BILINEAR CONSTRAINTS BETWEEN ELEMENTS OF THE 4 X 4 MUELLER-JONES TRANSFER MATRIX OF POLARIZATION THEORY

Richard BARAKAT

Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA and Bolt Beranek and Newman Inc., Cambridge, MA 02238, USA

Received 2 March 1981

When the 4 X 4 Mueller matrix of an optical system represents a situation in which the polarization of the light is unchanged, then the elements of the Mueller matrix are derivable from a 2 X 2 Jones matrix. The general Mueller matrix contains 16 independent parameters, whereas the Jones matrix contains 7 independent parameters. Consequently, 9 identities exist between the 16 matrix elements. The purpose of this note is to display explicitly these 9 nonlinear (bilinear) equations. These equations are then used to study the number of independent parameters of a recent experimental determination by Howell of the Mueller matrix of a collimator-radiometer system.

Consider a non-image forming optical system (e.g., scattering system, crystalline plates, etc.) and let the input Stokes parameters $(s_0, s_1, s_2, s_3)^+ = \mathsf{S}$ and the Stokes parameters $(s_0', s_1', s_2', s_3')^+ = \mathsf{S}'$ be linearly related.

$$\mathsf{S}' = \mathsf{MS}\,, \tag{1}$$

where M is a 4 X 4 matrix with 16 real matrix elements, now commonly termed the Mueller matrix [1]. M, in general, contains 16 independent parameters. However under various symmetry conditions the number of independent parameters can be less than 16 [2]. In the special, but important case where the optical system represented by M does not depolarize, then as Jones [3] has pointed out, 9 identities exist among the 16 matrix elements of M so that only 7 of them are independent. Furthermore, in this special case, M can be expressed in terms of the 2 X 2 Jones matrix $\mathsf{T_J}$. Consequently it should be possible to obtain these 9 equations explicitly by using properties of physically realizable $\mathsf{T_J}$ matrices and relating them to the M matrix. The purpose of the present communication is to carry out such an analysis and explicitly display the 9 nonlinear (bilinear) equations.

We will let $\mathsf{T_M}$ denote the Mueller matrix derived from the Jones matrix $\mathsf{T_J}$. $\mathsf{T_M}$, termed the Mueller–Jones matrix, can be expressed in terms of $\mathsf{T_J}$ by the known formula [4]

$$(\mathsf{T_M})_{ij} = \tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{\sigma}_i \mathsf{T_J} \boldsymbol{\sigma}_j \mathsf{T_J^\dagger}), \quad (i,j = 0,1,2,3)\,, \tag{2}$$

where $\boldsymbol{\sigma}$'s are the Pauli matrices. However a more convenient expression for our purpose is

$$\mathsf{T_M} = \mathsf{A}(\mathsf{T_J} \otimes \mathsf{T_J^\dagger})\mathsf{A}^{-1}\,, \tag{3}$$

where $\otimes$ denotes the Kronecker matrix product and

$$\mathsf{A} \equiv \begin{vmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{vmatrix}\,. \tag{4}$$

Note that the matrix A differs somewhat from the corresponding matrix T in [5] and A in [6]. This is due to the fact that we use the standard version of $\boldsymbol{\sigma}_3$.

The output coherency matrix $\boldsymbol{\Phi}'$ is related to the input coherency matrix $\boldsymbol{\Phi}$ for a system governed by a (physically realizable) Jones matrix $\mathsf{T_J}$ by [7]

$$\boldsymbol{\Phi}' = \mathsf{T_J}\boldsymbol{\Phi}\mathsf{T_J^\dagger}, \quad \det \mathsf{T_J} \neq 0\,, \tag{5}$$

where we have omitted the explicit dependence of these matrices on frequency. The input coherency matrix (spectral density matrix) is

$$\mathbf{\Phi} = \begin{vmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{vmatrix} = \begin{vmatrix} \phi_{11} & c_{21} - \mathrm{i}q_{21} \\ c_{21} + \mathrm{i}q_{21} & \phi_{22} \end{vmatrix} . \quad (6)$$

Here ϕ_{11} and ϕ_{22} are the power spectral densities associated with the first and second components; and ϕ_{12} $(= \phi_{21}^*)$, ϕ_{21} are the cross power spectral densities. It is usually more convenient to work with the corresponding cospectral density c_{21} and quadspectral density q_{21}, both of which are real (see ref. [7] for more details). The Stokes parameters $s_l (l = 0, 1, 2, 3)$ are given by

$$s_0 = \tfrac{1}{2}(\phi_{11} + \phi_{22}), \quad s_1 = \tfrac{1}{2}(\phi_{11} - \phi_{22}) ,$$
$$s_2 = c_{21}, \qquad s_3 = q_{21} . \quad (7)$$

In ref. [7], the Stokes parameters are termed x_1, x_2, x_3, x_4, however it is more convenient to employ the standard s-notation and terminology for the Stokes parameters. The coherency matrix can be expressed directly in terms of the Stokes parameters

$$\mathbf{\Phi} = \begin{vmatrix} s_0 + s_1 & s_2 - \mathrm{i}s_3 \\ s_2 + \mathrm{i}s_3 & s_0 - s_1 \end{vmatrix} , \quad (8)$$

The corresponding expression for the output coherency matrix $\mathbf{\Phi}'$ follows by putting primes on the matrix elements of the input coherency matrix.

