An eigenvalue criterion for matrices transforming Stokes parameters

C. V. M. van der Mee[a)]
Department of Physics and Astronomy, Free University, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

(Received 3 March 1993; accepted for publication 30 June 1993)

Simple eigenvalue tests are given to ascertain that a given real 4×4 matrix transforms the four-vector of Stokes parameters of a beam of light into the four-vector of Stokes parameters of another beam of light, and to determine whether a given 4×4 matrix is a weighted sum of pure Mueller matrices. The latter result is derived for matrices satisfying a certain symmetry condition. To derive these results indefinite inner products are applied.

I. INTRODUCTION

The intensity and the state of polarization of a (partially) polarized beam of light are completely determined by the four Stokes parameters I, Q, U, and V. The Stokes parameters form the components of a real four-vector $\mathbf{I}=\{I,Q,U,V\}$, called the Stokes vector,[1] which satisfies the Stokes criterion, i.e., we have

$$I\geqslant(Q^2+U^2+V^2)^{1/2}, \tag{1}$$

so that the degree of polarization

$$p=(Q^2+U^2+V^2)^{1/2}/I\leqslant 1. \tag{2}$$

For an introduction of Stokes parameters we refer to Refs. 2–4.

In physics one encounters a plethora of real 4×4 matrices $\mathbf{M}$ that represent a linear transformation of (the Stokes parameters of) a beam of light into (the Stokes parameters of) another beam of light. Such matrices are called Mueller matrices. In order that a Mueller matrix $\mathbf{M}$ be physically meaningful, it must have the following necessary property: For every column vector $\mathbf{I}_0=\{I_0,Q_0,U_0,V_0\}$ satisfying the inequality

$$I_0\geqslant(Q_0{}^2+U_0{}^2+V_0{}^2)^{1/2}, \tag{3}$$

the product vector $\mathbf{I}=\mathbf{M}\mathbf{I}_0$ satisfies Eq. (1). If this holds $\mathbf{M}$ is said to satisfy the Stokes criterion.

In reality, Mueller matrices may satisfy more restrictions than the Stokes criterion only. They are usually derived as the averages of so-called pure Mueller matrices, which describe the transformation of the four-vector of Stokes parameters of an incident beam of strictly monochromatic light into the four-vector of Stokes parameters of an outgoing beam of light if the corresponding complex electric vectors are transformed into each other by a complex 2×2 matrix called the Jones matrix.[3,5,6] More precisely, they are usually derived in the form

$$\mathbf{M}=\sum_{i=1}^{N} c_i\mathbf{M}_{\mathbf{J}_i}, \tag{4}$$

[a)]Permanent address: Dipartimento di Matematica, Università di Cagliari, Via Ospedale 72, 09124 Cagliari, Italy.

0022-2488/93/34(11)/5072/17/$6.00

where $c_1,...,c_N \geqslant 0$ and $\mathbf{M}_{\mathbf{J}_i}$ is the pure Mueller matrix corresponding to the Jones matrix $\mathbf{J}_i$. The connection between $\mathbf{M}_\mathbf{J}$ and $\mathbf{J}$ is given in, e.g., Sec. 5.14 of Ref. 3. In this article the term "Mueller matrix" will pertain to a matrix of the form (4).

The first major purpose of this article is to derive necessary and sufficient conditions for a real 4×4 matrix $\mathbf{M}$ to satisfy the Stokes criterion in terms of the eigenvalue structure of $\mathbf{M}$. The problem of finding necessary and sufficient conditions for $\mathbf{M}$ to satisfy the Stokes criterion has been studied by Konovalov,[7] Van der Mee and Hovenier,[8] and Nagirner.[9] According to Refs. 7–9, it is sufficient to prove that a real 4×4 matrix $\mathbf{M}$ maps all real vectors $\xi=\{1,q,u,v\}$ satisfying

