S Afr Optom 2007 **66**(2) 56-61

Transformations of ray transferences of optical systems to augmented Hamiltonian matrices and the problem of the average system*

JR Cardoso and **WF Harris*****

**Instituto Superior de Engenharia de Coimbra, Rua Pedro Nunes, 3030-199 Coimbra, Portugal and Institute of Systems and Robotics, Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal

***Optometric Science Research Group, Department of Optometry, University of Johannesburg, PO Box 524, Auckland Park, 2006 South Africa

**<jocar@isec.pt>
***<wfh@uj.ac.za>

Abstract

The first-order optical nature of an optical system (including an eye) is completely characterized by a 5×5 matrix called the ray transference. It is known that the image of a ray transference by the matrix logarithm function is an augmented Hamiltonian matrix. It turns out that there are other ways of transforming transferences into augmented Hamiltonian matrices. They include Cayley transforms and modified Cayley transforms. This paper will describe these transforms with a view to finding the most suitable one for quantitative analyses of eyes and other systems in augmented Hamiltonian spaces. In particular we look at the calculation of average systems.

Key words: ray transference, symplectic matrix, Hamiltonian matrix, matrix functions, average

Introduction

An important issue in the quantitative analysis of optical systems is the question of how to calculate an average of a set of eyes or other optical systems. One of us[1] has proposed the *exp-mean-log-transference* and we[2] have examined associated conditions of existence, uniqueness and symplecticity. (The abbreviations *exp* and *log* concern the matrix exponential and the matrix logarithm respectively; they will be used throughout the paper.) The method is implemented in Matlab® making use of the functions `expm` and `logm`. The procedure works as follows: the log function transforms ray transferences into augmented Hamiltonian matrices; one computes their arithmetic mean, another augmented Hamiltonian matrix, which in turn is transformed back into a ray transference via the exp function. This latter matrix is the average resulting from the application of the exp-mean-log. The method takes advantage of the fact that augmented Hamiltonian matrices define a linear or vector space and so allow the operations involved in basic quantitative analysis. One drawback of the exp-mean-log is the restriction required on the spectrum of the ray transferences involved: it can only be applied to sets of systems none of whose transference has a negative eigenvalue.[2] In practice this is not a problem for sets of eyes but it does limit the generality of the method.

The purpose of this paper is to find other pairs of transformations that work in the same way, that is, one of them transforms ray transferences into augmented Hamiltonian matrices and the other (the inverse)

*Paper originally presented at Mopane 2006A, Visual Optics, an international conference held at Mopani Camp, Kruger National Park, Limpopo Province, South Africa, August 6-9, 2006.

Received 23 November 2006; revised version accepted 7 June 2007

transforms augmented Hamiltonian matrices into ray transferences. In other words we are seeking other transformations that allow the same calculations in augmented Hamiltonian space but that overcome the drawback of the exp-mean-log. In particular we look at the calculation of an average system.

An accompanying paper[3] suggests suitable bases for quantitative analysis in augmented Hamiltonian space and a second paper[4] illustrates application to the cornea in particular.

Ray transferences and their logarithms

The first-order optical nature of an optical system (including an eye) is completely characterized by a 5×5 matrix **T** (the ray transference) which has the form

$$\mathbf{T} = \begin{pmatrix} \mathbf{S} & \boldsymbol{\delta} \\ \mathbf{o}' & 1 \end{pmatrix}. \quad (1)$$

S is a 4×4 symplectic matrix, that is, it obeys

$$\mathbf{S}'\mathbf{E}\mathbf{S} = \mathbf{E} \quad (2)$$

where

$$\mathbf{E} = \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix} \quad (3)$$

and **O** and **I** are 2×2 null and identity matrices respectively, $\boldsymbol{\delta}$ is 4×1 and the bottom row of **T** is the trivial $\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

