Symplecticity and relationships among the fundamental properties in linear optics

WF Harris

Department of Optometry, University of Johannesburg, PO Box 524, Auckland Park, 2006 *South Africa*

<wharris@uj.ac.za>

Received 4 January 2010; revised version accepted 5 March 2010

Abstract

Because of symplecticity the four fundamental first-order optical properties of an optical system are not independent. Relationships among them reduce the number of degrees of freedom of a system's transference from 16 to 10. There are many such relationships, they are not easy to remember, they take many forms and they are often needed in derivations. The purpose of this paper is to provide in one place a comprehensive collection of those that have proved useful in linear optics generally and in the context of the eye particularly. The paper also offers aids to memorizing some of the results, derives most of them and along the way introduces the basic notions underlying symplecticity. The relationship to another important class of matrices, the Hamiltonian matrices, is discussed together with their potential role in statistical analysis of the eye. Augmented symplectic matrices are also defined and their relationship to augmented Hamiltonian matrices described. An appendix gives numerical examples of symplectic and Hamiltonian matrices and shows how they may be recognized and constructed. (*S Afr Optom* 2010 **69**(1) 3-13)

Key words: symplecticity, Schur complement, symmetric product, Hamiltonian matrix, augmented symplectic matrix, augmented Hamiltonian matrix

Introduction

Symplecticity is of profound significance to modern science and to optometry in particular. It can even be argued (see towards the end of this paper) that it is the reason why refractive errors can be compensated by means of conventional spherocylindrical lenses and is, therefore, no less than a *sine qua non* of optometry.

Symplecticity implies particular relationships among the fundamental optical properties of an optical system. Although there are several good sources that deal with symplecticity[1-4] they tend to be mathematically sophisticated and not readily accessible for most people working in visual optics. Because the relationships can take many unfamiliar forms, are not easy to remember, and are often needed in analyses of optical problems it would seem useful to have a compact and more accessible summary. Accordingly the objective of this paper is to supply such a summary.

Basic results of linear algebra

We make use of the basic results of linear algebra as presented in introductory texts[5-8]. Our matrices are all real, that is, their entries are all real numbers. In particular for square matrices $\mathbf{A}$ and $\mathbf{B}$

$$(\mathbf{AB}^{\mathrm{T}}) = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \tag{1}$$

where $\mathbf{A}^{\mathrm{T}}$ is the matrix transpose of $\mathbf{A}$. Also if

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}\,, \tag{2}$$

where **I** is an identity matrix, then **B** is the inverse of **A** and is written $\mathbf{A}^{-1}$. Then

$$(\mathbf{AB}^{-1}) = \mathbf{B}^{-1}\mathbf{A}^{-1}, \tag{3}$$

provided the inverses exist. Also

$$(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}. \tag{4}$$

We use the common abbreviation $\mathbf{A}^{-\mathrm{T}}$ for either side of Equation 4. Also

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B} \tag{5}$$

and

$$\det \mathbf{A}^{\mathrm{T}} = \det \mathbf{A}. \tag{6}$$

Partitioned matrices feature importantly in symplecticity. They take the form

$$\mathbf{S} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \tag{7}$$

a matrix of matrices as it were. They are no different from ordinary matrices, however, and obey the usual rules of matrix algebra. Equation 7 does not imply that **S** is 2×2 though it is 2×2 if each of the submatrices **A**, **B**, **C** and **D** is 1×1. **S** is 4×4 if each of the submatrices is 2×2. In optical applications **S** represents the transference of an optical system and the submatrices the (2×2) fundamental linear optical properties of the system.

Multiplication takes the form

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE}+\mathbf{BG} & \mathbf{AF}+\mathbf{BH} \\ \mathbf{CE}+\mathbf{DG} & \mathbf{CF}+\mathbf{DH} \end{pmatrix} \tag{8}$$

as might be expected. Multiplication by a scalar is as expected:

$$s\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} s\mathbf{A} & s\mathbf{B} \\ s\mathbf{C} & s\mathbf{D} \end{pmatrix}. \tag{9}$$

The transposition operator is taken inside the partitioned matrix but the off-diagonal submatrices are also interchanged:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} \mathbf{A}^{\mathrm{T}} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{D}^{\mathrm{T}} \end{pmatrix}. \tag{10}$$

It might be supposed that similar simple expressions can be written for the inverse and determinant of a partitioned matrix; that, however, is not the case in general. (There are expressions but they are much more complicated[9].)

The symplectic unit matrix

Consider the partitioned matrix

$$\mathbf{E} := \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix}. \tag{11}$$

I and **O** are $n\times n$ identity and null matrices respectively. In general n is any positive integer although only $n=1$ and $n=2$ are needed on the optical context; $n=1$ applies in Gaussian optics (the simplest optics in the plane) and $n=2$ applies in linear optics (the simplest optics in three dimensions and the simplest proper approach to astigmatism). **E** itself is $2n\times2n$. It is sometimes called the *symplectic unit matrix*[10]. Although symbols vary, more often than not it is represented by **J**. (We use **E** instead because **J** is already used for one of the basic matrices in the set **I**, **J**, **K** and **L**.)

