1040-5488/07/8401-0072/0 VOL. 84, NO. 1, PP. E72–E78
OPTOMETRY AND VISION SCIENCE

TECHNICAL REPORT

Paraxial Optics of Astigmatic Systems: Relations Between the Wavefront and the Ray Picture Approaches

EVA ACOSTA, PhD, and RALF BLENDOWSKE, PhD

Universidad de Santiago de Compostela, Santiago de Compostela, Spain (EA), and Hochschule Darmstadt, University of Applied Sciences, Darmstadt, Germany (RB)

ABSTRACT

Purpose. The paraxial propagation of astigmatic wavefronts through astigmatic optical systems can be described by the augmented step-along method (ASAM). Its equivalence to the linear ray optics approach is considered in detail.
Methods. The ASAM is exploited to derive paraxial ray paths through a general coaxial astigmatic system.
Results. Starting from the information inherent in the ASAM all 2×2 submatrices rendering the general 4×4 transference of linear optics can be generated. This proves the complete equivalence of both approaches. Additionally, we show that the symplectic relations are automatically obeyed in the ASAM.
Conclusions. The ASAM offers a complete alternative to describe the paraxial optics of astigmatic optical systems. According to the ASAM, an optical system is fully characterized by the back vertex vergence and the angular magnification matrix. Hence, a complete description of the paraxial optics of an eye should not only report the state of refraction but the angular magnification matrix as well, although it is not yet very common. The magnification matrix might be important in cases of anisometropia or the design of progressive addition lenses. Yet, a simple clinical procedure to determine the angular magnification matrix is missing.
(Optom Vis Sci 2007;84:E72–E78)

Key Words: astigmatic wavefronts, paraxial wavefront propagation, augmented step-along method, ray transference, symplectic relations, angular magnification

Although exact numerical ray-tracing through nearly any optical system is available today, paraxial optics is still key to an intuitive understanding of optical systems. The traditional Gaussian theory deals with a two-dimensional setting. Essentially, rays are considered in one plane only. This is appropriate for rotational symmetric systems. However, ophthalmic optics requires more elaborate paraxial tools because the mentioned symmetry of optical systems is lacking due to astigmatic elements. In principle, this calls for a strict three-dimensional treatment. Nevertheless, often a quite simple approach is chosen and a quite strong assumption is introduced: all elements of the considered systems are aligned in such a way that all principal axes of the elements throughout the system have the same orientation. For such systems each of the two meridians containing all the principal axes are dealt with separately and the traditional Gaussian theory is applied twice. This may be considered a poor approach to the three-dimensional problem.

Daily practice in ophthalmic optics, however, dictates the understanding of more sophisticated situations, where the principal axes of the elements under consideration may have arbitrary orientations. *A fortiori*, even decentered or tilted elements have to be considered. These problems call for an appropriate treatment and accordingly the related theory of linear optics of astigmatic systems has been extensively developed by Harris,[1] Keating,[2] Long,[3] and Fick[4] (in reverse historical order). To accommodate the three-dimensional nature of the optical setup, the scalar optical quantities known from rotational symmetric systems change their character and become vectors and matrices in the general case. These developments happened to be based mainly on ray optics, perhaps, because one light ray represents a quite intuitive concept.

A wavefront represents an infinite set of rays all being normal to the considered wavefront. Thus, if a wavefront is propagated, a whole bundle of rays is propagated at the same time. In other words: all rays belonging to one wavefront are not allowed to propagate freely but have to follow the collective rule governed by the wavefront. This has been explained recently by the authors in previous works[5,6] showing that the related equations are quite plain. In most cases, they are direct generalizations of their paraxial relatives from the Gaussian theory.

According to this paraxial wavefront approach, a general astigmatic system is characterized by three quantities: the incident and emergent vergences, which are symmetric 2×2 matrices, and by the angular magnification, which might be an asymmetric 2×2 matrix in the general case.

Given an on-axis incident wavefront with a vergence $\boldsymbol{L}_{in}$, the on-axis emergent wavefront is characterized by its vergence $\boldsymbol{L}_{out}$. All incident off-axis wavefronts with the same vergence $\boldsymbol{L}_{in}$ leave the system with the vergence $\boldsymbol{L}_{out}$ as well. However, the directions of the emergent wavefronts are connected to the ones of the incident wavefronts by the angular magnification matrix $\boldsymbol{N}$. In other words, all relevant information about the optical system under consideration is available from these three entities.

In the well-established linear optics approach, the optical system is characterized by the ray transference matrix. In this article, the explicit relation between both approaches, the ray and the wavefront picture, is developed to demonstrate their equivalence.

Thereby, the following assumptions valid for both models are applied throughout the whole article. First, our discussion is restricted to coaxial systems. Decentrations and tilts are not considered in detail. They may be included easily, as can be seen from the results. Second, we do not consider the topic of monochromatic aberrations, which by definition are outside the paraxial approach. As well, we leave the topic of chromatic aberrations aside.