As explained elsewhere[2], the principal matrix logarithm of **T** is

$$\operatorname{Log}\mathbf{T} = \begin{pmatrix} \operatorname{Log}\mathbf{S} & \hat{\boldsymbol{\delta}} \\ \mathbf{o}' & \operatorname{Log}1 \end{pmatrix},$$

where $\hat{\boldsymbol{\delta}}$ is a certain 4×1 vector. Writing $\hat{\mathbf{T}} = \operatorname{Log}\mathbf{T}$ and $\hat{\mathbf{S}} = \operatorname{Log}\mathbf{S}$, we abbreviate it as

$$\hat{\mathbf{T}} = \begin{pmatrix} \hat{\mathbf{S}} & \hat{\delta} \\ \mathbf{o}' & 0 \end{pmatrix}, \quad (4)$$

the bottom row now being a row of zeros. The 4×4 submatrix $\hat{\mathbf{S}}$ is Hamiltonian, that is, it obeys

$$\hat{\mathbf{S}}'\mathbf{E} = \mathbf{E}'\hat{\mathbf{S}}. \quad (5)$$

The added fifth row and column make $\hat{\mathbf{T}}$ what we call an *augmented Hamiltonian matrix*. Thus Log transforms a transference into an augmented Hamiltonian matrix.

We now seek other transformations that convert a transference into an augmented Hamiltonian matrix. We consider first transformations that convert a symplectic matrix into a Hamiltonian matrix.

Transformations of symplectic to Hamiltonian matrices

Given a square matrix **A** and a scalar analytic function $f(z)$, we can define the matrix function $f(\mathbf{A})$. Scalar functions such as the inverse of a number, rational functions, exponential, logarithm and cosine may define matrix functions. Linear algebra textbooks[5, 6] include both the general theory associated with these functions and a thorough analysis.

Such functions have two important properties:

$$f(\mathbf{A}') = [f(\mathbf{A})]' \quad (6)$$

and

$$f\left(\mathbf{BAB}^{-1}\right) = \mathbf{B}f(\mathbf{A})\mathbf{B}^{-1} \quad (7)$$

for any non-singular matrix **B**. Assume now that **B** is non-singular and that $f(\mathbf{B})$ and $f(\mathbf{B}^{-1})$ are defined. Then all matrix functions satisfying

$$f(\mathbf{B}^{-1}) = -f(\mathbf{B}) \quad (8)$$

transform symplectic matrices into Hamiltonian matrices. To justify this claim, we assume that **B** is symplectic, that is, $\mathbf{B}'\mathbf{E}\mathbf{B} = \mathbf{E}$, or, equivalently, $\mathbf{E}^{-1}\mathbf{B}'\mathbf{E} = \mathbf{B}^{-1}$, so that we can write $f(\mathbf{E}^{-1}\mathbf{B}'\mathbf{E}) = f(\mathbf{B}^{-1})$. From equations 6, 7 and 8 it follows that

$$\mathbf{E}^{-1}\left[f(\mathbf{B})\right]'\mathbf{E} = -f(\mathbf{B}),$$

that is,

$$\left[f(\mathbf{B})\right]'\mathbf{E} = -\mathbf{E}\,f(\mathbf{B}),$$

which proves that $f(\mathbf{B})$ is Hamiltonian.

Our task now is to look for matrix functions satisfying equation 8. A well known example is the matrix logarithm, because $\mathrm{Log}(\mathbf{B}^{-1}) = -\mathrm{Log}(\mathbf{B})$, for any non-singular **B** with no negative eigenvalue[7]. Another example is the class of rational functions

$$R(\mathbf{A}) = p(\mathbf{A})\left[q(\mathbf{A})\right]^{-1}, \qquad (9)$$

where $p(\mathbf{A})$ and $q(\mathbf{A})$ are matrix polynomials in **A**, both with degree *m*, obeying

$$p(\mathbf{A}^{-1}) = -(\mathbf{A}^m)^{-1}p(\mathbf{A}), \qquad q(\mathbf{A}^{-1}) = (\mathbf{A}^m)^{-1}q(\mathbf{A}). \qquad (10)$$