The output and input Stokes parameters are related by the quadratic forms

$$(s_0'^2 - s_1'^2 - s_2'^2 - s_3'^2)$$
$$= |\det \mathrm{T_J}|^2 (s_0^2 - s_1^2 - s_2^2 - s_3^2) , \quad (9)$$

obtained from eq. (5) by taking determinants of both sides. As noted in ref. [7], eq. 9 is essentially a statement of the invariance of the Stokes parameters under a Lorentz transformation.

To proceed further, we recall some facts about Lorentz transformations pertinent to our needs. Consider the quadratic form

$$\mathrm{X^+GX} = x_0^2 - x_1^2 - x_2^2 - x_3^2 , \quad (10)$$

where

$$\mathrm{X} = \begin{vmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{vmatrix} , \quad \mathrm{G} = \begin{vmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{vmatrix} \quad (11)$$

A real homogeneous linear transformation

$$\mathrm{X_2} = \Lambda \mathrm{X_1} \quad (12)$$

is termed a Lorentz transformation if it leaves the quadratic form $\mathrm{X^+GX}$ invariant (i.e., if $\mathrm{X_2^+GX_2} = \mathrm{X_1^+GX_1}$). The 4×4 matrix Λ is called the Lorentz matrix. The criterion for an arbitrary 4×4 real matrix to be Lorentz is [8,9]

$$\Lambda^+ \mathrm{G} \Lambda = \hat{\mathrm{G}} . \quad (13)$$

We now show that Λ can be written in terms of $\mathrm{T_M}$ so that eq. (13) will furnish the various nonlinear relations between the matrix elements of $\mathrm{T_M}$.

The possible Lorentz matrices Λ form a group. We are interested only in a subgroup called the proper orthochronous subgroup L_+ consisting of those Λ which obey the conditions det $\Lambda = 1$ and $\mu_{00} > 0$. The reasons for requiring these conditions is discussed in the Appendix. Every Λ matrix in L_+ can be written as

$$\Lambda = \mathrm{A}(T_\mathrm{J} \otimes T_\mathrm{J}^+)\mathrm{A}^{-1} , \quad (14)$$

where T_J is an unimodular 2×2 matrix and A is given by eq. (4). The expression for $\mathrm{T_M}$, given by eq. (3), is essentially the same as the right hand side of eq. (14). In order to express $\hat{\mathrm{T}}_\mathrm{M}$ in terms of T_J, we set

$$T_\mathrm{J} = \frac{1}{(\det \mathrm{T_J})} \mathrm{T_J} \quad (15)$$

so that det $T_\mathrm{J} = 1$. Upon equating eqs. (13) and (14), we have

$$\mathrm{T_M} = |\det \mathrm{T_J}|^2 \Lambda . \quad (16)$$

Consequently, the basic condition for Λ to be Lorentz, eq. (13) now reads

$$\mathrm{T_M^+GT_M} = |\det \mathrm{T_J}|^4 \mathrm{G} \quad (17)$$

in terms of $\mathrm{T_M}$.

Upon equating matrix elements of both sides of eq. (17) we obtain sixteen bilinear relations between the sixteen matrix elements of $\mathrm{T_M}$. The off-diagonal elements furnish twelve equations of which six are redundant, the remaining six are:

$$\mu_{00}\mu_{01} - \mu_{10}\mu_{11} - \mu_{20}\mu_{21} - \mu_{30}\mu_{31} = 0\,, \tag{18a}$$

$$\mu_{00}\mu_{02} - \mu_{10}\mu_{12} - \mu_{20}\mu_{22} - \mu_{30}\mu_{32} = 0\,, \tag{18b}$$

$$\mu_{00}\mu_{03} - \mu_{10}\mu_{13} - \mu_{20}\mu_{23} - \mu_{30}\mu_{33} = 0\,, \tag{18c}$$

$$\mu_{01}\mu_{02} - \mu_{11}\mu_{12} - \mu_{21}\mu_{22} - \mu_{31}\mu_{32} = 0\,, \tag{18d}$$

$$\mu_{01}\mu_{03} - \mu_{11}\mu_{13} - \mu_{21}\mu_{23} - \mu_{31}\mu_{33} = 0\,, \tag{18e}$$

$$\mu_{02}\mu_{03} - \mu_{12}\mu_{13} - \mu_{22}\mu_{23} - \mu_{32}\mu_{33} = 0\,. \tag{18f}$$