We begin with a very brief account of the linear optics approach, that is the paraxial ray picture. This is followed by a short recall of the wavefront approach as described in the ASAM. However, the reader is referred to the literature[5,6] for a deeper understanding. We explore the relation of both approaches showing that the ASAM renders all relevant data appearing in the ray-based approach. A discussion of the practical impact while characterizing astigmatic systems ends this article.

Paraxial Ray Optics

We consider a paraxial ray in a given plane orthogonal to the optical axis. This is the z-axis, defined by the centers of curvature of all interfaces. At incidence onto a system the ray has transverse position $\mathbf{r}_1 = (x_1, y_1)^T$ and slope or direction $\mathbf{a}_1 = (a_{1,x}, a_{1,y})^T$. Both position and direction are relative to the optical axis. At emergence, the ray has position $\mathbf{r}_K$ and direction $\mathbf{a}_K$, where a system with K interfaces is considered. Instead of direction we work with reduced or optical direction, that is, direction multiplied by the index of refraction n of the surrounding medium, or $\boldsymbol{\alpha} = n\mathbf{a}$. In nonparaxial optics these numbers are usually called optical direction cosines. That is why we prefer the term *optical* instead of *reduced*.

The state of the ray is represented by the ray vector, a vector with position and optical direction as components. In particular, the incident ray vector is

$$\boldsymbol{\rho}_1 = \begin{pmatrix} \mathbf{r}_1 \\ \boldsymbol{\alpha}_1 \end{pmatrix} \tag{1}$$

and the emergent ray vector is

$$\boldsymbol{\rho}_K = \begin{pmatrix} \mathbf{r}_K \\ \boldsymbol{\alpha}_K \end{pmatrix} \tag{2}$$

The ray transference of an optical system is a 4×4 matrix (or linear operator) that converts the incident ray vector to the emergent ray vector according to

$$\boldsymbol{\rho}_K = \mathbf{S}\boldsymbol{\rho}_1 \tag{3}$$

The ray transference fully characterizes the paraxial optical nature of the system and may be divided up according to

$$\mathbf{S} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \tag{4}$$

Vectors are denoted by small boldface letters and matrices by capital boldfaced ones. The fundamental 2×2 matrices **A**, **B**, **C**, and **D** are not independent entries but obey the following so-called sympletic relations

$$\mathbf{A}^T\mathbf{C} - \mathbf{C}^T\mathbf{A} = \mathbf{0}$$

$$\mathbf{B}^T\mathbf{D} - \mathbf{D}^T\mathbf{B} = \mathbf{0}$$

$$\mathbf{A}^T\mathbf{D} - \mathbf{C}^T\mathbf{B} = \mathbf{I} \tag{5}$$

where the 2×2 identity and null matrices are denoted as **0** and **I**, respectively, and $\mathbf{A}^T$ is the matrix transpose of **A**.

The ray transference can be calculated by the repeated multiplication of the 4×4 transfer and refraction matrices which describe the free propagation and the refraction of a ray at an interface. A reflecting interface may be considered as a special case of refraction. By the process of refraction and transfer, the optical construction parameters like surface powers and distances between interfaces enter the calculation.

Applying the ray transference of the system, we arrive at a pair of vector equations relating the output and input quantities

$$\mathbf{r}_K = \mathbf{A}\mathbf{r}_1 + \mathbf{B}\boldsymbol{\alpha}_1 \tag{6}$$

$$\boldsymbol{\alpha}_K = \mathbf{C}\mathbf{r}_1 + \mathbf{D}\boldsymbol{\alpha}_1 \tag{7}$$

In the output reference plane in image space, the position of a ray and its direction are known once the entrance data of the ray in the input plane in object space and the ray transference are given. These relations completely determine all paraxial features of the considered optical system for a given pair of reference planes.

Paraxial Optics of Wavefronts

In a given coordinate system, a wavefront w is described in paraxial approximation by the equation

$$w(\mathbf{r}) = \frac{1}{2}\mathbf{r}^T\mathbf{L}\mathbf{r} - \mathbf{r}^T\mathbf{p} \tag{8}$$

where again $\mathbf{r}^T = (x, y)$ denotes the transpose of the position vector **r**. The tilt **p** of the wavefront at the optical axis is given by the normal to the wavefront at $\mathbf{r} = 0$ multiplied by the refractive index of the local medium. Be aware that the tilt **p** of a wavefront at the

z-axis must not be confused with the optical direction $\boldsymbol{\alpha}$ of an arbitrary ray. The vergence $\mathbf{L}$ of the wavefront is described by

$$\mathbf{L} = \begin{pmatrix} S + C\sin^2\phi & -C\cos\phi\sin\phi \\ -C\cos\phi\sin\phi & S + C\cos^2\phi \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} \\ L_{12} & L_{22} \end{pmatrix} \tag{9}$$

where $S = nc_1$, $C = n(c_2 - c_1)$, represent a combination of the principal curvatures and ϕ the angle of the orientation of c_1.

