Integrating Wave-Optics and 5x5 Ray Matrices for More Accurate Optical System Modeling

Justin D. Mansell*, Robert Suizu, Robert Praus, Brent Strickler, Anthony Seward, and Steve Coy
MZA Associates Corporation, 2021 Girard Ste 150, Albuquerque, NM 87106

ABSTRACT

Two common techniques for modeling optical systems are wave-optics and ray matrices. Usually only one of these techniques is used based on the desired fidelity of the model and the desired model results. We present here a 5x5 ray matrix formalism that includes more optical effects than its smaller ray matrix predecessors. We then present a methodology to combine it with wave-optics to obtain a more complete optical model.

Keywords: Wave-optics, ray matrices, image rotation, image translation, optical system modeling.

1. INTRODUCTION

Wave-optics and ray matrices are two techniques that have been used for many years to model complex optical systems. The ray matrix technique is a fast way to gain an understanding of how a ray of light propagates through a series of optics.[1] In the traditional 2x2 form, they can be used to determine things like whether a system is imaging, if it adds curvature to a beam, or the system magnification. Expanding the 2x2 form to 3x3 allows for modeling of additional tilt and offset from the optical axis to capture misalignment and perturbations of optics. Expansion to a 4x4 form allows for modeling of image rotation and image inversion, but not misalignment and perturbation of optics. If more detailed system performance is required, wave-optics techniques are typically employed.

Wave-optics (a.k.a. Fourier optics) models simulate light using the Fresnel approximation of the Huygens-Fresnel principle.[2] Light propagated between two planes can be modeled using a two-dimensional Fourier transform of a grid of samples of the complex electric field. Although this technique is much more computationally intensive, it allows much more detail about the beam to be modeled including the effects of higher-order aberrations and the resulting system transfer functions like the modulation transfer function (MTF) or the optical transfer function (OTF). Unfortunately, wave-optics models are deficient in some ways. The traditional wave-optics model does not have any ability to determine the effects of image rotation or inversion and no simple way of adding perturbations and misalignments of the optics to the model without adversely affecting the propagation mesh requirements.

We present here a way to extend the ray matrix formalism to a 5x5 ray matrix that allows image rotation, inversion, translation, and tilt addition to be modeled. Furthermore, this technique allows perturbations to the optics to be modeled to first order. We then describe how it can be used to augment a wave-optics model so that the effects that it lacks can be integrated.

1.1. Ray Matrices

Ray matrices are a common technique for modeling simple optical systems. Using the traditional ray matrix technique, a ray of light is represented by its distance from the optical axis and its angle relative to the optical axis. A 2x2 ray matrix, often referred to as an ABCD matrix, is then used to represent each optical operation on the ray of light. This formalism is sufficient for modeling the behavior of simple optical systems, but lacks the ability to handle effects such as multi-dimensionality, image inversion, image rotation, image translation, and tilt addition.

There are formalisms that exist which expand the simple 2x2 matrices to handle some of these effects. Gerrard and Burch present a 3x3 matrix formalism that has been expanded to handle image translation and tilt addition.[3] A 4x4 matrix formalism is presented by Siegman that allows orthogonal axes to be modeled so that things such as simple

* jmansell@mza.com; phone 1 505 245-9970 x122

astigmatism and image rotation can be modeled.[1] We present here an integration of these two formalisms into a single 5x5 ray matrix formalism that enables modeling of effects such as tilt addition, image translation, and image rotation.

2. 5X5 RAY MATRIX FORMALISM

2.1. 5x5 Ray Matrix Structure

The 5x5 ray matrix formalism borrows heavily from the existing formalisms. The general form of a 5x5 matrix is applied to a 5-element ray as

$$\begin{bmatrix} x' \\ \theta_x' \\ y' \\ \theta_y' \\ 1 \end{bmatrix} = \begin{bmatrix} A_x & B_x & 0 & 0 & E_x \\ C_x & D_x & 0 & 0 & F_x \\ 0 & 0 & A_y & B_y & E_y \\ 0 & 0 & C_y & D_y & F_y \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} x \\ \theta_x \\ y \\ \theta_y \\ 1 \end{bmatrix} \qquad (1)$$

where x and y are the distances from the origin along the x and y axes, θ_x and θ_y are the angles of the ray from the optical axis in the x and y directions. The variables x, y, θ_x, and θ_y describe a ray. In the matrix, A, B, C, and D correspond to the components of the traditional 2x2 ray matrix formalism where their subscripts represent the effect in the x and y axes. Table 1 shows examples of many of the common 2x2 ray matrices.[1] E and F represent displacement from the optical axis and added tilt in the axes of their subscripts. The use of this form of 5x5 matrix allows for maximum leveraging of the existing traditional ray matrix formalism while adding only enough complexity to keep track of image rotation and image inversion.