Applying Equation 8 we find that

$$\mathbf{E}^2 = -\mathbf{I}. \tag{12}$$

(Notice that **I** does not have the same meaning in Equations 11 and 12: in the former it is $n\times n$, in the latter $2n\times2n$.) Hence

$$\mathbf{E}(-\mathbf{E}) = (-\mathbf{E})\mathbf{E} = \mathbf{I}. \tag{13}$$

It follows from Equations 2 and 13 that

$$\mathbf{E}^{-1} = -\mathbf{E}. \tag{14}$$

It follows from Equation 10 that

$$\mathbf{E}^{\mathrm{T}} = -\mathbf{E}. \tag{15}$$

From the definition[5-8] of the determinant it turns out that

$$\det \mathbf{E} = 1. \tag{16}$$

Symplectic matrices

By definition a matrix **S** is symplectic if

$$\mathbf{S}^{\mathrm{T}}\mathbf{ES} = \mathbf{E}. \tag{17}$$

The transference of an optical system is symplectic[1, 4]. Different optical systems can have the same transference. For every 2×2 or 4×4 symplectic matrix it is possible to have an optical system whose transference is that matrix.[11, 12]

Suppose $\mathbf{S}_1$ and $\mathbf{S}_2$ are both symplectic (and have the same n). Consider the product $\mathbf{S}_1\mathbf{S}_2$. Substituting it into the left-hand side of Equation 17 we obtain

$$(\mathbf{S}_1\mathbf{S}_2)^{\mathrm{T}}\mathbf{E}(\mathbf{S}_1\mathbf{S}_2)$$

$$= \mathbf{S}_2^{\mathrm{T}}\mathbf{S}_1^{\mathrm{T}}\mathbf{E}_1\mathbf{S}_2 \text{ by Equation 1}$$

$$= \mathbf{S}_2^{\mathrm{T}}\mathbf{ES}_2 \text{ because } \mathbf{S}_1 \text{ is symplectic}$$

$$= \mathbf{E} \text{ because } \mathbf{S}_2 \text{ is symplectic.}$$

In other words $\mathbf{S}_1\,\mathbf{S}_2$ satisfies Equation 17 and so is symplectic. This means that the product of symplectic matrices is symplectic or, in other words, symplectic matrices are closed under multiplication.

On the other hand symplectic matrices are not in general closed under addition or multiplication by a scalar. (The reader is encouraged to show this by working out some simple examples.) This means, in particular, that an arithmetic average of symplectic matrices is not in general symplectic, a fact that has important implications for basic statistics. If one wants to calculate an average eye, for example, one would want the average to be a possible eye. The fact that the arithmetic average of transferences is not symplectic in general implies that, strictly speaking, it is not meaningful to calculate an average eye that way. The problem of how to calculate an average eye is not a simple one. The relationship between symplectic and Hamiltonian matrices, to be discussed below, appears to offer a solution[13-20]. Despite what we have said here the arithmetic average may, in some cases, be sufficiently close to being symplectic for it to be a good enough approximation.

It is easy to see that **I** and **E** are themselves symplectic: they satisfy Equation 17. On the other hand **O** does not and, therefore, is not symplectic.

Making use of Equations 6, 16 and 17 we find that

$$(\det \mathbf{S})^2 = 1 . \tag{18}$$

This suggests that

$$\det \mathbf{S} = \pm 1 . \tag{19}$$

In fact it turns out that

$$\det \mathbf{S} = 1 . \tag{20}$$

(The proof is not simple and will not be attempted here. Proofs are given elsewhere[2-4].) This means a symplectic matrix **S** is never singular (that is, $\det \mathbf{S} \neq 0$) and its inverse $\mathbf{S}^{-1}$ always exists.

Premutiplying both sides of Equation 17 by $\mathbf{S}^{-\mathrm{T}}$ and postmultiplying by $\mathbf{S}^{-1}$ one obtains

$$\mathbf{E} = (\mathbf{S}^{-1})^{\mathrm{T}} \mathbf{E} \mathbf{S}^{-1} . \tag{21}$$

Comparing this with Equation 17 we see that, if **S** is symplectic, then so is $\mathbf{S}^{-1}$.

Inverting both sides of Equation 17 and premultiplying by **S** and postmultiplying by $\mathbf{S}^{-1}$ one obtains

$$\mathbf{E} = (\mathbf{S}^{\mathrm{T}})^{\mathrm{T}} \mathbf{E} \mathbf{S}^{\mathrm{T}} \tag{22}$$

which shows that $\mathbf{S}^{\mathrm{T}}$ is symplectic.

Let symplectic matrix S be partitioned as in Equation 7. Then, from Equation 10,

$$\mathbf{S}^{\mathrm{T}} = \begin{pmatrix} \mathbf{A}^{\mathrm{T}} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{D}^{\mathrm{T}} \end{pmatrix} . \tag{23}$$

Substitution into Equation 17 results in

$$\begin{pmatrix} -\mathbf{C}^{\mathrm{T}}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{C} & -\mathbf{C}^{\mathrm{T}}\mathbf{B} + \mathbf{A}^{\mathrm{T}}\mathbf{D} \\ -\mathbf{D}^{\mathrm{T}}\mathbf{A} + \mathbf{B}^{\mathrm{T}}\mathbf{C} & -\mathbf{D}^{\mathrm{T}}\mathbf{B} + \mathbf{B}^{\mathrm{T}}\mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix} . \tag{24}$$

Equating the four blocks on the left and right we see that

$$\mathbf{A}^{\mathrm{T}}\mathbf{C} = \mathbf{C}^{\mathrm{T}}\mathbf{A} , \tag{25}$$

$$\mathbf{B}^{\mathrm{T}}\mathbf{D} = \mathbf{D}^{\mathrm{T}}\mathbf{B} , \tag{26}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{D} - \mathbf{C}^{\mathrm{T}}\mathbf{B} = \mathbf{I} \tag{27}$$

and

$$\mathbf{D}^{\mathrm{T}}\mathbf{A} - \mathbf{B}^{\mathrm{T}}\mathbf{C} = \mathbf{I} . \tag{28}$$

Transposition applied to Equation 28 shows that it is equivalent to Equation 27.