$$q^2+u^2+v^2=1 \tag{5}$$

into vectors $\mathbf{I}=\{I,Q,U,V\}$ satisfying Eq. (1). If we introduce

$$\mathbf{G}=\mathrm{diag}(1,-1,-1,-1) \tag{6}$$

and the real symmetric matrix

$$\mathbf{N}=\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}, \tag{7}$$

where $\tilde{\mathbf{M}}$ denotes the matrix transpose of $\mathbf{M}$, then $\mathbf{M}$ satisfies the Stokes criterion if and only if

$$M_{11} \geqslant (M_{12}{}^2+M_{13}{}^2+M_{14}{}^2)^{1/2} \tag{8}$$

and

$$D(q,u,v)=N_{11}+2N_{12}q+2N_{13}u+2N_{14}v+N_{22}q^2+N_{33}u^2+N_{44}v^2+2N_{23}qu+2N_{24}qv+2N_{34}uv \tag{9}$$

has a non-negative minimum if $\{q,u,v\}$ satisfies the constraint (5). For the special real block-diagonal matrices

$$\mathbf{F}=\begin{pmatrix} a_1 & b_1 & 0 & 0 \\ b_1 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & b_2 \\ 0 & 0 & -b_2 & a_4 \end{pmatrix}, \tag{10}$$

which occur, e.g., as scattering matrices of certain particles,[3] Mueller matrices of certain optical devices,[5] and reflection matrices of one-dimensionally rough surfaces,[10] the minimum of $D(q,u,v)$ under the constraint (5) can be computed analytically to find necessary and sufficient conditions for $\mathbf{F}$ to satisfy the Stokes criterion.[7,9] In Ref. 8 a diagonalization method was used to find necessary and sufficient conditions for $\mathbf{F}$ to satisfy the Stokes criterion.

A different approach to finding conditions for a real matrix to satisfy the Stokes criterion has been suggested by Nagirner[9] and Xing:[11] a diagonalization of the given real matrix $\mathbf{M}$, followed by a simple comparison test of its eigenvalues. It has been applied to matrices of the form (10) in Refs. 8, 9. Said in simple terms, if one were to factorize $\mathbf{M}$ in the form

$$\mathbf{M}=\mathbf{U}_1^{-1}\,\mathrm{diag}(\lambda_0,\lambda_1,\lambda_2,\lambda_3)\mathbf{U}_2, \tag{11}$$

where the exponent -1 denotes the matrix inverse and $\mathbf{U}_1$ and $\mathbf{U}_2$ are real 4×4 matrices satisfying

$$\tilde{\mathbf{U}}\mathbf{G}\mathbf{U}=\mathbf{G}, \quad [\mathbf{U}]_{11}>0, \tag{12}$$

then **M** satisfies the Stokes criterion if and only if the diagonal matrix on the right-hand side of Eq. (11) does, and the latter is true if and only if $\lambda_0 \geqslant \max(|\lambda_1|, |\lambda_2|, |\lambda_3|)$. Studying only factorizations of **M** where $\mathbf{U}_1=\mathbf{U}_2$, Nagirner[9] either had to restrict himself to the class of matrices **M** for which **GM** is real symmetric, or to state a similar criterion in terms of the eigenvalues of $\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}$. Neither author indicates whether the diagonalization (11) is possible and, if this is the case, how it is to be obtained. We note that the real matrices **U** satisfying Eq. (12) form the orthochronous subgroup of the Lorentz group.[12]

At this point, we devote a few lines to the physical significance of **G**-symmetric matrices. If **M** is the scattering matrix of an assembly consisting of one kind of particles where for each particle in one position a particle in the reciprocal position is found, $\mathrm{diag}(1,1,-1,1)\mathbf{M}$ is real symmetric.[3] Thus $\mathrm{diag}(1,-1,1,-1)\mathbf{M}$ becomes real symmetric on premultiplication by **G**. Since a matrix **M** satisfies the Stokes criterion whenever $\mathrm{diag}(1,-1,-1,1)\mathbf{M}$ does, the class of real matrices **M** such that **GM** is real symmetric is physically relevant.