Table 1 – Select Common 2x2 Ray Matrices

Matrix Type	Form	Variables
Propagation	$\begin{bmatrix} 1 & L/n \\ 0 & 1 \end{bmatrix}$	L = physical length n = refractive index
Lens	$\begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}$	f = effective focal length
Curved Mirror (normal incidence)	$\begin{bmatrix} 1 & 0 \\ -2/R & 1 \end{bmatrix}$	R = effective radius of curvature
Curved Dielectric Interface (normal incidence)	$\begin{bmatrix} 1 & 0 \\ -(n_2 - n_1)/R & 1 \end{bmatrix}$	n_1 = starting refractive index n_2 = ending refractive index R = effective radius of curvature

Although the traditional 2x2 ray matrix formalism can be heavily leveraged in the 5x5 formalism, several additional ray matrices were required to model some effects or simplify the representation of other effects. Table 2 shows the additional ray matrices added to the 5x5 formalism. Tilt addition and image translation were added in the same form as the 3x3 formalism. The image rotation matrix is borrowed from the 4x4 formalism.

Table 2 - Additional 5x5 Ray Matrices

Matrix Type	Form	Variables
Image Rotation	$\begin{bmatrix} C & 0 & S & 0 & 0 \\ 0 & C & 0 & S & 0 \\ -S & 0 & C & 0 & 0 \\ 0 & -S & 0 & C & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$C = \cos(\theta_{image})$ $S = \sin(\theta_{image})$ θ_{image} = image rotation angle
Magnification	$\begin{bmatrix} M_x & 0 & 0 & 0 & 0 \\ 0 & 1/M_x & 0 & 0 & 0 \\ 0 & 0 & M_y & 0 & 0 \\ 0 & 0 & 0 & 1/M_y & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	M_x = magnification along x axis M_y = magnification along y axis
Reflection Inversion (reflection from a mirror in the plane of the x-axis)	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	
Image Translation	$\begin{bmatrix} 1 & 0 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \Delta y \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	Δx = translation along the x axis Δy = translation along the y axis
Tilt Addition	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta\theta_x \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta\theta_y \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\Delta\theta_x$ = added tilt along the x axis $\Delta\theta_y$ = added tilt along the y axis

2.2. Applying the 5x5 Ray Matrix Formalism

Application of the 5x5 ray matrix formalism can be accomplished by multiplying successive matrices in the same way as the 2x2, 3x3, or 4x4 formalisms, but this requires that the user keep track of the image rotation manually and apply it as needed. The goal of this work was to automate the entire process so that the user did not have to do any piece of it manually. For this reason, additional procedural steps were added to enable the automatic inclusion of the image rotation effects.

The automated 5x5 ray matrix procedure starts with the optical system defined in three-dimensional space in a global Cartesian coordinate system by the nominal locations of each of the optics. The first coordinate is the optical source and the last coordinate is the target or sensor. The nominal ray vectors are determined by subtracting the next coordinate from the previous coordinate, or

$$r_j = C_j - C_{j-1} = (x, y, z)_j - (x, y, z)_{j-1} \qquad (2)$$

where r_j is the direction of the ray from optic j-1 to optic j, C_j is the coordinate of optic j, and C_{j-1} is the coordinate of optic j-1. The orientation of the optics can be determined by making sure that the reflection from each optic directs the beam from the previous optic to the next optic. The normal of each optic is determined by the normalized difference between the normalized ray toward that optic from the normalized ray away from that optic, or

$$\overline{N}_j = \overline{\left(\overline{r}_{j-1} - \overline{r}_j\right)} \qquad (3)$$

where N_j is the j^{th} optic normal, $\overline{r}_{j-1}$ is the normalized ray toward the j^{th} optic, and $\overline{r}_j$ is the normalized ray leaving the j^{th} optic. The single line above a vector represents normalization. The initial optic normal is aligned with the first ray. The final optic, which is typically a sensor or target, has a normal aligned with the ray that leads to it. Figure 1 shows an example of a simple optical system in which the rays and the normal vector of the optic has been calculated.

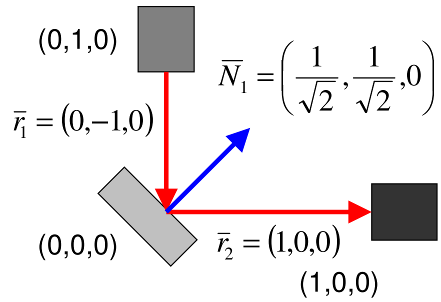

Figure 1 - Example of a simple optical system

2.3. Coordinate Transforms

There are three coordinate systems that are used throughout this procedure: global coordinates, optic local coordinates, and beam coordinates. Transforming between these coordinate systems is accomplished here using linear algebra matrix transforms. The global to optic local, henceforth referred to as local, coordinates will be addressed first.