When **S** is 2×2 Equations 25 and 26 are trivially true and Equation 28 reduces to Equation 20. If the four entries of a 2×2 matrix are chosen arbitrarily then the matrix is usually not symplectic. To construct a 2×2 symplectic matrix one is free to choose at most three of the entries arbitrarily; the fourth is determined by Equation 20. One can say that symplecticity implies a loss of one degree of freedom from four to three. Note that Equation 20 implies that a 2×2 symplectic matrix cannot have a row or a column of zeros.

When **S** is 4×4 Equations 25 and 26 imply a loss of one degree of freedom each and Equation 27 a loss of four degrees of freedom. Thus, instead of 16 degrees of freedom, a 4×4 symplectic matrix has only 10 degrees of freedom. Constructing a 4×4 symplectic matrix, however, is not easy. A couple of methods will be described later.

Because $\mathbf{S}^{\mathrm{T}}$ is symplectic Equation 22 shows that

$$\begin{pmatrix} -\mathbf{B}\mathbf{A}^{\mathrm{T}} + \mathbf{A}\mathbf{B}^{\mathrm{T}} & -\mathbf{B}\mathbf{C}^{\mathrm{T}} + \mathbf{A}\mathbf{D}^{\mathrm{T}} \\ -\mathbf{D}\mathbf{A}^{\mathrm{T}} + \mathbf{C}\mathbf{B}^{\mathrm{T}} & -\mathbf{D}\mathbf{C}^{\mathrm{T}} + \mathbf{C}\mathbf{D}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix} . \tag{29}$$

Hence

$$\mathbf{AB}^{\mathrm{T}} = \mathbf{BA}^{\mathrm{T}}, \tag{30}$$

$$\mathbf{CD}^{\mathrm{T}} = \mathbf{DC}^{\mathrm{T}}, \tag{31}$$

$$\mathbf{AD}^{\mathrm{T}} - \mathbf{BC}^{\mathrm{T}} = \mathbf{I} \tag{32}$$

and

$$\mathbf{DA}^{\mathrm{T}} - \mathbf{CB}^{\mathrm{T}} = \mathbf{I}. \tag{33}$$

Equation 33 is equivalent to Equation 32.

In order to test whether a given matrix is symplectic or not one can proceed as follows. One can directly test whether it obeys Equation 17. Alternatively one can test whether all three of Equations 25 to 27 (or Equations 30 to 32) are obeyed. In either case if the matrix does obey the equation or equations then it is symplectic; if it does not then it is not symplectic. Numerical examples are presented in the Appendix.

Symmetric products

Applying Equation 1 to the right-hand side of Equation 25 we obtain

$$\mathbf{A}^{\mathrm{T}}\mathbf{C} = (\mathbf{A}^{\mathrm{T}}\mathbf{C})^{\mathrm{T}}. \tag{34}$$

In other words $\mathbf{A}^{\mathrm{T}}\mathbf{C}$ is symmetric. The same approach can be applied to Equations 26, 30 and 31. One reaches the conclusion that, although **A**, **B**, **C** and **D** in **S** (Equation 7) can be symmetric or asymmetric, all of the following products are necessarily symmetric:

$$\mathbf{A}^{\mathrm{T}}\mathbf{C},\ \mathbf{C}^{\mathrm{T}}\mathbf{A},\ \mathbf{B}^{\mathrm{T}}\mathbf{D},\ \mathbf{D}^{\mathrm{T}}\mathbf{B},$$

$$\mathbf{AB}^{\mathrm{T}},\ \mathbf{BA}^{\mathrm{T}},\ \mathbf{CD}^{\mathrm{T}},\ \mathbf{DC}^{\mathrm{T}}.$$

Notice that, in these cases, the two submatrices in a product occupy the same row or the same column of **S**. Furthermore, if the two submatrices share the same column of **S** then the first of the two is transposed; if the two are in the same row then the second is transposed. As an aid to memory one might say 'rows, transpose second; columns, transpose first'. Other products, including $\mathbf{AC}$, $\mathbf{BD}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{T}}\mathbf{B}$, $\mathbf{C}^{\mathrm{T}}\mathbf{B}$, $\mathbf{AB}$, $\mathbf{BC}$, etc., are not symmetric in general.

Postmultiplying Equation 25 by $\mathbf{A}^{-1}$ and premultiplying by $\mathbf{A}^{-\mathrm{T}}$ results in

$$\mathbf{CA}^{-1} = \mathbf{A}^{-\mathrm{T}}\mathbf{C}^{\mathrm{T}} \tag{35}$$

provided **A** is nonsingular. Hence

$$\mathbf{CA}^{-1} = (\mathbf{CA}^{-1})^{\mathrm{T}} \tag{36}$$

showing that $\mathbf{CA}^{-1}$ is symmetric. In the same way one finds other products that are symmetric. Thus

$$\mathbf{AC}^{-1},\ \mathbf{CA}^{-1},\ \mathbf{BD}^{-1},\ \mathbf{DB}^{-1}$$

$$\mathbf{A}^{-1}\mathbf{B},\ \mathbf{B}^{-1}\mathbf{A},\ \mathbf{C}^{-1}\mathbf{D},\ \mathbf{D}^{-1}\mathbf{C},$$

are all necessarily symmetric. One might say 'rows, inverse first; columns, inverse second'. The initial letters spell *rifcis*. Other products, including $\mathbf{A}^{-1}\mathbf{C}$, $\mathbf{B}^{-1}\mathbf{D}$, $\mathbf{AB}^{-1}$, $\mathbf{CD}^{-1}$ are not generally symmetric.