We consider a system of $k = 1, \ldots, K$ interfaces, separated by gaps with the reduced thicknesses t_k. An incoming vergence is traced through that system by repeated application of the equations for refraction and transfer given by[5]

$$\mathbf{L}'_k = \mathbf{L}_k + \mathbf{F}_k \tag{10}$$

$$\mathbf{L}_{k+1} = \frac{1}{\Delta_k}\mathbf{L}'_k - \frac{t_k}{\Delta_k}(\det\mathbf{L}'_k)\mathbf{I} = \frac{\mathbf{L}'_k}{\mathbf{I} - t_k\mathbf{L}'_k} \tag{11}$$

where k is the surface number, $\mathbf{F}_k$ the dioptric power matrix of that surface, and the diagonal identity matrix is denoted by $\mathbf{I}$. As usual vergences after a surface are distinguished by a prime, the quantity Δ is calculated by

$$\Delta = 1 - t(L'_{11} + L'_{22}) + t^2((\mathrm{L}'_{11}\mathrm{L}'_{22} - (\mathrm{L}'_{12})^2) \tag{12}$$

where the index k for all quantities has been dropped.

Finally, the angular magnification $\mathbf{N}$ is determined by the equation

$$\mathbf{N} = ([\mathbf{I} - t_1\mathbf{L}'_1] \times \ldots \times [\mathbf{I} - t_{K-1}\mathbf{L}'_{K-1}])^{-1} \tag{13}$$

This angular magnification governs the relation between the wavefront tilts in object and image space:

$$\mathbf{p}_K = \mathbf{N}\mathbf{p}_1 \tag{14}$$

It might be worth mentioning that the tilt vector does not change across a surface but across a gap only.

Rays and Wavefronts

In the revision of this article we became aware that the connection between rays and wavefronts and specially between a tilt and a decentration of a wavefront, have already been discussed by Deschamps[7] and Bastiaans.[8]

Since rays are normal to a wavefront, the local normal vector to a wavefront describes the direction or slope of a ray. This can be calculated from the local gradient. In paraxial approximation, we determine the optical direction at the position $\mathbf{r}$ by the equation

$$\boldsymbol{\alpha}(\mathbf{r}) = -\nabla w(\mathbf{r}) = -\begin{pmatrix} \partial w/\partial x \\ \partial w/\partial y \end{pmatrix} \tag{15}$$

Applying the paraxial definition of a wavefront, equation 8, yields the relation

$$\boldsymbol{\alpha}(\mathbf{r}) = -\mathbf{L}\mathbf{r} + \mathbf{p} \tag{16}$$

where $\mathbf{p}$ represents the tilt of the wavefront for $\mathbf{r} = \mathbf{0}$.

A ray having position $\mathbf{r}$ and optical direction $\boldsymbol{\alpha}$ can be transferred to the next interface, separated by the reduced thickness t. The position vector at the next interface (denoted by $+1$) is given by the equation

$$\mathbf{r}_{+1} = t\boldsymbol{\alpha} + \mathbf{r} \tag{17}$$

or inserting equation 16

$$\mathbf{r}_{+1} = (\mathbf{I} - t\mathbf{L})\mathbf{r} + t\mathbf{p} \tag{18}$$

For a more compact notation, we introduce the transfer matrix for the space between interface k and $k + 1$, denoted by

$$\mathbf{T}_k = \mathbf{I} - t_k\mathbf{L}'_k \tag{19}$$

This leads us to

$$\mathbf{r}_{k+1} = \mathbf{T}_k\mathbf{r}_k + t_k\mathbf{p}_k \tag{20}$$

To our knowledge, the matrix form of this equation has been used for the first time by Fick[4] and we propose to refer to equation 20 as the *Fick equation of ray transfer*. An example for the application of this equation and related marginal rays, the properties of Sturm's conoid, is presented in the appendix (available online at www.optvissci.com).

The transfer matrix is closely related to the angular magnification matrix of equation 13 which may be rewritten as a product of these transfer matrices

$$\mathbf{N} = (\mathbf{T}_1\mathbf{T}_2 \ldots \mathbf{T}_{K-1})^{-1} \tag{21}$$

Ray Transference and ASAM

We now study in some detail the propagation of rays through an astigmatic optical system applying the ASAM. The repeated application of equation 20 to a given ray leads to a recursive system of equations describing the ray coordinates throughout the system. To illustrate the development of the final formula, we explicitly write down the first three equations up to and including the fourth interface

$$\mathbf{r}_2 = \mathbf{T}_1\mathbf{r}_1 + t_1\mathbf{p}_1 \tag{22}$$

$$\mathbf{r}_3 = \mathbf{T}_2\mathbf{T}_1\mathbf{r}_1 + t_1\mathbf{T}_2\mathbf{p}_1 + t_2\mathbf{p}_2 \tag{23}$$

$$\mathbf{r}_4 = \mathbf{T}_3\mathbf{T}_2\mathbf{T}_1\mathbf{r}_1 + t_1\mathbf{T}_3\mathbf{T}_2\mathbf{p}_1 + t_2\mathbf{T}_3\mathbf{p}_2 + t_3\mathbf{p}_3 \tag{24}$$

Generally, for a system of K interfaces, we arrive at the following expression for the ray position at the last surface

$$\mathbf{r}_K = (\mathbf{T}_{K-1} \ldots \mathbf{T}_2\mathbf{T}_1)\mathbf{r}_1 + \sum_{l=1}^{K-1} t_l(\mathbf{T}_{K-1} \ldots \mathbf{T}_{l+1})\mathbf{p}_l \tag{25}$$

where $\mathbf{T}_{K-1} \ldots \mathbf{T}_{l+1} = \mathbf{I}$, when $l = K - 1$.