2.3.1. Switching between Local and Global Coordinate Spaces

The first step in coordinate transforms is to establish three coordinate axes of the local space in the global coordinate system. In the local coordinate space, the x axis is the axis normal to the optic. The y axis is the result of cross product of the x axis and the global z axis [0 0 1]. If the magnitude of the cross product is small (indicating that the local x and global z axes are aligned) the y axis is set to [1 0 0]. The z axis is the cross product of the x and y axes. All axis vectors are normalized.

The global to local conversion matrix is established by first creating a 3x3 matrix of the global coordinates of the local axes with each axis as a column. Then the matrix is transposed to map from global to local. In the special case of a point being mapped to the front surface of an optic to determine the beam displacement (a.k.a beam walk) on the optic surface, the first row of the resulting matrix is removed to leave a 2x3 matrix.

An example transform using the system shown in Figure 1 is described here. The normal to the optic becomes the x axis ([1 1 0]/√2). The y axis is then [1 -1 0]/√2. Finally, the z-axis is given by [0 0 -1]. The matrix of axes is

$$M^{\text{local to global}}_{\textit{transform}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{bmatrix} \Big/ \sqrt{2}\,. \qquad (4)$$

The matrix is then transposed and the first row is removed to create the global three-space to local two-space transform matrix of

$$M^{\text{global to local}}_{\textit{transform}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{bmatrix} \Big/ \sqrt{2}\,. \qquad (5)$$

This matrix will allow a point in global three-space to be projected onto the two-space of the front of the optic.

2.3.2. Beam and Local Coordinate Spaces

Beam coordinate space uses the propagation direction as the z axis. The x axis is then made equal to the z axis of the local coordinate space. The y axis is set equal to the y axis of the local coordinate space. Projections of individual global coordinates onto the beam axes are done using dot products between the beam axes and the differential motion or rays in the optic local coordinate space using global coordinates. This process is described in the section below on extracting the 5x5 ray matrix from the results of the simplified three-space ray tracing.

2.4. Propagation through the System

The propagation of these rays through the system is done using a simplified ray tracing algorithm. This algorithm assumes that the optics are thin so that they are effectively planar when determining the position that the ray intersects the next optic. If the optic is curved, the field curvature is applied to the ray.

After the nominal positions of the optics have been determined, the intersection of the ray launched from a point in the source plane to plane of the first optic is determined. Each plane is defined by a point on the plane, which is the coordinate C_i of optic i, and the normal to the plane, which is the vector N_i. Each ray is defined by a point in the plane of the source, P_i, and a vector direction, v_i, which are both defined in beam coordinate space. The ray is converted into global coordinates using the transform matrix defined above that takes the two-space coordinates of the ray position and direction in beam coordinates and converts that into global coordinates. Then the intersection determination can be done in global coordinates.

There are many algorithms for determining the intersection of a ray with a plane. The algorithm implemented in our code is

$$
\begin{aligned}
&line = \{P_i, \overline{v}_i\},\ plane = \{C_{i+1}, \overline{N}_{i+1}\} \\
&s_{\text{intersection}} = \left(-\frac{\overline{N}_{i+1} \bullet (P_i - C_{i+1})}{\overline{N}_{i+1} \bullet \overline{v}_i}\right) \\
&P_{i+1} = P_i + (s_{\text{intersection}}) \cdot \overline{v}_i
\end{aligned}
\qquad (6)
$$

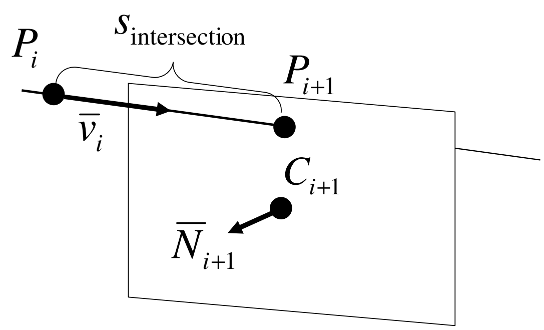

Figure 2 - Variables used in calculating the intersection of a plane and a line

where P_i and $\overline{v}_i$ are the coordinate and ray direction that define the line in global coordinates, C_{i+1} and $\overline{N}_{i+1}$ are the coordinate and normal that define the plane in global coordinates, and $s_{intersection}$ is the scalar distance along the line from the point P_i to the intersection point, and the large dot ($\bullet$) in the equations is the dot product, and P_{i+1} is the coordinate of intersection of the ray with the $(i+1)^{th}$ optic. Figure 2 shows the variables used in the calculation.