For an eye $\mathbf{B}^{-1}\mathbf{A}$ is the corneal-plane refractive compensation[21] and $\mathbf{A}^{-1}\mathbf{B}$ is the optical structure (a point or interval of Sturm) conjugate to the retina[22]. The negatives of $\mathbf{CA}^{-1}$ and $\mathbf{D}^{-1}\mathbf{C}$ are back- and front-vertex powers of an optical system[23].

The inverted symplectic matrix

Because of Equations 25 to 28 multiplication shows that

$$\begin{pmatrix} \mathbf{D}^{\mathrm{T}} & -\mathbf{B}^{\mathrm{T}} \\ -\mathbf{C}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \mathbf{I} \tag{37}$$

for a symplectic matrix S partitioned as in Equation 7. Similarly, because of Equations 30 to 32, one finds that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\begin{pmatrix} \mathbf{D}^{\mathrm{T}} & -\mathbf{B}^{\mathrm{T}} \\ -\mathbf{C}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} \end{pmatrix} = \mathbf{I}. \tag{38}$$

By Equation 2 then we see that

$$\mathbf{S}^{-1} = \begin{pmatrix} \mathbf{D}^{\mathrm{T}} & -\mathbf{B}^{\mathrm{T}} \\ -\mathbf{C}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} \end{pmatrix}. \tag{39}$$

for any symplectic matrix **S**. Thus the inverse of a symplectic matrix is what one might expect for a 2×2 matrix except that the submatrices are also transposed.

The Schur complements

For **A**, **B**, **C** and **D** the submatrices of a matrix **S** as in Equation 7 the expression $\mathbf{A}-\mathbf{D}^{-\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{C}$ is known as the *Schur complement* of **A** in **S**. Similarly there are Schur complements of **B**, **C** and **D** in **S**. Schur complements arose out of the work of I Schur[24] and are of considerable modern scientific interest[25]. It is no surprise that they should arise in visual optics.

If **S** is symplectic then the Schur complements reduce to particularly neat expressions:

$$\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=\mathbf{D}^{-\mathrm{T}}, \tag{40}$$

$$\mathbf{B}-\mathbf{A}\mathbf{C}^{-1}\mathbf{D}=-\mathbf{C}^{-\mathrm{T}}, \tag{41}$$

$$\mathbf{C}-\mathbf{D}\mathbf{B}^{-1}\mathbf{A}=-\mathbf{B}^{-\mathrm{T}}, \tag{42}$$

$$\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}=\mathbf{A}^{-\mathrm{T}}. \tag{43}$$

These results were apparently first obtained by Dopico and Johnson[26]. In connection with eyes they have been involved in several papers[22, 27-29] although not always recognised as such.

To prove Equation 40 we premultiply each side of Equation 28 by $\mathbf{D}^{-\mathrm{T}}$ to give

$$\mathbf{A}-\mathbf{D}^{-\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{C}=\mathbf{D}^{-\mathrm{T}} \tag{44}$$

and apply Equation 1 to give

$$\mathbf{A}-(\mathbf{B}\mathbf{D}^{-1})^{\mathrm{T}}\mathbf{C}=\mathbf{A}^{-\mathrm{T}}. \tag{45}$$

Equation 40 follows because $\mathbf{B}\mathbf{D}^{-1}$ is symmetric. Equations 41 to 43 follow similarly.

Thus the Schur complements of the block-diagonal submatrices of transference **S** are the transposed inverses of their opposites while the Schur complements of the other submatrices are the negatives of the transposed inverses of their opposites. To recall the sequence of submatrices on the left-hand sides of Equations 40 to 43 the following may be helpful: beginning diagonally across in $\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}$ from the submatrix in question one writes the submatrices in cyclical order going clockwise for the block-diagonal matrices and anticlockwise for the others.

Hamiltonian matrices

There is another class of matrices which is important in modern science and which bears a surprising relationship to symplectic matrices: it is the class of Hamiltonian matrices. A matrix **H** is Hamiltonian if it obeys[10, 30-34]

$$\mathbf{H}^{\mathrm{T}}\mathbf{E}=\mathbf{E}^{\mathrm{T}}\mathbf{H}. \tag{46}$$

E remains the symplectic unit matrix defined by Equation 11 above. In general **H** is $2n\times2n$ but, as with symplectic matrices, we are interested only in Hamiltonian matrices with $n=1$ (Gaussian optics) or $n=2$ (linear optics).

It is a remarkable fact that the principal matrix logarithm of a symplectic matrix is a Hamiltonian matrix and the matrix exponential of a Hamiltonian matrix is symplectic[10, 17, 32-34]. The exponential of a real square matrix X is the real infinite convergent series

$$\mathrm{e}^{\mathbf{X}}:=\sum_{j=0}^{\infty}\frac{1}{j!}\mathbf{X}^{j}. \tag{47}$$

Any matrix **X** which satisfies

$$\mathrm{e}^{\mathbf{X}}=\mathbf{A} \tag{48}$$

is called a logarithm of **A**; in general there is more than one. However, if none of the eigenvalues of **A** is zero or a negative real number, then **A** does have a unique real logarithm the magnitude of whose eigenvalues are less than π. It is the *principal logarithm* and is written $\mathrm{Log}\mathbf{A}$. So, if **S** is symplectic then

$$\mathrm{Log}\mathbf{S}=\mathbf{H} \tag{49}$$

is Hamiltonian and if **H** is Hamiltonian then

$$e^{\mathbf{H}} = \mathbf{S} \tag{50}$$

is symplectic.

It is important to note that these are not the familiar logarithm and exponential simply applied to the entries of the matrix separately. (In Matlab they are given by the functions *logm* and *expm* as opposed to *log* and *exp* which operate separately on the entries of the matrix.) See Example 3 in the Appendix.