The diagonal elements furnish four addition relations of which three are independent

$$\mu_{01}^2 - \mu_{11}^2 - \mu_{21}^2 - \mu_{31}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0\,, \tag{19a}$$

$$\mu_{02}^2 - \mu_{12}^2 - \mu_{22}^2 - \mu_{32}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0\,, \tag{19b}$$

$$\mu_{03}^2 - \mu_{13}^2 - \mu_{23}^2 - \mu_{33}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0\,. \tag{19c}$$

In deriving the last three relations, we used

$$\mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = |\det \mathrm{T_J}|^4\,, \tag{20}$$

which is a statement of the equality of the (00) matrix elements. Thus we have obtained nine nonlinear equations between the sixteen $\mathrm{T_M}$ elements. Consequently there are seven independent parameters left to characterize the Mueller-Jones matrix $\mathrm{T_M}$.

As an application of these bilinear relations between the matrix elements, consider the recent experimental determination by Howell [10] of the Mueller matrix of a collimator-radiometer system. He obtains

$$\mathrm{M} = \begin{vmatrix} 0.7599 & -0.0623 & 0.0295 & 0.1185 \\ -0.0573 & 0.4687 & -0.1811 & -0.1863 \\ 0.0384 & -0.1714 & 0.5394 & 0.0282 \\ 0.1240 & -0.2168 & -0.0120 & 0.6608 \end{vmatrix} \tag{21}$$

(see eq. (6) of his paper). The question we wish to ask is: does this matrix contain 16 independent parameters? Direct calculation of the left hand sides of eqs. (18) and (19) yields

LHS(18a) = 0.0129, LHS(18d) = 0.1729 ,

LHS(18b) = –0.0072, LHS(18e) = 0.2280 ,

LHS(18c) = –0.0036, LHS(18f) = –0.0375 ,

and

LHS(19a) = 0.2617 . LHS(19b) = 0.2342 ,

LHS(19c) = 0.0711 .

It would appear that 4 of the eqs. (18) are approximately zero, whereas all of the eqs. (19) are nonzero. Consequently, only 4 of the 9 bilinear relations are satisfied and thus M as given in eq. (21) cannot be a Mueller–Jones matrix. The number of independent parameters characterizing Howell's matrix is (16 – 4) = 12.

Appendix

The Lorentz group leaves the quadratic form $\mathrm{X^+GX}$ invariant, but such a group is wider than our situation permits. The transformation

$$s_0 \to s_0', \quad s_1 \to s_1', \quad s_2 \to s_2', \quad s_3 \to s_3' \tag{A.1}$$

leaves $\mathrm{X^+GX}$ invariant but changes the intensity (an intrinsically nonnegative quantity) to a negative quantity. This is easily avoided by requiring that μ_{00} in Λ, or equivalently in $\mathrm{T_M}$, be nonnegative.

The other transformation we want to exclude is the inversion

$$s_0 \to s_0', \quad s_1 \to -s_1', \quad s_2 \to -s_2', \quad s_3 \to -s_3'\,. \tag{A.2}$$

This happens when the input space coordinate system is taken to be right handed and the output space coordinate system is taken to be left handed (or vice versa). Imposition of the constraint det $\Lambda = +1$ forbids this situation.

This work was supported by the Naval Research Laboratory under Contract No. N00-173-79-C-0474.

[1] P. Soleillet, Ann. Phys. 12 (1929) 23.
[2] F. Perrin, J. Chem. Phys. 10 (1942) 415.
[3] R.C. Jones, J. Opt. Soc. Am. 37 (1947) 107.
[4] J.W. Simmons and M. Guttmann, States, waves and photons (Addison-Wesley, Reading, Mass., 1970) p. 79.
[5] E.L. O'Neil, Introduction to statistical optics (Addison-Wesley, Reading, Mass., 1963) p. 143.
[6] R.M.A. Azzam and N.M. Bashara, Ellipsometry and polarized light (North-Holland, New York, 1977) p. 149.
[7] R. Barakat, J. Opt. Soc. Am. 53 (1963) 317.
[8] M. Carmelli and M. Shimon, Representations of the rotation and Lorentz groups (Dekker, New York, 1976).
[9] F.D. Murnaghan, The theory of group representations (Johns Hopkins Univ. Press, Baltimore, 1938) Chap. 9.
[10] B.J. Howell, Appl. Optics 18 (1979) 809.