We note that $R(\mathbf{A})$ is defined for all matrices whose eigenvalues are such that the polynomial $q(z)$ in the denominator does not vanish and that every pair of polynomials in **A** commutes. $R(\mathbf{A})$ satisfies equation 8 because

$$\begin{aligned} R(\mathbf{A}^{-1}) &= p(\mathbf{A}^{-1})\left[q(\mathbf{A}^{-1})\right]^{-1} \\ &= -(\mathbf{A}^m)^{-1}p(\mathbf{A})\left[(\mathbf{A}^m)^{-1}q(\mathbf{A})\right]^{-1} \\ &= -(\mathbf{A}^m)^{-1}p(\mathbf{A})\left[q(\mathbf{A})\right]^{-1}\mathbf{A}^m \\ &= -R(\mathbf{A}). \end{aligned}$$

From equation 9 we can construct infinitely many rational matrix functions that transform symplectic matrices into Hamiltonian matrices. We mention a few examples below.

Example 1: $C(\mathbf{A}) = (\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1}$

This is the so-called *Cayley transform*, which plays an important role in several fields of mathematics and engineering.

Example 2: $C_\alpha(\mathbf{A}) = \alpha(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1}$

where α is a nonzero real number. We call this the *modified Cayley transform*.

Example 3: $f(\mathbf{A}) = (a\mathbf{A}^2 - a\mathbf{I})(b\mathbf{A}^2 + c\mathbf{A} + b\mathbf{I})^{-1}$ where *a*, *b*, *c* are real.

Example 4: $g(\mathbf{A}) = \mathbf{A} - \mathbf{A}^{-1}$.

(Note that $g(\mathbf{A}) = (\mathbf{A}^2 - \mathbf{I})\,\mathbf{A}^{-1}$.)

Transformations of ray transferences to augmented Hamiltonian matrices

The image of a transference **T** under a matrix function *f* is of the form

$$f(\mathbf{T}) = \begin{pmatrix} f(\mathbf{S}) & \hat{\boldsymbol{\delta}} \\ \mathbf{o}' & f(1) \end{pmatrix},$$

where matrix $\hat{\boldsymbol{\delta}}$ is 4×1. If *f* satisfies Equation 8 the image of **S**, *f*(**S**), is a Hamiltonian matrix. Moreover, $f(1) = 0$ (because $f(1^{-1}) = -f(1)$), and then *f*(**T**) is an augmented Hamiltonian matrix. Consequently, all the matrix functions satisfying equation 8 (e.g., matrix logarithm, matrix exponential, Cayley transform, …) transform a ray transference into an augmented Hamiltonian matrix.

Averaging ray transferences

Given a set of ray transferences $\mathbf{T}_1, \ldots, \mathbf{T}_N$, the arithmetic-mean

$$\bar{\mathbf{T}} = \frac{1}{N}\sum_{j=1}^{N} \mathbf{T}_j$$

may not be a ray transference because of the symplectic condition. This has been one of the reasons for the increasing interest in the exp-mean-log-transference

$$\tilde{\mathbf{T}} = \exp\left(\frac{1}{N}\sum_{j=1}^{N} \operatorname{Log} \mathbf{T}_j\right),$$

which gives an average that is in fact a ray transference. Indeed, the top-left 4×4 submatrix of $\tilde{\mathbf{T}}$ obeys the symplectic condition, provided that none of the ray transferences $\mathbf{T}_j$ has a negative eigenvalue. An important feature of the exp-mean-log is that it does not depend on the order in which the $\mathbf{T}_j$ are taken.