Let **H** be a Hamiltonian matrix and s a scalar. It is easy to show that the matrix $s\mathbf{H}$ obeys Equation 46 and so is Hamiltonian. (The left-hands side of Equation 46 equals $s\mathbf{H}^{\mathrm{T}}\mathbf{E}$ and the right-hand side equals $s\mathbf{E}^{\mathrm{T}}\mathbf{H}$ but the two sides are equal because of Equation 46.) Hence Hamiltonian matrices are closed under multiplication by a scalar.

Now suppose $\mathbf{H}_1$ and $\mathbf{H}_2$ are each Hamiltonian and have the same *n*. Then Equation 46 applies to each:

$$\mathbf{H}_1^{\mathrm{T}}\mathbf{E} = \mathbf{E}^{\mathrm{T}}\mathbf{H}_1 \tag{51}$$

and

$$\mathbf{H}_2^{\mathrm{T}}\mathbf{E} = \mathbf{E}^{\mathrm{T}}\mathbf{H}_2 \,. \tag{52}$$

Addition leads to

$$\left(\mathbf{H}_1^{\mathrm{T}} + \mathbf{H}_2^{\mathrm{T}}\right)\mathbf{E} = \mathbf{E}^{\mathrm{T}}\left(\mathbf{H}_1 + \mathbf{H}_2\right) \tag{53}$$

and so

$$\left(\mathbf{H}_1 + \mathbf{H}_2\right)^{\mathrm{T}}\mathbf{E} = \mathbf{E}^{\mathrm{T}}\left(\mathbf{H}_1 + \mathbf{H}_2\right) \tag{54}$$

which shows that $\mathbf{H}_1 + \mathbf{H}_2$ obeys Equation 46. Hence Hamiltonian matrices are closed under addition.

That Hamiltonian matrices are closed under addition and multiplication by a scalar means that the arithmetic average of Hamiltonian matrices is itself Hamiltonian.

It is apparent from Equation 46 that **O** and **E** are Hamiltonian but that **I** is not.

Suppose **H** is partitioned as

$$\mathbf{H} = \begin{pmatrix} \mathbf{M} & \mathbf{N} \\ \mathbf{P} & \mathbf{Q} \end{pmatrix}. \tag{55}$$

If **H** is Hamiltonian then according to Equations 46 and 11

$$\begin{pmatrix} \mathbf{M} & \mathbf{N} \\ \mathbf{P} & \mathbf{Q} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix} = \begin{pmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathbf{M} & \mathbf{N} \\ \mathbf{P} & \mathbf{Q} \end{pmatrix} \tag{56}$$

which multiplies out to

$$\begin{pmatrix} -\mathbf{P}^{\mathrm{T}} & \mathbf{M}^{\mathrm{T}} \\ -\mathbf{Q}^{\mathrm{T}} & \mathbf{N}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} -\mathbf{P} & -\mathbf{Q} \\ \mathbf{M} & \mathbf{N} \end{pmatrix}. \tag{57}$$

This shows that

$$\mathbf{P}^{\mathrm{T}} = \mathbf{P}\,, \tag{58}$$

$$\mathbf{N}^{\mathrm{T}} = \mathbf{N} \tag{59}$$

and

$$\mathbf{Q} = -\mathbf{M}^{\mathrm{T}}\,. \tag{60}$$

In other words, for a Hamiltonian matrix partitioned as in Equation 55, the off-diagonal submatrices **N** and **P** are necessarily symmetric and one diagonal submatrix is the negative of the transpose of the other.

For a 2×2 Hamiltonian matrix Equations 58 and 59 are trivially true and Equation 60 is equivalent to the statement that

$$\mathrm{tr}\mathbf{H} = 0 \tag{61}$$

or, in other words, that a Hamiltonian matrix has zero trace or is traceless as it is sometimes called. Thus, as in the case of symplectic matrices, there is a loss of one degree of freedom from four to three. In fact the 2×2 Hamiltonian matrices define a three-dimensional vector or linear space which means that it is possible to draw three-dimensional graphs representing the space. Individual matrices can be plotted in the space to form trajectories and clusters. An arithmetic mean is a point in the same space at the centre of a cluster. 3×3 variance-covariance matrices provide measures of spread and variation in the space.

For 4×4 Hamiltonian matrices Equations 58 and

59 each represent a loss of a degree of freedom and Equation 60 a loss of four degrees of freedom. Thus, also as for symplectic matrices, there is a loss of six degrees of freedom from 16 to 10. The 4×4 Hamiltonian matrices define a 10-dimensional vector space. Although one cannot make proper drawings of such a space one can certainly work in the space mathematically and calculate arithmetic means and 10×10 variance-covariance matrices.[18-20] (Equation 61 is, in fact, true for Hamiltonian matrices of any size.)

In order to test whether or not a particular matrix is Hamiltonian one can check whether or not it obeys Equation 46. Or it may be easy to check whether all three Equations 58 to 60 are obeyed, that is, whether the two off-diagonal submatrices are symmetric and one diagonal submatrix is the negative of the transpose of the other. See the Appendix for examples.

Constructing a Hamiltonian matrix is just as easy. One chooses **N** and **P** in Equation 55 arbitrarily except that they must be symmetric. One can then choose **M** arbitrarily in which case **Q** is the negative of the transpose of **M**.

How to construct a symplectic matrix

One occasionally wishes to construct a symplectic matrix. As mentioned above this is not difficult if the matrix is 2×2: one can usually chose any three of the four entries and then calculate the fourth from the requirement that the matrix has a unit determinant (Equation 20). The task is much harder if the matrix is 4×4. We describe two methods below.