The beam walk (bw), or the displacement of the beam on the optic relative to the nominal ray position, is given then by the distance between the nominal ray point at the $(i+1)^{th}$ optic, C_{i+1}, and the actual intersection, P_{i+1}, or

$$
bw_{i+1}^{\;local} = M_{\;transform}^{\text{global to local}} \cdot bw_{i+1}^{\;global} = M_{\;transform}^{\text{global to local}} \cdot (C_{i+1} - P_{i+1}) \qquad (7)
$$

where C_i is the nominal coordinate of the i^{th} optic and bw_{i+1}^{local} is the two-space beam walk in the plane of the optic.

The beam walk is then translated from global coordinates into optic coordinates (defined above) by the appropriate transform matrix. The effect of the optic on the beam is then determined based on the optic type. If the optic is a flat, the output vector direction is given by

$$\overline{v}_{i+1} = \vec{v}_i - 2 \cdot \overline{N}_i \qquad (8)$$

where $\overline{r}_{out}$ is the output ray direction, $\overline{r}_{in}$ is the input ray direction, and $\overline{N}_i$ is the normal of the i[th] optic. If the optic is curved, the output ray also needs to be modified by the effect of the curvature on the ray direction. To accomplish this, the ray direction is modified as follows:

$$\begin{aligned}
\Delta\theta_{local} &= \begin{bmatrix} 0 & 1/f \\ -1/f & 0 \end{bmatrix} \cdot bw_{i+1}{}^{local} \\
\Delta\theta_{global} &= M_{transform} \cdot \Delta\theta_{local} \qquad (9) \\
\overline{v}_{i+1} &= \overline{v}_{i+1} + \Delta\theta_{global} \times \overline{v}_{i+1}
\end{aligned}$$

where f is the optic's effective focal length of the curved optic and the large x ($\times$) indicates cross product. This algorithm can be fairly easily extended to handle astigmatic optics that have an astigmatic axis along the beam axes, but that will not be addressed here. We will limit this description of the formalism by not addressing any form of astigmatism. Although there should be sufficient degrees of freedom and reason to do so relative to the mesh parameter requirements, astigmatism is already treated in our wave-optics theory.

Once the output ray coordinate, P_{i+1}, and the output ray direction, $\overline{v}_{i+1}$, have been determined, the propagation to the next optic can be initiated. This ray propagation algorithm is repeated until the ray intersects the plane of the source.

2.5. Determining the 5x5 Ray Matrix from Ray Tracing

Once the nominal position of the optical system has been determined, the system is probed with a set of 5 probe rays to determine the optical system's effect on a beam. The five probe vectors are chosen so as to match with the five columns of the 5x5 ray matrix. The first probe vector is the nominal vector. It starts on axis and propagates along the optical axis. The four subsequent rays are displaced a small amount, Δ, angularly and spatially from this nominal ray. Table 3 and Figure 3 describe the coordinates and directions of the five probe rays in beam coordinate space, where the ray is

Table 3 - The Five Probe Rays

Ray	Coordinate	Direction
a	C_1	$\overline{r}_1$
b	C_1	$\overline{\overline{r}_1 + \Delta\overline{x}}$
c	C_1	$\overline{\overline{r}_1 + \Delta\overline{y}}$
d	$C_1 + \Delta\overline{x}$	$\overline{r}_1$
e	$C_1 + \Delta\overline{y}$	$\overline{r}_1$

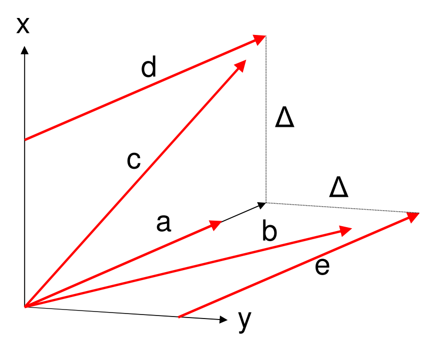

Figure 3 - Diagram of Probe Rays

always propagating along the z-axis. The small displacement should be chosen such that it is a small fraction of the size of the system, but large enough not to be near the edge of the dynamic range of the variables. When using double precision floating point numbers, we chose the magnitude of delta to be 10^{-6} m. After these five rays are propagated through the system, the difference in the four displaced rays, b to e, final coordinates (ΔP_b to ΔP_e) and the final direction vectors ($\Delta\overline{v}_b$ to $\Delta\overline{v}_e$) relative to the nominal vector, a, are calculated by