The appearing products of transfer matrices for the subsystems may be handled easier by introducing the shorthand notation

$$\tilde{\mathbf{T}}^{(l)} = \begin{cases} \mathbf{T}_{K-1} \ldots \mathbf{T}_l & \text{for } l \leq K - 1 \\ \mathbf{I} & \text{for } l = K \end{cases} \tag{26}$$

If the complete system is considered, we simply use

$$\mathbf{T} = \tilde{\mathbf{T}}^{(1)} \tag{27}$$

The tilt vectors $\mathbf{p}_l$ of the wavefronts at each surface in equation 25 can be related to the entrance tilt vector $\mathbf{p}_1$ by applying equation 14 for the subsystem with interfaces from 1 to l

$$\mathbf{p}_l = (\mathbf{T}_{l-1}^{-1} \ldots \mathbf{T}_1^{-1})\mathbf{p}_1 \quad l > 1 \tag{28}$$

We now rewrite equation 25 as

$$\mathbf{r}_K = \mathbf{T}\mathbf{r}_1 + \mathbf{W}\mathbf{p}_1 \tag{29}$$

where we have introduced the abbreviation

$$\mathbf{W} = \sum_{l=1}^{K-1} t_l \tilde{\mathbf{W}}^{(l)} \tag{30}$$

with

$$\tilde{\mathbf{W}}^{(l)} = (\mathbf{T}_{K-1} \ldots \mathbf{T}_{l+1}\mathbf{T}_l)(\mathbf{T}_l^{-1}\mathbf{T}_{l-1}^{-1} \ldots \mathbf{T}_1^{-1}) \tag{31}$$

The redundancy of the factors $\mathbf{T}_l\mathbf{T}_l^{-1} = \mathbf{I}$ is introduced for simplicity in forthcoming usage. Note that $\tilde{\mathbf{W}}^{(1)} = \mathbf{I}$.

In the next step, we relate the ASAM results of equations 14 and 29 to the linear optics approach, described by equations 6 and 7. To this end, we exploit equation 16 which gives the relation between the ray angle, the vergence of the wavefront, and the wavefront tilt. First, we calculate the local direction of a ray in image space at position $\mathbf{r}_K$ at the last surface K

$$\boldsymbol{\alpha}_K = -\mathbf{L}_K'\mathbf{r}_K + \mathbf{p}_K \tag{32}$$

Now the ray quantities of surface K have to be related to the ray quantities of surface 1 by equation 14 and 29 leading to

$$\boldsymbol{\alpha}_K = -\mathbf{L}_K'(\mathbf{T}\mathbf{r}_1 + \mathbf{W}\mathbf{p}_1) + \mathbf{N}\mathbf{p}_1 \tag{33}$$

In equations 29 and 33, the tilt $\mathbf{p}_1$ at surface 1 may be expressed via equation 16 by the relation

$$\mathbf{p}_1 = \boldsymbol{\alpha}_1 + \mathbf{L}_1\mathbf{r}_1 \tag{34}$$

After collection of terms, the substitution of equation 34 into equation 33 and 29 leads us to

$$\boldsymbol{\alpha}_K = [-\mathbf{L}_K'(\mathbf{T} + \mathbf{W}\mathbf{L}_1) + \mathbf{N}\mathbf{L}_1]\mathbf{r}_1 + [\mathbf{N} - \mathbf{L}_K'\mathbf{W}]\boldsymbol{\alpha}_1 \tag{35}$$

and

$$\mathbf{r}_K = (\mathbf{T} + \mathbf{W}\mathbf{L}_1)\mathbf{r}_1 + \mathbf{W}\boldsymbol{\alpha}_1 \tag{36}$$

Comparing these results to equations 6 and 7, we can determine the entries of the ray transference by the following relations

$$\mathbf{A} = \mathbf{T} + \mathbf{W}\mathbf{L}_1 \tag{37}$$

$$\mathbf{B} = \mathbf{W} \tag{38}$$

$$\mathbf{C} = -\mathbf{L}_K'\mathbf{A} + \mathbf{N}\mathbf{L}_1 \tag{39}$$

$$\mathbf{D} = \mathbf{N} - \mathbf{L}_K'\mathbf{B} \tag{40}$$

We summarize the result as follows: all vergences throughout the system and especially the back vertex vergence $\mathbf{L}_K'$ have to be calculated. An arbitrary incident wavefront vergence $\mathbf{L}_1$ can be used to this end. Then all transfer matrices $\mathbf{T}_l$ including their product $\mathbf{T}$ can be determined. This knowledge, including the inverse matrices $\mathbf{T}_l^{-1}$, allows for the calculation of the remaining quantities $\mathbf{N}$ and $\mathbf{W}$. Hence, for an arbitrary incident ray the coordinates and the direction of the related emergent ray can be determined by the ASAM without making reference to any other information, but the vergences provided by the step-along method applied to an on-axis vergence in object space.