The second major purpose of this article is to derive necessary and sufficient conditions for a real 4×4 matrix to be a Mueller matrix. Such necessary and sufficient conditions have already been given by Cloude,[13,14] who transformed a given real 4×4 matrix **M** into a Hermitian 4×4 matrix **T**, the so-called coherency matrix, and showed **M** to be a Mueller matrix if and only if all of the eigenvalues of **T** are non-negative. From the eigenvalues λ_i and an orthonormal basis $\{k_0^{(i)},k_1^{(i)},k_2^{(i)},k_3^{(i)}\}$ of corresponding eigenvectors of **T** one finds

$$\mathbf{M}=\sum_{i=1}^{4}\lambda_i \mathbf{M}_{\mathbf{J}_i}, \tag{13}$$

where for $i=1,2,3,4$

$$\mathbf{J}_i=\begin{pmatrix} k_0^{(i)}+k_1^{(i)} & k_2^{(i)}-ik_3^{(i)} \\ k_2^{(i)}+ik_3^{(i)} & k_0^{(i)}-k_1^{(i)} \end{pmatrix}. \tag{14}$$

We present details on the coherency matrix in the Appendix. In this article we try to find necessary and sufficient conditions in terms of the eigenvalue structure of **M**.

Conditions for a real 4×4 matrix to be a (pure) Mueller matrix have been given by many authors,[11–20] but not all of these authors have specified all of their conditions, some have given necessary conditions and others sufficient conditions. In Refs. 13, 15, 17, and 20 necessary and sufficient conditions were given concerning pure Mueller matrices and in Refs. 13 and 14 concerning general Mueller matrices. In Refs. 11 and 12 the importance of the Lorentz group was stressed.

In the present article, we answer both questions asked about the Stokes criterion: A diagonalization of the form (11) is not always possible, but, if it is, a procedure is given to find it. If it is not possible, a different "normal form" is found that also allows one to reduce the verification of the Stokes criterion to an eigenvalue comparison test. To do the job, we apply the theory of matrices real symmetric with respect to an indefinite scalar product.[21,22] In fact, we introduce the indefinite scalar product

$$[\mathbf{I}_1,\mathbf{I}_2]=I_1I_2^*-Q_1Q_2^*-U_1U_2^*-V_1V_2^*, \tag{15}$$

where $\mathbf{I}_j=\{I_j,Q_j,U_j,V_j\}$ with $j=1,2$ and the asterisk denotes complex conjugation. Then, obviously, **M** satisfies the Stokes criterion if $[\mathbf{MI},\mathbf{MI}]\geqslant0$ and $[\mathbf{MI}]_1\geqslant0$ for every real vector **I** with $[\mathbf{I},\mathbf{I}]\geqslant 0$ and $[\mathbf{I}]_1\geqslant 0$. A complex indefinite scalar product is defined, because the theory of indefinite scalar products is formulated for complex matrices, in spite of the fact that the Stokes parameters are always real.

If $\mathbf{M}$ is a real 4×4 matrix such that $\mathbf{GM}$ is symmetric, the above normal form of $\mathbf{M}$ can be used to find necessary and sufficient conditions for $\mathbf{M}$ to be a Mueller matrix. More precisely, we first prove that all real 4×4 matrices satisfying Eq. (12) and $\det \mathbf{U}=+1$ are pure Mueller matrices, which reduces the problem to tackling the normal form of $\mathbf{M}$. Using Cloude's coherency matrix $\mathbf{T}$, we easily find necessary and sufficient conditions for $\mathbf{M}$ to be a Mueller matrix in terms of the eigenvalue structure of $\mathbf{M}$.

Throughout the article, $\mathrm{col}(\xi_0,\xi_1,...,\xi_{m-1})$ denotes the $m\times m$ matrix whose columns are the m-vectors $\xi_0,...,\xi_{m-1}$.

II. REDUCTION TO NORMAL FORM

The four-dimensional vector space with indefinite scalar product (15) is in fact the usual Minkowski space, possibly extended to deal with complex vectors. Following the terminology of Refs. 21, 22, a vector $\mathbf{I}$ is called *positive* if $[\mathbf{I},\mathbf{I}]>0$, *negative* if $[\mathbf{I},\mathbf{I}]<0$, *non-negative* if $[\mathbf{I},\mathbf{I}]\geqslant 0$, *nonpositive* if $[\mathbf{I},\mathbf{I}]\leqslant 0$, and *neutral* if $[\mathbf{I},\mathbf{I}]=0$. Linear subspaces consisting of only non-negative vectors have dimension of at most one and linear subspaces consisting of only nonpositive vectors have dimension of at most three. Also, if $[\mathbf{I},\mathbf{J}]=0$ for every vector $\mathbf{J}$, then $\mathbf{I}=0$.