Consider the matrix $\begin{pmatrix} \mathbf{I} & \zeta\mathbf{I} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$ where ζ is a scalar. Testing shows that it is symplectic. Now consider the matrix $\begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{C} & \mathbf{I} \end{pmatrix}$. In general it is not symplectic because Equation 25 fails. However if we restrict **C** to being symmetric then we see that all of Equations 25 to 27 are obeyed and so the matrix is symplectic. Because symplectic matrices are closed under multiplication we can construct symplectic matrices by multiplying any number of matrices of these two forms. (The first matrix is the transference of a homogeneous gap and the second that of a thin system. Instead of the first of these two matrices one can also use any matrix of the form $\begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$ where **B** is symmetric but an optical system with this transference is not simple.)

A second method of constructing a symplectic matrix can be quicker and easier. We first construct a Hamiltonian matrix as described above and then take its matrix exponential using Matlab. The result is symplectic. See Example 2 in the Appendix.

Augmented symplectic matrices

For heterocentric systems it is sometimes convenient to work with augmented symplectic matrices. They take the form

$$\mathbf{T} = \begin{pmatrix} \mathbf{S} & \boldsymbol{\delta} \\ \mathbf{o}^{\mathrm{T}} & 1 \end{pmatrix} \tag{62}$$

where **S** is symplectic. Thus an augmented symplectic matrix is a symplectic matrix with an additional right-hand column and an additional trivial bottom row. If **S** is 4×4 then **T** is 5×5. Submatrix $\boldsymbol{\delta}$ is any 4×1 matrix and **o** is the 4×1 null matrix. The bottom row of **T** consists of four 0s and a 1. For every augmented symplectic matrix there corresponds an optical system.[12]

From the definition of the determinant $\det\mathbf{T} = 1\times\det\mathbf{S}$. Hence from Equation 20

$$\det\mathbf{T} = 1. \tag{63}$$

$\mathbf{T}^{\mathrm{T}}$ is not an augmented symplectic matrix.

Multiplication according to Equation 2 shows that

$$\mathbf{T}^{-1} = \begin{pmatrix} \mathbf{S}^{-1} & -\mathbf{S}^{-1}\boldsymbol{\delta} \\ \mathbf{o}^{\mathrm{T}} & 1 \end{pmatrix} \tag{64}$$

which is an augmented symplectic matrix.

The product of two augmented symplectic matrices is an augmented symplectic matrix as the following shows:

$$\mathbf{T}_1\mathbf{T}_2 = \begin{pmatrix} \mathbf{S}_1\mathbf{S}_2 & \mathbf{S}_1\boldsymbol{\delta}_2 + \boldsymbol{\delta}_1 \\ \mathbf{o}^{\mathrm{T}} & 1 \end{pmatrix}. \tag{65}$$

Thus, like symplectic matrices, augmented symplectic matrices are closed under matrix multiplication. They are also not closed under addition or under multiplication by a scalar.

An augmented symplectic matrix can be partitioned as

$$\mathbf{T}=\begin{pmatrix}\mathbf{A} & \mathbf{B} & \mathbf{e}\\ \mathbf{C} & \mathbf{D} & \boldsymbol{\pi}\\ \mathbf{o}^{\mathrm{T}} & \mathbf{o}^{\mathrm{T}} & 1\end{pmatrix}. \qquad (66)$$

It follows from Equations 39 and 64 that

$$\mathbf{T}^{-1}=\begin{pmatrix}\mathbf{D}^{\mathrm{T}} & -\mathbf{B}^{\mathrm{T}} & -\mathbf{D}^{\mathrm{T}}\mathbf{e}+\mathbf{B}^{\mathrm{T}}\boldsymbol{\pi}\\ -\mathbf{C}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} & \mathbf{C}^{\mathrm{T}}\mathbf{e}-\mathbf{A}^{\mathrm{T}}\boldsymbol{\pi}\\ \mathbf{o}^{\mathrm{T}} & \mathbf{o}^{\mathrm{T}} & 1\end{pmatrix}. \qquad (67)$$

Because submatrix **S** of augmented symplectic matrix **T** is symplectic all of the results for symplectic matrices (symmetric products, Schur complements, etc.) apply in the context of augmented symplectic matrices as well.

Augmented Hamiltonian matrices

As for symplectic matrices one can define an augmented Hamiltonian matrix to be a Hamiltonian matrix with an additional right-hand column and an additional bottom row. An augmented Hamiltonian matrix takes the form

$$\mathbf{G}=\begin{pmatrix}\mathbf{H} & \boldsymbol{\beta}\\ \mathbf{o}^{\mathrm{T}} & 0\end{pmatrix} \qquad (68)$$

where **H** is Hamiltonian and $\boldsymbol{\beta}$ is arbitrary. The bottom row is a row of zeros.

It is obvious that augmented Hamiltonian matrices are closed under addition and multiplication by a scalar. Thus the arithmetic average of augmented Hamiltonian matrices is augmented Hamiltonian.

Augmented Hamiltonian matrices bear the same relationship to augmented symplectic matrices as Hamiltonian matrices do to symplectic matrices.[18] That is, if **T** is augmented symplectic then

$$\operatorname{Log}\mathbf{T}=\mathbf{G} \qquad (69)$$

is augmented Hamiltonian and if G is augmented Hamiltonian then

$$\mathrm{e}^{\mathbf{G}}=\mathbf{T} \qquad (70)$$

is augmented symplectic.