For a special choice of the incident wavefront, namely $\mathbf{L}_1 = 0$, the equations 37 and 39 render a simpler form, because they reduce to $\mathbf{A} = \mathbf{T}$ and $\mathbf{C} = -\mathbf{L}_K'\mathbf{A}$.

This concludes the demonstration of the equivalence of both approaches: the ray transference and the ASAM. As an additional check, we will show in the appendix that the symplectic relations are automatically fulfilled in the ASAM.

Since the involved steps in the derivation of the above results might be considered lengthy, we repeat the main argument restricted to a simple example of a thick lens.

Example: Thick Lens

We consider the relationship between the ray transference and the ASAM in more detail for a thick lens. The lens is made up of two surfaces with surface powers $\mathbf{F}_1$, $\mathbf{F}_2$ and a reduced distance t. The reference planes for the ray transference are chosen to be identical with the two surfaces. The ray transference is calculated according to the product

$$\mathbf{S} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{F}_2 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & t\mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{F}_1 & \mathbf{I} \end{pmatrix} \tag{41}$$

which becomes

$$\mathbf{S} = \begin{pmatrix} \mathbf{I} - t\mathbf{F}_1 & t\mathbf{I} \\ -(\mathbf{F}_1 + \mathbf{F}_2 - t\mathbf{F}_2\mathbf{F}_1) & \mathbf{I} - t\mathbf{F}_2 \end{pmatrix} \tag{42}$$

According to the ASAM, applying $\mathbf{L}_1 = 0$ for simplicity, we calculate the vergences

$$\mathbf{L}_1' = \mathbf{F}_1 \tag{43}$$

$$\mathbf{L}_2 = \frac{\mathbf{F}_1}{\mathbf{I} - t_1\mathbf{F}_1} = \mathbf{F}_1\mathbf{T}_1^{-1} \tag{44}$$

$$\mathbf{L}_2' = \mathbf{L}_2 + \mathbf{F}_2 = \mathbf{F}_1\mathbf{T}_1^{-1} + \mathbf{F}_2 \tag{45}$$

From these equations, we calculate the entries of the ray transference according to the equations 37 through 40 yielding the following results

$$\mathbf{A} = \mathbf{T}_1 = \mathbf{I} - t_1\mathbf{F}_1 \tag{46}$$

$$\mathbf{B} = t_1\mathbf{W}^{(1)} = t_1\mathbf{I} \tag{47}$$

$$\mathbf{C} = -\mathbf{L}_2'\mathbf{T}_1 = -(\mathbf{F}_1\mathbf{T}_1^{-1} + \mathbf{F}_2)\mathbf{T}_1 \tag{48}$$

$$= -(\mathbf{F}_1 + \mathbf{F}_2\mathbf{T}_1) \tag{49}$$

$$= -(\mathbf{F}_1 + \mathbf{F}_2 - t\mathbf{F}_2\mathbf{F}_1) \tag{50}$$

All three entries are identical to those of the ray transference $\mathbf{S}$. We now go for the last missing entry

$$\mathbf{D} = \mathbf{N} - \mathbf{L}_2'\mathbf{W} \tag{51}$$

$$= \mathbf{T}_1^{-1} - t_1\mathbf{L}_2'\mathbf{I} \tag{52}$$

$$= (\mathbf{I} - t_1\mathbf{L}_2'\mathbf{T}_1)\mathbf{T}_1^{-1} \tag{53}$$

$$= \mathbf{I} - t\mathbf{F}_2 \tag{54}$$

In the last step, equation 49 for $\mathbf{L}_2'\mathbf{T}$ has been exploited. In conclusion, all entries of the ray transference have been reproduced from the ASAM information. The whole calculation may be repeated with $\mathbf{L}_1 \neq \mathbf{0}$. Besides some more intermediate steps, the results remain the same, as can be expected.

DISCUSSION

We have shown that the fundamental paraxial properties of a coaxial astigmatic system can be described by two equivalent ways. One option is given by the 4×4 ray transference, mapping ray positions and directions in an entrance and an exit plane of that optical system. The alternative is offered by the wavefront approach describing the propagation of vergences, given by 2×2 matrices, throughout the system.

In both approaches, the dioptric power matrix is applied to describe the action of an interface. The transfer is a common procedure as well, albeit once applied to a single ray and once applied to a wavefront. From these two elementary operations, shortly termed refraction and transfer, the properties of an astigmatic optical system are generated. The main difference between the two approaches lies in the way those properties are described.

The linear ray optics approach offers the four 2×2 **A**-, **B**-, **C**-, **D**-submatrices which build the ray transference. First, these matrices are not completely independent because of the symplectic relations. These equations introduce the coupling between rays which belong to one wavefront and must not be considered separately. While switching to the wavefront picture these constraints become obsolete. Second, the properties in which optometrists are interested have to be derived from the submatrices. These include the back vertex vergence and the magnification. They can be generated quite easily but nevertheless are derived quantities.