In order to be meaningful an average transference must have the same physical meaning for any change in units. This translates in mathematical terms as follows: Consider systems with 4×4 transferences

$$\mathbf{S}_j = \begin{pmatrix} \mathbf{A}_j & \mathbf{B}_j \\ \mathbf{C}_j & \mathbf{D}_j \end{pmatrix}.$$

Suppose they have an average

$$\tilde{\mathbf{S}} = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{C}} & \tilde{\mathbf{D}} \end{pmatrix}.$$

Then for

$$\mathbf{S}_j(\beta) = \begin{pmatrix} \mathbf{A}_j & \mathbf{B}_j/\beta \\ \mathbf{C}_j\beta & \mathbf{D}_j \end{pmatrix},$$

(β is a scalar) the average must be

$$\tilde{\mathbf{S}}(\beta) = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}}/\beta \\ \tilde{\mathbf{C}}\beta & \tilde{\mathbf{D}} \end{pmatrix}.$$

Therefore, a meaningful average **T** of N transferences $\mathbf{T}_1$, $\mathbf{T}_2$, … $\mathbf{T}_N$ should satisfy the following:

(i) **T** is a transference;

(ii) **T** is a function of $\mathbf{T}_1$, $\mathbf{T}_2$, … $\mathbf{T}_N$ that is independent of the order in which the $\mathbf{T}_j$ are taken;

(iii) **T** is invariant under change of units.

Examples of average transferences

Below, we describe three potentially useful means, some of them using the matrix functions addressed in a previous section. We note that some matrix functions (for instance, the attractive function $f(\mathbf{A}) = \mathbf{A} - \mathbf{A}^{-1}$) are a poor choice for defining a mean because they do not have unique inverses.

Example 1: The Cayley Mean

The definition of the Cayley mean requires the inverse of the Cayley transform,

$$C^{-1}(\mathbf{A}) = (\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A})^{-1},$$

which transforms a Hamiltonian matrix **A** (that does not have 1 as an eigenvalue) into a symplectic matrix. Given a set of ray transferences $\mathbf{T}_1$, …, $\mathbf{T}_N$, none of which has -1 as an eigenvalue, the *Cayley mean* is defined by

$$\tilde{\mathbf{T}}_c = C^{-1}\left(\frac{1}{N}\sum_{j=1}^{N} C(\mathbf{T}_j)\right).$$

Example 2: The Cayley Closest Mean

Let $\mathbf{A}^E := \mathbf{E}^{-1}\mathbf{A}'\mathbf{E}$. Every square matrix **A** with even size can be split as a sum of a Hamiltonian matrix and an anti-Hamiltonian matrix:

$$\mathbf{A} = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^E\right) + \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^E\right).$$

The matrix $\mathbf{H} = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^E\right)$ is Hamiltonian (i.e., $\mathbf{H}^E = -\mathbf{H}$) and can be seen as the Hamiltonian part of **A**; it is also the closest Hamiltonian matrix to **A** with respect to some matrix norms.[8] The matrix $\mathbf{M} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^E\right)$ is anti-Hamiltonian (that is, $\mathbf{M}^E = \mathbf{M}$) and is the anti-Hamiltonian part of **A**. Using the Hamiltonian part of a matrix, we can define the *Cayley closest mean* of the ray transferences $\mathbf{T}_1$, …, $\mathbf{T}_N$ by the following steps:

1) Take the arithmetic mean of the ray transferences, obtaining $\overline{\mathbf{T}} = \begin{pmatrix} \overline{\mathbf{S}} & \overline{\boldsymbol{\delta}} \\ \mathbf{o}' & 1 \end{pmatrix}$;

2) If -1 is not an eigenvalue of $\overline{\mathbf{T}}$, compute its Cayley transform $C(\overline{\mathbf{T}}) = \begin{bmatrix} \hat{\mathbf{K}} & \hat{\boldsymbol{\delta}} \\ \mathbf{0}' & 0 \end{bmatrix}$;

3) Since $\hat{\mathbf{K}}$ may not be Hamiltonian, find the closest Hamiltonian matrix to $\hat{\mathbf{K}}$,

$$\hat{\mathbf{H}} = \frac{1}{2}\left(\hat{\mathbf{K}} - \hat{\mathbf{K}}^E\right);$$

4) Let $\hat{\mathbf{X}} = \begin{bmatrix} \hat{\mathbf{H}} & \hat{\boldsymbol{\delta}} \\ \mathbf{0}' & 0 \end{bmatrix}$. Then $\tilde{\mathbf{T}}_{cc} = C^{-1}(\hat{\mathbf{X}})$ satisfies the previously stated conditions (i), (ii) and (iii) for an average to be meaningful.