A real 4×4 matrix $\mathbf{M}$ is called $\mathbf{G}$ *symmetric* if $\mathbf{GM}$ is real symmetric, i.e., if

$$\tilde{\mathbf{M}}=\mathbf{GMG}. \tag{16}$$

As known, a real symmetric matrix has only real eigenvalues, is diagonalizable, and can in fact be diagonalized by a real orthogonal matrix of determinant 1. These properties are lost for $\mathbf{G}$-symmetric matrices: They may have complex conjugate pairs of eigenvalues and may not be diagonalizable. Instead, it is possible to find a real 4×4 matrix $\mathbf{U}$ such that $\mathbf{GU}$ is orthogonal and $\mathbf{U}^{-1}\mathbf{MU}$ has a special form.[22]

The $\mathbf{G}$-orthogonal matrices, i.e., the matrices $\mathbf{U}$ such that $\tilde{\mathbf{U}}\mathbf{GU}=\mathbf{G}$, form the so-called Lorentz group.[23] Any such matrix $\mathbf{U}_0$ can be written in one of the forms $\mathbf{U}_0=\mathbf{U}$, $\mathbf{U}_0=-\mathbf{U}$, $\mathbf{U}_0=\mathbf{GU}$ or $\mathbf{U}_0=-\mathbf{GU}$, where

$$\tilde{\mathbf{U}}\mathbf{GU}=\mathbf{G}, \quad \det \mathbf{U}=+1, \quad [\mathbf{U}]_{11}>0. \tag{17}$$

Denoting by $\mathscr{L}$ the set of all real 4×4 matrices U satisfying (17), we seek a normal form of $\mathbf{G}$-symmetric matrices where the diagonalizing transformation $\mathbf{U}\in\mathscr{L}$. This requires the use of the Jordan normal form of a matrix.

Any $n\times n$ matrix $\mathbf{B}$ can be brought to the Jordan normal form, i.e., there exists an invertible matrix $\mathbf{S}$ such that

$$\mathbf{S}^{-1}\mathbf{BS}=\mathrm{diag}(\mathbf{J}_{m_1}(\lambda_1),...,\mathbf{J}_{m_r}(\lambda_r)), \tag{18}$$

where $\lambda_1,...,\lambda_r$ are the eigenvalues of $\mathbf{B}$, the $m\times m$ matrix $\mathbf{J}_m(\lambda)$ is defined by $[\mathbf{J}_m(\lambda)]_{ij}=\lambda\delta_{i,j}+\delta_{i+1,j}$ if $m\geqslant 2$, and by (λ) if $m=1$. The columns of $\mathbf{S}$ are composed of the so-called Jordan chains of eigenvectors and generalized eigenvectors of $\mathbf{B}$. That is, corresponding to the Jordan block $\mathbf{J}_m(\lambda)$ we have the "chain" of vectors $\xi_0,...,\xi_{m-1}$ with $\xi_0\neq 0$ satisfying

$$
\begin{aligned}
\mathbf{B}\xi_{m-1}&=\lambda\xi_{m-1}+\xi_{m-2},\\
\mathbf{B}\xi_{m-2}&=\lambda\xi_{m-2}+\xi_{m-3},\\
&\vdots\\
\mathbf{B}\xi_{2}&=\lambda\xi_{2}+\xi_{1},\\
\mathbf{B}\xi_{1}&=\lambda\xi_{1}+\xi_{0},\\
\mathbf{B}\xi_{0}&=\lambda\xi_{0}.
\end{aligned}
\tag{19}
$$

For $m=1$ we only have the last equation to deal with.