According to the ASAM, the fundamental paraxial properties of an astigmatic optical system are given by three quantities: the vergence of the entering wavefront, the back vertex matrix, and the angular magnification matrix. Keeping the object distance fixed, we need to consider only two quantities. First, we discuss the back vertex vergence which is given as a symmetric matrix. Therefore, choosing a proper rotation of the coordinate system this matrix becomes diagonal. In this frame of reference, we may apply the notions of traditional Gaussian theory to each of the principal meridians in image space. Thus, traditional terms like sphere, cylinder, and axis evolved to characterize the emerging on-axis wavefront but obviously not the complete system. This on-axis approach, restricted to the imaging of one field point only, may be considered sufficient as long as it is restricted to the on-axis matching of wavefronts. The procedure of on-axis matching forms the basis for the prescription of optical correction like spectacle or contact lenses: the wavefront that emerges from the lens has to be of the same form as the incident wavefront which enters the eye and generates a point-like image. Since these prescriptions play an important role in optometric practice, we can understand why the simple Gaussian theory outlived its limitations in so many cases.

Second, we have to consider the angular magnification matrix. This quantity describes how the object field is imaged to image space or in a simpler way: how the image is related to the object. The angular magnification matrix might in general deviate from what is termed magnification elsewhere, e.g., for near objects. A thorough discussion is given by Harris.[9,10]

Independent of the question whether the ASAM or the linear ray optics approach is used, it should be clear that the traditional two-dimensional Gaussian theory contributes nothing in this region if an astigmatic system is under consideration.

Since the angular magnification is given as an asymmetric matrix in the general case, the impact of an astigmatic system on the imagery can be described by the vocabulary of affine geometry. The angular magnification matrix represents an affine transformation, which maps lines to lines and parallel lines to parallel lines. However, circles, for example, might be imaged as rotated ellipses and vice versa. In general, the affine group includes a translation as well. Therefore noncoaxial optical systems with decentered or tilted elements are included as well, since the result of decentering or tilting is merely a translation of the optical direction in image space. As a side issue, prismatic corrections might be mentioned as well, which contribute to the translation of optical direction in image space. The paraxial off-axis optics of a general astigmatic system is therefore completely characterized by an affine transformation determined by the angular magnification and a translation. We do not step into the questions of singular angular magnification matrices here, because they are not of practical concern in the field of optometry.

From the perspective of the ASAM, a normal prescription, describing the state of refraction, is giving only half the information necessary to characterize an astigmatic optical system in the paraxial region. Only the information on the back vertex vergence of an optical correction is given. Missing is any information on magnification. In addition, there exist clinical problems with the correction of cylinders even in the monocular case, and binocular corrections including eyes with strong astigmatism might be quite difficult. One might speculate whether the change of image forms and sizes are responsible for the difficulty of accepting certain corrections. The understanding of the related magnification matrix is mandatory for solving these problems. However, yet there is no simple way to determine the magnification of astigmatic eyes experimentally. Nevertheless, it might be worth some effort to go in this direction.

Much effort has been spent to characterize the on-axis properties of emerging wavefronts, leaving the eye beyond the paraxial approximation. Nowadays higher order aberrations of the wavefront are considered and measured, e.g., by Hartmann-Shack-sensors, as well. The results render a better understanding of the point spread function related to the aberrations. These results are beyond the scope of paraxial calculations. However, this detailed knowledge is restricted to only one point concerning the object field of the optical system. Frequently, the effect on the imagery of extended objects, like letter charts, is visualized by a convolution of the object with the point spread function. Obviously, blurred letters are the results and the amount of blurring depends on the amount of aberrations. This procedure implicitly rests on two assumptions. First, the aberrations do not change drastically with the field position and second the magnification of the optical systems; besides, a simple scaling does not change the form of the letter. The last assumption however is not true if astigmatic systems are under consideration. Therefore, the shape of the displayed letters is at best unknown. At least paraxial information on the magnification should be applied to determine the shape of the imaged letter and in a second step the effect of aberrations can be included. Again, the information on off-axis behavior of astigmatic systems is often not included yet.

ACKNOWLEDGMENTS

The work of E. Acosta was supported by Ministerio de Educación y Ciencia, grant AYA2004-07,773-C02-02 (Spain), and FEDER. R. Blendowske gratefully acknowledges the support by the ZFE at the Darmstadt University of Applied Sciences and a grant from the Forschungsgemeinschaft Deutsche Augenoptik.

Received November 14, 2005; accepted October 12, 2006.

APPENDIX
Symplectic Relations in the ASAM

The symplectic relations, equations 5, are related to the choice of rays as the basis to describe an optical system. The symplectic relations then represent the constraints according to the fact that rays do not behave completely independent of each other, once they are connected to the same wavefront. If, instead of rays, wavefronts are considered, these constraints are obeyed automatically. This shows that the introduced quantities in the ASAM are a good choice to characterize the optics of paraxial astigmatic systems.