$$\begin{aligned}
\Delta P_b &= P_{last}(b) - P_{last}(a) & \Delta v_b &= \overline{v}_{last}(b) - \overline{v}_{last}(a) \\
\Delta P_c &= P_{last}(c) - P_{last}(a) & \Delta v_c &= \overline{v}_{last}(c) - \overline{v}_{last}(a) \\
\Delta P_d &= P_{last}(d) - P_{last}(a) & \Delta v_d &= \overline{v}_{last}(d) - \overline{v}_{last}(a) \qquad (10) \\
\Delta P_e &= P_{last}(e) - P_{last}(a) & \Delta v_e &= \overline{v}_{last}(e) - \overline{v}_{last}(a)
\end{aligned}$$

The final calculation is done to determine the difference of the system from its unperturbed state. The nominal ray, a, is compared to another nominal ray operating on an unperturbed system. This results in the following differences

$$\begin{aligned} \Delta P_a &= P_{last}{}^{perturbed}(a) - P_{last}{}^{unperturbed}(a) \\ \Delta v_a &= \bar{v}_{last}{}^{perturbed}(a) - \bar{v}_{last}{}^{unperturbed}(a) \end{aligned} \tag{11}$$

The resulting 5x5 ray matrix,

$$\begin{bmatrix} A1 & A2 & A3 & A4 & A5 \\ B1 & B2 & B3 & B4 & B5 \\ C1 & C2 & C3 & C4 & C5 \\ D1 & D2 & D3 & D4 & D5 \\ E1 & E2 & E3 & E4 & E5 \end{bmatrix},$$

is determined as follows:

$A1 = \Delta P_d \bullet \bar{x}_{beam}{}^{global}$ $A2 = \Delta P_b \bullet \bar{x}_{beam}{}^{global}$ $A3 = \Delta P_e \bullet \bar{x}_{beam}{}^{global}$ $A4 = \Delta P_c \bullet \bar{x}_{beam}{}^{global}$ $A5 = \Delta P_a \bullet \bar{x}_{beam}{}^{global}$	$B1 = \Delta v_d \bullet \bar{x}_{beam}{}^{global}$ $B2 = \Delta v_b \bullet \bar{x}_{beam}{}^{global}$ $B3 = \Delta v_e \bullet \bar{x}_{beam}{}^{global}$ $B4 = \Delta v_c \bullet \bar{x}_{beam}{}^{global}$ $B5 = \Delta v_a \bullet \bar{x}_{beam}{}^{global}$	$C1 = \Delta P_d \bullet \bar{y}_{beam}{}^{global}$ $C2 = \Delta P_b \bullet \bar{y}_{beam}{}^{global}$ $C3 = \Delta P_e \bullet \bar{y}_{beam}{}^{global}$ $C4 = \Delta P_c \bullet \bar{y}_{beam}{}^{global}$ $C5 = \Delta P_a \bullet \bar{y}_{beam}{}^{global}$
$D1 = \Delta v_d \bullet \bar{y}_{beam}{}^{global}$ $D2 = \Delta v_b \bullet \bar{y}_{beam}{}^{global}$ $D3 = \Delta v_e \bullet \bar{y}_{beam}{}^{global}$ $D4 = \Delta v_c \bullet \bar{y}_{beam}{}^{global}$ $D5 = \Delta v_a \bullet \bar{y}_{beam}{}^{global}$	$E1 = 0$ $E2 = 0$ $E3 = 0$ $E4 = 0$ $E5 = 1$	

(12)

where $\bar{x}_{beam\_global}$ is the x axis of the beam space in global coordinates, $\bar{y}_{beam\_global}$ is the y axis of the beam space in global coordinates.

2.6. Sequential Application of the Algorithm

We describe applying the algorithm to an entire system at once, as a collective piece, but we also developed an algorithm for applying this algorithm sequentially through a set of optics. Each propagation of the sequential algorithm modeled the propagation of a ray from a source plane to a target plane. In every propagation but the first and last, the source and target were virtual. The virtual target plane of each propagation was determined by knowing the direction of the ray toward the next optic, r_i, and calculating the point along the nominal trajectory toward this second optic such as was done above with the intersection of a line with a plane as $C_{\text{intermediate}} = C_i + \Delta \cdot \bar{r}_i$. This plane becomes the virtual source plane of the next propagation. To maintain accuracy in the propagation distance, the distance Δ is subtracted from each propagation distance.

3. EXTRACTING WAVE OPERATIONS FROM 5X5 RAY MATRIX

Once a 5x5 ray matrix is determined for an optical system, the effect of that system on the transmitted rays is determined by using test rays. A set of seven test rays, in the form of ray vectors, are multiplied by the system ray matrix, M_{system}, to determine the system's effect on the test rays. The test rays are defined in Table 4 and Figure 4.