THEOREM 2.1: *Let* **B** *be a* **G**-*symmetric matrix. Then we have the following options*:

(*1*) **B** *has the four real eigenvalues* λ_0, λ_1, λ_2, *and* λ_3 *with the corresponding positive eigenvector* ξ_0 *and negative eigenvectors* ξ_1, ξ_2, *and* ξ_3. *We may normalize the eigenvectors in such a way that the matrix* $\mathbf{U}=col(\xi_0,\xi_1,\xi_2,\xi_3)\in\mathscr{L}$.

(*2*) **B** *has the real eigenvalues* λ, μ, *and* ν *but is not diagonalizable. The eigenvectors* η *corresponding to* μ *and* ζ *corresponding to* ν *are negative, whereas to the double eigenvalue* λ *there corresponds one Jordan block of size 2 with corresponding Jordan chain* (ξ_0,ξ_1). *We may choose the eigenvectors and generalized eigenvectors and the real constants c and d in such a way that the matrix* $\mathbf{U}=col(\xi_1+c\xi_0,\xi_1+d\xi_0,\eta,\zeta)\in\mathscr{L}$.

(*3*) **B** *has the real eigenvalues* λ *and* μ *but is not diagonalizable. The eigenvector* η *corresponding to* μ *is negative, whereas to the triple eigenvalue* λ *there corresponds one Jordan block of size 3 with corresponding Jordan chain* (ξ_0,ξ_1,ξ_2). *We may choose the eigenvectors and generalized eigenvectors and the constants a, b, c, and d in such a way that the matrix* $\mathbf{U}=col(\xi_2+a\xi_1+c\xi_0,(\xi_1+b\xi_0)/\sqrt{-\alpha},\xi_2+a\xi_1+d\xi_0,\eta)\in\mathscr{L}$ *for some* $\alpha<0$. *It is possible to take* $a=b$.

(*4*) **B** *has the two simple complex conjugate eigenvalues* $x\pm iy$ *and the two real eigenvalues* μ *and* ν. *The eigenvectors* $\xi_\pm$ *pertaining to* $x\pm iy$ *are neutral, and the eigenvectors* η *pertaining to* μ *and* ζ *pertaining to* ν *are negative. One may choose the eigenvectors and a complex constant c in such a way that the matrix* $\mathbf{U}=col(\frac{1}{2}(c\xi_++c^*\xi_-)(i/2)(c\xi_+-c^*\xi_-),\eta,\zeta)\in\mathscr{L}$.

This theorem is easily derived by people familiar with the results of Refs. 21, 22. However, we seek a proof in terms of elementary matrix algebra aimed at a nonspecialist in indefinite inner product spaces and we are also interested in the explicit form of **U**. Hence some elaboration follows.

Proof: In the first case we simply normalize the eigenvectors such that

$$[\xi_r,\xi_s]=(2\delta_{r,0}-1)\delta_{r,s},\quad r,s=0,1,2,3. \tag{20}$$

Replacing some ξ_r by $-\xi_r$, we may choose them in such a way that $[\xi_0]_1>0$ and the matrix $\mathbf{U}=\mathrm{col}(\xi_0,\xi_1,\xi_2,\xi_3)\in\mathscr{L}$.

In the second case we easily verify the following:

$$
\begin{gathered}
[\eta,\eta]<0,\quad [\zeta,\zeta]<0,\quad [\eta,\zeta]=0,\\
[\xi_0,\xi_0]=0,\quad [\xi_0,\xi_1]\neq 0,\\
[\xi_0,\eta]=[\xi_1,\eta]=[\xi_0,\zeta]=[\xi_1,\zeta]=0,\\
\mathbf{B}\xi_1=\lambda\xi_1+\xi_0,\quad \mathbf{B}\xi_0=\lambda\xi_0.
\end{gathered}
\tag{21}
$$

We then normalize η and ζ to get $[\eta,\eta]=[\zeta,\zeta]=-1$, where one may still change the signs of η and ζ. We then define $c=\{1-[\xi_1,\xi_1]\}/2[\xi_0,\xi_1]$ and $d=-\{1+[\xi_1,\xi_1]\}/2[\xi_0,\xi_1]$ and check that $[\xi_1+c\xi_0,\xi_1+c\xi_0]=1$, $[\xi_1+d\xi_0,\xi_1+d\xi_0]=-1$, and $[\xi_1+c\xi_0,\xi_1+d\xi_0]=0$. Replacing η by $-\eta$, ζ by $-\zeta$, and (ξ_0,ξ_1) by $(-\xi_0,-\xi_1)$ if necessary, the matrix $\mathbf{U}=\mathrm{col}(\xi_1+c\xi_0,\xi_1+d\xi_0,\eta,\zeta)\in\mathscr{L}$.