One might compare the case of symplectic relations to the description of a pendulum in cartesian coordinates. This choice is feasible, but since the length of the pendulum is fixed, the two cartesian coordinates are coupled to each other, because an additional constraint has to be obeyed. If we switch to polar coordinates the constraint is related to one coordinate only, the radius, while the other coordinate, the angle, is describing the motion of the pendulum independently.

In the following we will show that all symplectic relations, equations 5, are inherent in the ASAM. We will make frequent use of the quantities defined in equations 37 through 40, albeit in the simpler form rendered by the choice $\mathbf{L_1} = \mathbf{0}$. It might be worth recalling that the matrix transpose of a product of matrices reverses the order of the matrices, for example $(\mathbf{AC})^T = \mathbf{C}^T\mathbf{A}^T$. We start with the first relation

$$\mathbf{A}^T\mathbf{C} - \mathbf{C}^T\mathrm{A} = 0$$

yielding

$$\tilde{\mathbf{T}}^T(-\mathbf{L}'_K\tilde{\mathbf{T}}) - (-\mathbf{L}'_K\tilde{\mathbf{T}})^T\tilde{\mathbf{T}} = 0$$

or

$$\tilde{\mathbf{T}}^T(\mathbf{L}'_K - (\mathbf{L}'_K)^T)\tilde{\mathbf{T}} = 0$$

Since the vergence of the emerging wavefront $\mathbf{L}'_K$ is always symmetric, we have $\mathbf{L}'_K - (\mathbf{L}'_K)^T = 0$. Thus the first relation holds.

The second relation

$$\mathbf{B}^T\mathbf{D} - \mathbf{D}^T\mathbf{B} = 0$$

leads us to

$$\mathbf{W}^T(\tilde{\mathbf{N}} - \mathbf{L}'_K\mathbf{W}) - (\mathbf{N} - \mathbf{L}_K{}'\mathbf{W})^T\mathbf{W} = 0$$

Again we use the symmetry property of the vergence matrix and get

$$\mathbf{W}^T\mathbf{N} - \mathbf{N}^T\mathbf{W} = 0$$

After right hand multiplication with $\cdot\tilde{\mathbf{N}}^{-1}$ this equation can be re-written as

$$\mathbf{W}^T = \mathbf{N}^T\mathbf{W}\mathbf{N}^{-1} \tag{55}$$

By definition, see equations 30 and 31, the quantity W is given as a sum. We will show that for each term in this sum the following relation holds

$$(\mathbf{W}^{(l)})^T = \mathbf{N}^T\mathbf{W}^{(l)}\mathbf{N}^{-1} \tag{56}$$

If this relation is respected from each term in the sum it follows that the sum as whole will respect equation 55. The left hand side of equation 56 is given by

$$(\mathbf{W}^{(l)})^T = (\mathbf{T}_1^{-1}\ldots\mathbf{T}_{l-1}^{-1})\,(\mathbf{T}_{l+1}\ldots\mathbf{T}_{K-1})$$

The right hand side of the considered relation may be evaluated as follows

$$\begin{aligned}\mathbf{N}^T\mathbf{W}\mathbf{N}^{-1} &= (\mathbf{T}_1^{-1}\ldots\mathbf{T}_{K-1}^{-1})\,(\mathbf{T}_{K-1}\ldots\mathbf{T}_{l+1})\,(\mathbf{T}_1\ldots\mathbf{T}_{l-1})^{-1}\,(\mathbf{T}_1\ldots\mathbf{T}_{K-1})\\ &= (\mathbf{T}_1^{-1}\ldots\mathbf{T}_l^{-1})\,(\mathbf{T}_1\ldots\mathbf{T}_{l-1})^{-1}\,(\mathbf{T}_1\ldots\mathbf{T}_{K-1})\\ &= (\mathbf{T}_1^{-1}\ldots\mathbf{T}_{l-1}^{-1})\,(\mathbf{T}_{l+1}\ldots\mathbf{T}_{K-1})\end{aligned}$$

This result equals the one of the left hand side. Thus the relation 56 holds for each contribution of the sum and hence for the whole sum. This completes the proof of the second relation.

The third relation

$$\mathbf{A}^T\mathbf{D} - \mathbf{C}^T\mathbf{B} = \mathbf{I}$$

provides the equation

$$\mathbf{T}^T(\mathbf{N} - \mathbf{L}'_K\mathbf{W}) - (-\mathbf{L}'_K\mathbf{T})^T\mathbf{W} = \mathbf{I}$$

With the help of $\mathbf{L}'_K = (\mathbf{L}'_K)^T$ we arrive at

$$\mathbf{T}^T\mathbf{N} = \mathbf{I}$$

or after application of a left multiplication by $(\mathbf{T}^T)^{-1}$ we get

$$\mathbf{N} = (\mathbf{T}^T)^{-1}$$

By inspection of the definitions of both quantities, equations 13 and 26, we can see the validity of this equation.

In conclusion all three symplectic relations are respected automatically within the ASAM approach. They impose no additional condition on the basic quantities involved in the ASAM.