Table 4 - Test Rays

Ray	x	θ_x	y	θ_y
A	0	0	0	0
B	1	0	0	0
C	0	1	0	0
D	0	0	1	0
E	0	0	0	1
F	-1	0	0	0
G	0	0	-1	0
H	0	-1	0	0
I	0	0	0	-1

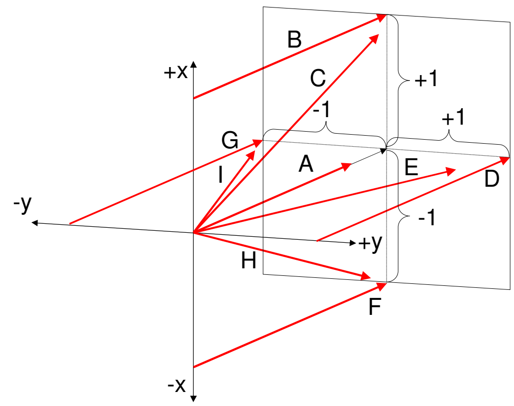

Figure 4 - Test Rays

The resulting rays found by the propagation of the test rays through the system can be used to determine the presence of several effects on a beam propagated through the system. The tilt and offset added to a beam is determined by looking at the position and angle of the nominal ray in the output plane, A', which is given by

$$\begin{bmatrix} \Delta x \\ \Delta\theta_x \\ \Delta y \\ \Delta\theta_y \\ 1 \end{bmatrix} = A' = \begin{bmatrix} x' \\ \theta_x' \\ y' \\ \theta_y' \\ 1 \end{bmatrix} = \begin{bmatrix} A1 & A2 & A3 & A4 & A5 \\ B1 & B2 & B3 & B4 & B5 \\ C1 & C2 & C3 & C4 & C5 \\ D1 & D2 & D3 & D4 & D5 \\ E1 & E2 & E3 & E4 & E5 \end{bmatrix} \begin{bmatrix} x \\ \theta_x \\ y \\ \theta_y \\ 1 \end{bmatrix} = M_{system} \cdot A \qquad (13)$$

where Δx is the image offset or translation along the x axis, $\Delta\theta_x$ is the tilt along the x axis, Δy is the image offset along the y axis, and $\Delta\theta_y$ is the tilt along the y axis.

The new beam axes, $+x_{axis}'$, $+y_{axis}'$, $-x_{axis}'$, and $-y_{axis}'$, are determined by subtracting the position of the optical axis ray, A, from the rays parallel to the optical axis, namely B, D, F, and G, or

$$\begin{aligned} +x_{axis}' &= \begin{bmatrix} x_B - x_A \\ y_B - y_A \end{bmatrix} & -x_{axis}' &= \begin{bmatrix} x_F - x_A \\ y_F - y_A \end{bmatrix} \\ +y_{axis}' &= \begin{bmatrix} x_D - x_A \\ y_D - y_A \end{bmatrix} & -y_{axis}' &= \begin{bmatrix} x_G - x_A \\ y_G - y_A \end{bmatrix} \end{aligned} . \qquad (14)$$

The presence of inversion can be determined by examining the sign of the cross product of the +x and +y axes. If the sign of the cross product is negative, then image inversion exists. The inversion is removed from the system matrix by multiplying by the inversion matrix shown in Table 2, which inverts the system about the x-axis. Then the system axes are calculated again before proceeding.

The system magnification is found by examining the length of the new +x and +y axes, or

$$\begin{aligned} M_x &= \left| + x_{axis}' \right| \\ M_y &= \left| + y_{axis}' \right| \end{aligned} . \qquad (15)$$

If we assume that no astigmatic terms are included in the ray matrix, M_x should equal M_y. The radius of curvature of the output beam, ROC, is calculated by taking the ratio of the magnitude of the new axes over the angle of the rays parallel to the optical axis in the positive direction, B and D, relative to the nominal ray, A, or

$$ROC_x = \left|+x_{axis}'\right| / \left(\theta_x(B) - \theta_x(A)\right)$$
$$ROC_y = \left|+y_{axis}'\right| / \left(\theta_y(D) - \theta_y(A)\right) \quad (16)$$

These radii of curvature should be the same as long as there are not any astigmatic terms in the matrix. The procedure for determining magnification and optical power are different if the output is a focal plane because the system magnification goes to zero. The output plane is checked to see if it is a focal plane by looking for zero displacement between the parallel rays A and B and A and D. In this case, the power and system magnification are set to zero.