If Hamiltonian matrix **H** is 4×4 then augmented Hamiltonian matrix **G** is 5×5. The bottom row being trivial **G** has 20 nontrivial entries. Submatrix $\boldsymbol{\beta}$ adds four degrees of freedom to the 10 of **H**. Thus **G** has 14 degrees of freedom.

The augmented Hamiltonian matrices define a 14-dimensional vector space in which one can calculate arithmetic means and 14×14 variance-covariance matrices.[19]

Concluding remarks

The intentions here have been to bring together in one place basic results in the context of symplecticity which continue to be of use in work in the optics of vision.

From the definition of a symplectic matrix we see that the inverse and transpose of a symplectic matrix are also symplectic and that symplectic matrices (of the same size) are closed under matrix multiplication but not under addition or multiplication by a scalar.

Although the fundamental properties may be symmetric or asymmetric symplecticity makes certain pairwise products necessarily symmetric. We now have the following rule: the product of two properties in the same row of a transference is symmetric if the second of the pair is transposed, and the product of two properties in the same column is symmetric if the first is transposed. Thus $\mathbf{CD}^{\mathrm{T}}$ (same row) and $\mathbf{C}^{\mathrm{T}}\mathbf{A}$ (same column), for example, are symmetric while $\mathbf{D}^{\mathrm{T}}\mathbf{C}$, $\mathbf{AC}^{\mathrm{T}}$, $\mathbf{DC}$ and $\mathbf{AT}$, for example, may or may not be symmetric.

The product of a fundamental property and the inverse of another fundamental property is also symmetric provided the two properties are in the same row with the first inverted or in the same column with the second inverted. For example $\mathbf{D}^{-1}\mathbf{C}$ and $\mathbf{CA}^{-1}$ (the negatives of the front- and back-vertex powers of a system[23]) are symmetric while $\mathbf{CD}^{-1}$ and $\mathbf{A}^{-1}\mathbf{C}$ are not generally symmetric.

Of course where inverses are involved the equations are meaningful only if the matrices in question are nonsingular. Furthermore one needs to be aware in numerical work that inversion of nearly-singular matrices can lead to spurious results.

$\mathbf{B}^{-1}\mathbf{A}$ is the corneal plane refractive compensation for an eye.[21] If it were not symmetric it would not be possible properly to compensate for the refractive error with thin lenses. But $\mathbf{B}^{-1}\mathbf{A}$ is symmetric because of symplecticity and so we have the remarkable fact that it is possible to compensate for the refractive error of an eye whose power ($\mathbf{F} = -\mathbf{C}$) is asymmetric[35] by means of thin lenses (the usual spherocylindrical lenses of optometry) whose powers are symmetric. It would seem, therefore, that were it not for symplecticity there might have been no optometry!

An subsequent paper[22] uses many of the results given here including all the Schur complements (Equations 40 to 43) and many of the symmetric products.

In numerical work it is often useful to be able to recognize and construct symplectic and Hamiltonian matrices. This paper shows how. Numerical examples are treated in the Appendix.

Acknowledgements

I gratefully acknowledge financial support from the National Research Foundation. I thank R D van Gool, T Evans and A Rubin for discussions and for commenting on the manuscript.

References

1. Guillemin V, Sternberg S. *Symplectic Techniques in Physics.* Cambridge University Press, Cambridge, 1984.
2. Serre D. *Matrices: Theory and Applications.* Springer-Verlag, New York, 2002.
3. Kauderer M. *Symplectic Matrices: First Order Systems and Special Relativity.* World Scientific, Singapore, 1994.
4. McDuff D, Salamon D. *Introduction to Symplectic Topology.* 2nd ed, Oxford University Press, Oxford, 1998.
5. Anton H, Rorres C. *Elementary Linear Algebra:* Applications Version. 9th ed, Wiley, Hoboken NJ, 2005.
6. Hohn FE. *Elementary Linear Algebra*. Macmillan, New York, 1973. Republished by Dove, Mineola NY, 2002.
7. Meyer CD. *Matrix Analysis and Applied Linear Algebra.* SIAM, Philadelphia, 2000.
8. Schneider H, Barker GP. Matrices and Linear Algebra. 2nd ed, Holt, Rinehart and Winston, New York, 1973. Republished by Dover, Mineola NY, 1989.
9. Graybill FA. *Matrices with Applications in Statistics*. 2nd ed, Wadsworth, Belmont CA, 1983, 183-201.
10. Torre A. *Linear Ray and Wave Optics in Phase Space.* Elsevier, Amsterdam, 2005, 11, 192.
11. Sudarshan ECG, Makunda N, Simon R. Realization of first order optical systems using thin lenses. *Optica Acta* 1985 **2** 855-872.
12. Harris WF. Realizability of systems of given linear optical character. *Optom Vis Sci* 2004 **81** 807-809.
13. Harris WF. The average eye. *Ophthal Physiol Opt* 2004 **24** 580-585.
14. van Gool RD, Harris WF. The concept of the average eye. *Optom Vis Sci* 2004 **81** (12S) 163.
15. van Gool RD, Harris WF. The concept of the average eye. *S Afr Optom* 2005 **64** 38-43.
16. Harris WF. The log-transference and an average Gaussian eye. *S Afr Optom* 2005 **64** 84-88.
17. 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.
18. Cardoso JR, Harris WF. Transformations of ray transferences of optical systems to augmented Hamiltonian matrices and the problem of the average system. *S Afr Optom* 2007 **66** 56-61.
19. Harris WF. Quantitative analysis of transformed transferences of optical systems in a space of augmented Hamiltonian matrices. *S Afr Optom* 2007 **66** 62-67.
20. Mathebula SD, Rubin A, Harris WF. Quantitative analysis in Hamiltonian space of the ray transference of a cornea. *S Afr Optom* 2007 **66** 68-76.
21. Harris WF. A unified paraxial approach to astigmatic optics. *Optom Vis Sci* 1999 **76** 480-499.
22. Harris WF. Special rays and structures in general optical systems. S Afr Optom 2010 69 submitted herewith.
23. Harris WF. Back- and front-vertex powers of astigmatic systems. *Optom Vis Sci* 2010 **87** 70-72.
24. Schur I. Über Potenzreihen, die im Innern des Einheitkreises beschränkt sind. *Reine Angew Math* 1917 **147** 205-232.
25. Zhang F. Editor of: *The Schur Complement and Its Applications.* Springer, New York, 2005.
26. Dopico FM, Johnson CJ. Complementary bases in symplectic matrices and a proof that their determinant is one. *Linear Algebra Its Appl* 2006 **419** 772-778.
27. 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.
28. Harris WF. Nodes and nodal points and lines in eyes and other optical systems. *Ophthal Physiol Opt* 2010 **30** 24-42.
29. Harris WF. Cardinal points and generalizations. *Ophthal Physiol Opt* 2010 **30** submitted.
30. Higham NJ. *Functions of Matrices: Theory and Computation.* SIAM, Philadelphia, 2008, 323.
31. Bernstein DS. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory.* Princeton Univ Press, Princeton, 2005, 85.
32. Horn RA, Johnson CR. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.
33. Cardoso JR. *Logaritmos de Matrizes: Aspectos Teóricos e Numéricos.* Doctoral thesis, Department of Mathematics, University of Coimbra, Coimbra, Portugal, 2003.
34. Cardoso JR. An explicit formula for the matrix logarithm. *S Afr Optom* 2005 **64** 80-83.
35. Harris WF. Dioptric power: its nature and its representation in three- and four-dimensional space. *Optom Vis Sci* 1997 **74** 349-366.