Sturm's Conoid Revised

As an example of how ray intercepts can be calculated from the ASAM and the Fick equation of ray transfer, equation 20, we consider Sturm's conoid. At incidence we assume a plane wave, with vergence $\mathbf{L} = \mathbf{0}$, which might be tilted and the related optical direction is given by $\mathbf{p}$. We assume a thin lens in air characterized by a dioptric power matrix **F**, leading to an emergent vergence matrix of $\mathbf{L}' = \mathbf{F}$. If the lens rim is hit by a ray at position $\mathbf{r}_0$ the related ray intercept with a plane at distance $t = z$ is given by

$$\mathbf{r}(z) = (\mathbf{I} - z\mathbf{L}')\mathbf{r}_0 + z\mathbf{p}$$

For illustrating purposes we choose the coordinate system in a way that the emerging vergence matrix is diagonal with diagonal elements L_1' and L_2'. Obviously, there are two focal lines at $z_1 = 1/L_1'$ and $z_2 = 1/L_2'$. Let us consider, for instance, the plane through z_1. The coordinates of the ray are then given by

$$x(z_1) = \frac{p_x}{L_1'}$$

$$y(z_1) = \left(1 - \frac{L_2'}{L_1'}\right)y_0 + \frac{p_y}{L_1'}$$

Obviously, the x-coordinate does not depend on the rim coordinates of the lens but on the tilt of the wavefront only, yielding an off-set p_x/L_1'. The length of the focal line l is proportional to the difference of the extreme y_0-coordinates

$$l = \left(1 - \frac{L_2'}{L_1'}\right)(y_{0,2} - y_{0,1})$$

Obviously, we have $l = 0$ for a spherical wave. If the lens has a circular shape its diameter is given by $2R = (y_{0,2} - y_{0,1})$ leading to

$$l = 2\left(1 - \frac{L_2'}{L_1'}\right)R$$

The second focal line shows the same properties as the first one and we omit a detailed discussion.

Next, we reconsider the circle of least confusion. As a first step we may ask at which position we can expect a scaled replica of the circular rim shape of the lens. A tilt of the incident wavefront does not change the overall shape of the conoid but only leads to an offset of coordinates. This has no influence on the following considerations and we therefore suppress the offset for simplicity. A replica of the circular lens shape is given if $\mathbf{r}$(z) is proportional to $\mathbf{r}^0$. This imposes the following condition on z

$$|1 - zL_1'| = |1 - zL_2'|$$

because the coordinates x^0 and y^0 have to be scaled by the same multiplicative factor, where a minus sign might be allowed for as well. The two possible solutions for the replica position z_R are

$$z_R = 0$$

$$z_R^{-1} = \frac{L_1' + L_2'}{2}$$

The first solution gives the position of the lens itself. The second one represents the well known result of the circle of least confusion, where the inverse of z_R is given by the arithmetic mean of the two curvatures L_1' and L_2'. It might be worth mentioning that the circle of least confusion is smaller than the lens itself if the relation $L_1' \cdot L_2' > 0$ holds. Otherwise the magnification of the replica, given by $|(L_1' - L_2')/(L_1' + L_2')|$ is >1.

REFERENCES

1. Harris WF. A unified paraxial approach to astigmatic optics. Optom Vis Sci 1999;76:480–99.
2. Keating MP. A system matrix for astigmatic optical systems. I. Introduction and dioptric power relations. Am J Optom Physiol Opt 1981;58:810–19.
3. Long WF. A matrix formalism for decentration problems. Am J Optom Physiol Opt 1976;53:27–33.
4. Blendowske R. HH Fick: Early contributions to the theory of astigmatic systems. S Afr Optom 2003;62:105–10. Available at: http://www.saoa.co.za/publications/saoptom/2003/sept/high/Journalvol62no3Fick.pdf. Accessed October 19, 2006.
5. Acosta E, Blendowske R. Paraxial propagation of astigmatic wavefronts in optical systems by an augmented stepalong method for vergences. Optom Vis Sci 2005;82:923–32.
6. Blendowske R, Acosta E. Paraxial propagation of astigmatic wavefronts through noncoaxial astigmatic optical systems. Optom Vis Sci 2006;83:119–22.
7. Deschamps GA. Ray techniques in electromagnetics. Proc IEEE 1972;60:1022–35. Available at: http://ieeexplore.ieee.org/iel5/5/31160/01450780.pdf. Accessed October 19, 2006.
8. Bastiaans MJ. Wigner distribution function and its application to first-order optics. J Opt Soc Am 1979;69:1710–16.
9. Harris WF. Magnification, blur, and ray state at the retina for the general eye with and without a general optical instrument in front of it: 1. Distant objects. Optom Vis Sci 2001;78:888–900.
10. Harris WF. Magnification, blur, and ray state at the retina for the general eye with and without a general optical instrument in front of it: 2. Near objects. Optom Vis Sci 2001;78:901–5.

Ralf Blendowske
Hochschule Darmstadt, Fachbereich Mathematik und Naturwissenschaften, Optotechnik und Bildverarbeitung
Schöfferstrasse 3
D-64295 Darmstadt, Germany
e-mail: blendowske@h-da.de