The image rotation can be determined by looking at the direction of the axes relative to their original positions. The image rotations are calculated as,

$$\theta_{+x}^{image} = \cos^{-1}\left(\begin{bmatrix} 1 & 0 \end{bmatrix} \bullet \frac{+x_{axis}'}{\left|+x_{axis}'\right|} \right)$$
$$\theta_{+y}^{image} = \cos^{-1}\left(\begin{bmatrix} 0 & 1 \end{bmatrix} \bullet \frac{+y_{axis}'}{\left|+y_{axis}'\right|} \right) \quad (17)$$

Although both angles are not needed, they are used for error checking. If the difference between the two angles is significantly non-zero, there must be an error in the input matrix.

4. FIELD RECTIFICATION: APPLYING RAY EFFECTS TO THE WAVE

The results of the ray optics approach must be combined with the wave-optics approach at any point where a non-ideal optic is being used or where the beam is being sensed. This commonly occurs at atmospheric propagations, interactions with deformable mirrors, apertures, and interaction with any sensor. There are two approaches that are commonly used to combine the ray effects with the wave. The first involves interpolation of the complex field onto a new mesh. This is a complicated process that should be avoided when possible because of the noise it can introduce. The second approach is to interpolate the effect onto the complex field mesh. This can be done with more accuracy since the effects are typically only grids of floating point numbers, not grids of complex numbers. The application of the ray effects onto the field must be done in the right order to avoid negative interactions between the effects.

4.1. Order of Operations

Once the principle optical effects that are not part of the normal wave-optics calculations are determined by analyzing the system ray matrix, they must be applied to the wave. The order of these operations should be done carefully to avoid any interference between the effects. For example, if the magnification is applied after the focal power, the focal power is decreased by the magnification.

To determine a proper order of operations for applying the ray information to the wave, an analysis of ray matrices was performed. The six key effects that were analyzed are image rotation, image inversion, image offset, tilt, magnification, and optical power addition. The 5x5 ray matrices representing each of these effects were multipled by each of the other ones to determine if the first matrix times the second matrix was the same as the second matrix times the first matrix. In other words, it was determined whether the matrix product was commutative. If the matrices were commutative, there are no interaction problems with them.

Then the results were analyzed to determine which operation should be performed first by comparing each of the two products ([M_1]*[M_2] and [M_2]*[M_1]) to an ideal combined matrix formed by putting the two matrices together to create a single matrix that did both operations, but had no interactions between effects. This combined matrix started with the diagonal elements. If the first matrix, M_1, had a non-unity diagonal element, then the new matrix, M_c, used the M_1 diagonal element; otherwise it used the second matrix, M_2, diagonal element. For the off-diagonal elements, if M_1 or M_2 had a zero element and the other one did not, the new matrix used the non-zero element. If both were non-zero but equal, the new matrix element used the non-zero value. If both were non-zero and not equal, the matrices could not be combined and the element was set to not-a-number (NaN). If M_1 times M_2 was equal to the combined matrix, M_c, then the operation represented by M_2 should be performed before M_1. If M_2 times M_1 was equal to the combined matrix, M_c, then the operation represented by M_1 should be performed before M_2. If the matrices could not be combined, the interaction was flagged as a potential problem and had to be analyzed manually.

The three interactions that were not clear from the automated analysis were rotate-invert, power-magnify, and rotate-power. Manual analysis of the interactions found that the inversion must be done before rotation. Power must be applied after magnification because if magnification is applied after the power, the power increases by the magnification. Assuming that the axis of rotation is the same as the center of the power application, which it should be for the limited cases that were considered, the power and rotation can be done in any order. Figure 5 shows the interaction results after these three interactions were included.

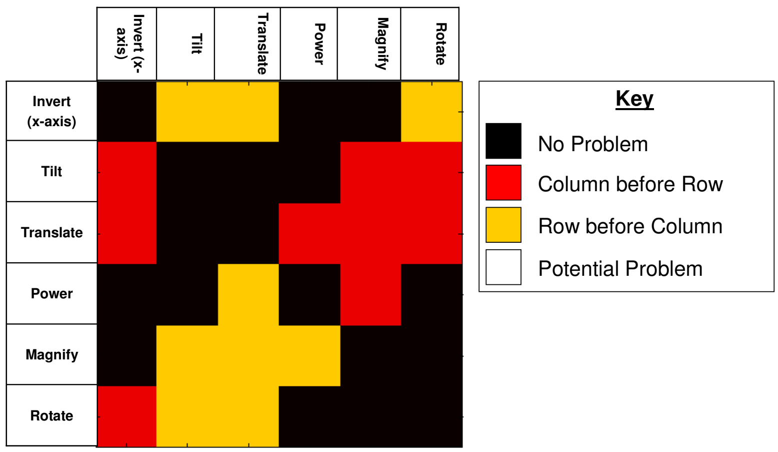

Figure 5 - Interaction results after manual analysis

The next step was to use the interactions to determine which of the many possible orders of operations were not in violation of any of the ordering rules. A program was written to generate every possible order (6!=720) and then each

