change the order of integration into a quadratic functions of the 4 elements of mask diffraction matrix, and the standard spectrum decomposition[?] can be applied to obtained a set of TCC-kernels. The details for 2 by 2 mask transfer matrix will be presented else.

If we further assume the thin-mask model, great simplification is achieved. We can find a set of TCC-kernels ϕ_k and corresponding eigenvalues λ_k dropping to zeros very fast, and the intensity be represented as

$$I = \sum_k \lambda_k |\phi_k \otimes m|^2 \tag{47}$$

where m as we recall, is the scalar transmission function. This formulation can be extended to cases where more than one transmission functions are used, for example in the rigorous solver. The details will be presented in a future work.

Once the optical image is obtained within the film stack, exposure step transform the intensity distribution into acid concentration, and during PEB, acid and base are diffused and quenched, and the concentration of resolution can be computed. Using one of the well-known development rate model, the developed resist profile can be computed. We omit the details here.

6. CONCLUSION

We presented a unified approach to modeling projection optics. We take a critical look at the modeling approach to each optical components, and point to the validity and limitations of the underlying assumptions. Future works will provide more details to some of the fine points briefly touched upon in the paper.

REFERENCES

[1] M. Born, E. Wolf, *Principle of Optics*, Cambridge University Press, 7th Ed. 1999

[2] J. Goodman, *Introduction to Fourier Optics*, 2nd ed, 1996.

[3] M. S. Yeung, *Modeling High Numerical Optical Lithography*, SPIE Vol 922 Optical/Laser Microlithography, 1988.

[4] M. S. Yeung, D. Lee, R. Lee and A.R. Neureuther, *Extension of the Hopkins Theory of Partially Coherent Imaging to Include Thin-Film Interference Effects*, SPIE Vol 1927, Optical/Laser Microlithography VI, 1993.

[5] D. G. Flagello, A.E. Rosenbluth, *Lithographic Tolerances Based On Vector Diffraction Theory*, J. Vac. Sci. Technol. B 10(b), 1992.

[6] D. G. Flagello, A.E. Rosenbluth, *Vector Diffraction Analysis of Phase-Mask Imaging in Photoresist Films*, SPIE Vol. 1927, Optical/Laser Microlithography VI, 1993.

[7] D. G. Flagello, *High Numerical Aperture Imaging in In Homogeneous Thin Films*, PhD Thesis, Univ. Of Arizona, 1993.

[8] D. C. Cole, E. Barouch, U. Hollerbach, S. Orszag, *Derivation and Simulation of High Numerical Aperture Scalar Aerial Images*, Jpn. J. Appl. Phys. Vol. 31, 1992

[9] D. C. Cole, E. Barouch, U. Hollerbach, S. Orszag, *Extending Scalar Aerial Image Calculations To Higher Numerical Apertures*, Jpn. J. Appl. Phys. Vol. 31, 1992

[10] G. M. Gallatin, *High-Numerical-Aperture Scalar Imaging*, Applied Optics, Vol. 40, No. 28, Oct. 2001.

[11] C. K. Chen, T.S. Gau, L.H. Shiu, Burn Lin *Mask Polarization Effects in Hyper NA Systems*, SPIE Vol. 5754, 2005.

[12] T. V. Pistor, A. R. Nuereuther, R. J. Socha, *Modeling Oblique Incidence Effects in Photomasks*, SPIE Vol. 4000, 2000

[13] F. Riesz, B. Sz.-Nagy, *Functional Analysis*, Dover, 1990.

[14] D. Jackson, *Classical Electrodynamics*, 3nd ed, 1999.