Example 3: Polar Mean

The mean we propose in this example produces a ray transference without involving the space of augmented Hamiltonian matrices. It is based on the so-called symplectic polar decomposition[8]: Given a nonsingular matrix of even size such that $\mathbf{A}^E\mathbf{A}$ has no negative eigenvalue, there exist a symplectic matrix **S** and an anti-Hamiltonian matrix **M** such that $\mathbf{A} = \mathbf{S}\,\mathbf{M}$. **M** is given explicitly by $\mathbf{M} = \left(\mathbf{A}^E\mathbf{A}\right)^{1/2}$ and **S** by $\mathbf{S} = \mathbf{A}\,\mathbf{M}^{-1} = \mathbf{A}\left[\left(\mathbf{A}^E\mathbf{A}\right)^{1/2}\right]^{-1}$. The notation $\mathbf{X}^{1/2}$ stands for the principal matrix square root of **X**, that is, the unique matrix whose square is **X** and has eigenvalues on the open half plane formed by the complex numbers with positive real parts. In Matlab® the function `sqrtm`(**X**) computes the principal matrix square root of **X**. The *polar mean* is defined as follows:

1) Take the arithmetic mean of the ray transferences, $\overline{\mathbf{T}} = \begin{pmatrix} \overline{\mathbf{S}} & \overline{\boldsymbol{\delta}} \\ \mathbf{o}' & 1 \end{pmatrix}$; recall that $\overline{\mathbf{S}}$ may not be symplectic;

2) If the eigenvalues of $\overline{\mathbf{S}}^E\,\overline{\mathbf{S}}$ are not negative or zero, find the symplectic matrix given by the symplectic polar decomposition of $\overline{\mathbf{S}}$,

$$\mathbf{S} = \overline{\mathbf{S}}\left[\left(\overline{\mathbf{S}}^E\,\overline{\mathbf{S}}\right)^{1/2}\right]^{-1};$$

3) The matrix $\tilde{\mathbf{T}}_p = \begin{pmatrix} \mathbf{S} & \overline{\boldsymbol{\delta}} \\ \mathbf{o}' & 1 \end{pmatrix}$ is a meaningful average of the ray transferences $\mathbf{T}_1$, …, $\mathbf{T}_N$.

Acknowledgements

This paper is based on research supported by the National Research Foundation under grant 2053699 to WF Harris and Institute of Systems and Robotics to JR Cardoso.

References

1. Harris WF. The average eye. *Ophthal Physiol Opt* 2004 **24** 580-585.
2. Harris WF, Cardoso JR. The exponential-mean-log-transference as a possible representation of the optical character of an average eye. *Ophthal Physiol Opt* 2006 **26** 380-383.
3. Harris WF. Quantitative analysis of transformed ray transferences of optical systems in a space of augmented Hamiltonian matrices. *S Afr Optom* 2007 **66**
4. Mathebula SD, Rubin A, Harris WF. Quantitative analysis in Hamiltonian space of the transformed ray transference of a cornea. *S Afr Optom* 2007 **66**
5. Meyer CD. *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.
6. Horn RA, Johnson CR. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991, ch 6.
7. Cardoso JR. *Logaritmos de Matrizes: Aspectos Teóricos e Numéricos.* Doctoral thesis, Department of Mathematics, University of Coimbra, Coimbra, Portugal, 2003.
8. Cardoso JR, Kenney CS, Leite F. Computing the square root and logarithm of a real P-orthogonal matrix., *Appl Numer Math* **46** 2003 173-196.