Appendix

We illustrate here recognition (Example 1) and construction (Example 2) of symplectic and Hamiltonian matrices. Example 3 compares the logarithm of the entries of a particular symplectic matrix with the matrix logarithm of the matrix.

Example 1 Classify each of the following as symplectic, Hamiltonian or both:

(a) $\begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 4 & 0 & 0 & 6 \\ -1 & 3 & 6 & 1 \\ 2 & -1 & -4 & 1 \\ 0 & 2 & 0 & -3 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & 0 & 0 & 6 \\ -1 & 3 & 6 & 1 \\ 2 & -1 & -4 & 1 \\ -1 & 2 & 0 & -3 \end{pmatrix}$

(f) $\begin{pmatrix} 4 & 0 & 0 & 6 \\ -1 & 3 & 6 & 1 \\ 2 & -1 & -4 & 0 \\ -1 & 2 & 1 & -3 \end{pmatrix}$

Consider the 2×2 matrices first. All we have to do is check the determinant and the trace: if the determinant is 1 then the matrix is symplectic; if the trace is zero then the matrix is Hamiltonian. We see that (a) is both symplectic and Hamiltonian. (b) is Hamiltonian and, because its determinant is not 1, it is not symplectic. (c) is neither symplectic nor Hamiltonian. Thus (a) could be the transference of an optical system but (b) and (c) could not. (a) and (b) could both be the log-transferences of optical systems.

Consider now the 4×4 matrices. We check first for Hamiltonicity. Partitioning according to Equation 55 we observe that submatrix $\mathbf{P}$ is not symmetric in (d). Hence (d) is not Hamiltonian. $\mathbf{N}$ and $\mathbf{P}$ are symmetric in (e) and $\mathbf{Q}$ is the negative of the transpose of $\mathbf{M}$; hence (e) is Hamiltonian. $\mathbf{Q}$ is not the negative of the transpose of $\mathbf{M}$ in (f) and so (f) is not Hamiltonian. Checking for symplecticity is not as easy. We resort to substituting into the left-hand side of Equation 17 and multiplying. For (d) we finally obtain

$$\begin{pmatrix} 0 & -6 & -16 & -5 \\ 6 & 0 & -12 & -5 \\ 16 & 12 & 0 & 6 \\ 5 & 5 & -6 & 0 \end{pmatrix}$$

which is not $\mathbf{E}$. Hence (d) is not symplectic. We also do not obtain $\mathbf{E}$ for either (e) or (f). Thus none of the matrices is symplectic. Thus none of (d), (e) and (f) could be the transference of an optical system but (e) could be the log-transference of one.

Example 2 Starting with each of the Hamiltonian matrices in Exercise 1 construct a symplectic matrix.

(a), (b) and (e) are the only Hamiltonian matrices in Exercise 1. Using Matlab we obtain the matrix exponentials. They turn out to be

(a) $\begin{pmatrix} 1.382 & 1.683 \\ -0.841 & -0.301 \end{pmatrix}$,

(b) $\begin{pmatrix} 2.718 & 0 \\ -1.175 & 0.368 \end{pmatrix}$

and

$$\text{(e)}\begin{pmatrix} 25.078 & 34.249 & 24.553 & 31.921 \\ 4.858 & 16.526 & 14.041 & 8.604 \\ 5.955 & 6.096 & 3.985 & 7.062 \\ -2.844 & 0.840 & 1.512 & -2.363 \end{pmatrix}$$

approximately.

Example 3 Compare the logarithm applied separately to the entries of the symplectic matrix $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ with the principal matrix logarithm of the matrix.

They are $\begin{pmatrix} 0 & 0.6932 \\ 0 & 1.0986 \end{pmatrix}$ and $\begin{pmatrix} -0.7603 & 1.5207 \\ 0.7603 & 0.7603 \end{pmatrix}$ respectively.