Table 5 - Summary of Possible Orders of Operations

	Order of Operations					
Order Number	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
1	invert	rotate	magnify	power	translate	tilt
2	magnify	invert	rotate	power	translate	tilt
3	invert	magnify	rotate	power	translate	tilt
4	magnify	power	invert	rotate	translate	tilt
5	magnify	invert	power	rotate	translate	tilt
6	invert	magnify	power	rotate	translate	tilt
7	invert	rotate	magnify	power	tilt	translate
8	magnify	invert	rotate	power	tilt	translate
9	invert	magnify	rotate	power	tilt	translate
10	magnify	power	invert	rotate	tilt	translate
11	magnify	invert	power	rotate	tilt	translate
12	invert	magnify	power	rotate	tilt	translate
13	invert	rotate	magnify	tilt	power	translate
14	invert	magnify	rotate	tilt	power	translate
15	invert	magnify	rotate	tilt	power	translate

order was analyzed relative to the interaction matrix. The 15 possible orders of operations found are summarized in Table 5.

We chose to use the second order of operations for our implementation of the order of operations.

5. USING THE COMBINED ABCD AND HUYGENS PROPAGATION INTEGRAL

The Huygen's propagation integral can be written in terms of the 2x2 element ray matrix (a.k.a., the ABCD matrix) results of an optical system as long as no non-linear elements, such as deformable mirrors or apertures are introduced as[4]

$$U_2(x_2, y_2) = \frac{\exp(jkL)}{j\lambda B} \int \int_{-\infty}^{\infty} U_1(x_1, y_1) \exp\left[\frac{jk}{2B} \left(A\left(x_1^{\;2} + y_1^{\;2}\right) - 2\left(x_1 x_2 + y_1 y_2\right) + D\left(x_2^{\;2} + y_2^{\;2}\right)\right)\right] dx_1 dy_1 \tag{18}$$

where U_2 is the output electric field, U_1 is the input electric field, L is the on-axis path optical path length through the system computed as the sum of each path length times its refractive index, A, B, C, and D are the four elements of the 2x2 ray matrix, λ is the wavelength, and k is the wave number ($2\pi / \lambda$). The formalism presented here handles only one ABCD matrix, but can be extended to handle two-axis propagation.[5]

This formalism handles the common effects of magnification and power addition, but still cannot be used to handle image offset, tilt addition, image rotation, and image inversion. The 5x5 element ray matrix can be reduced to a 2x2 form by determining the other effects and removing them with appropriate matrix multiplications. Assuming that there are no astigmatic elements in the system, the 5x5 element ray matrix will reduce to two identical 2x2 element ABCD matrices and the Fourier propagation described above can be applied.

Like the procedure described above, the order of the application of the counteracting ray matrices must be handled properly. Using the order of operations analysis performed above (summarized in Table 5) the two orders which are most conducive to the application of the ABCD matrix Fourier operations then the post-processing are 4 or 10 because they are the orders with the power and magnification first.

6. CONCLUSIONS

We present here a 5x5 ray matrix approach to optical system analysis based on combining the existing 2x2, 3x3, and 4x4 formalisms and a simplified form of ray tracing. We then show how key ray effects can be extracted from the 5x5 element ray matrix and applied to the wave-optics formalism to allow for the inclusion of effects that are not capable of being handled by wave-optics, such as image rotation and image inversion, or are typically left out of wave-optics analysis because of the complexity they induce (for example, tilt and image offset). This 5x5 ray matrix formalism requires minimal additional computational power and even enables simplification of the wave-optics modeling process.

Acknowledgements
This work was funded under contract to the ABL SPO. We offer our thanks to Dr. Salvatore Cusumano and Captain Jason Tellez for their continued contributions and support.

REFERENCES

[1] A. E. Siegman, *Lasers*, Ch. 15, University Science Books, Mill Valley, CA, 1986.
[2] J. Goodman, *Introduction to Fourier Optics*, Ch. 3-4, McGraw-Hill, New York, 1988.
[3] A. Gerrard and J.M. Burch, *Introduction to Matrix Methods in Optics*, Appendix B, Dover, New York, 1975.
[4] Hans Peter Herzig, *MicroOptics: Elements, systems and applications*, Ch.1, Taylor & Francis, London, 1997.
[5] A. E. Siegman, *Lasers*, Ch. 20, University Science Books, Mill Valley, CA, 1986.