In the third case we easily verify the following:

$$[\eta,\eta]<0,\quad [\eta,\xi_0]=[\eta,\xi_1]=[\eta,\xi_2]=0,$$

$$[\xi_0,\xi_0]=[\xi_0,\xi_1]=0,\quad [\xi_1,\xi_1]=[\xi_0,\xi_2]<0, \tag{22}$$

$$\mathbf{B}\xi_2=\lambda\xi_2+\xi_1,\quad \mathbf{B}\xi_1=\lambda\xi_1+\xi_0,\quad \mathbf{B}\xi_0=\lambda\xi_0.$$

We may normalize η to get $[\eta,\eta]=-1$, where one may still change the sign of η. Now let us seek real constants a, b, c, and d such that

$$[\xi_1+b\xi_0,\xi_2+a\xi_1+c\xi_0]=[\xi_1+b\xi_0,\xi_2+a\xi_1+d\xi_0]=0,$$

$$[\xi_2+a\xi_1+c\xi_0,\xi_2+a\xi_1+d\xi_0]=0,$$

$$[\xi_2+a\xi_1+c\xi_0,\xi_2+a\xi_1+c\xi_0]=1, \tag{23}$$

$$[\xi_2+a\xi_1+d\xi_0,\xi_2+a\xi_1+d\xi_0]=-1,$$

$$[\xi_1+b\xi_0,\xi_1+b\xi_0]=\alpha.$$

Using $[\xi_0,\xi_2]=[\xi_1,\xi_1]=\alpha$, $[\xi_1,\xi_2]=\beta$, and $[\xi_2,\xi_2]=\gamma$ with $\alpha<0$ and β, γ real, we get the equations

$$\beta+(a+b)\alpha=0,\quad \gamma+2a\beta+(c+d+a^2)\alpha=0,$$

$$\gamma+2a\beta+(2c+a^2)\alpha=1, \tag{24}$$

$$\gamma+2a\beta+(2c+a^2)\alpha=1,\quad \gamma+2a\beta+(2d+a^2)\alpha=-1.$$

We can now find a, b, c, and d, while $c-d=1/\alpha$. Changing the signs of some of the eigenvectors and generalized eigenvectors if necessary, we find the matrix $\mathbf{U}=\mathrm{col}(\xi_2+a\xi_1+c\xi_0,(\xi_1+b\xi_0)/\sqrt{-\alpha},\xi_2+a\xi_1+d\xi_0,\eta)\in\mathscr{L}$. It is easy to see that a, b, c, and d can be found in such a way that $a=b$.

In the fourth case the eigenvectors and generalized eigenvectors are easily found to satisfy

$$[\xi_+,\xi_+]=[\xi_-,\xi_-]=0,\quad [\xi_+,\xi_-]=[\xi_-,\xi_+]^*\neq 0,$$

$$[\eta,\eta]=[\zeta,\zeta]=-1,\quad [\eta,\eta]<0,\quad [\zeta,\zeta]<0,\quad [\eta,\zeta]=0, \tag{25}$$

$$[\xi_\pm,\eta]=[\xi_\pm,\zeta]=0,$$

where $B\xi_\pm=(x\pm iy)\xi_\pm$. By normalization we get $[\eta,\eta]=[\zeta,\zeta]=-1$, where one may still change the signs of η and ζ. Now let c be a complex constant such that $c^2[\xi_+,\xi_-]=2$. Then, changing the signs of (ξ_+,ξ_-), η, and ζ if necessary, the matrix $\mathbf{U}=\mathrm{col}(\tfrac{1}{2}(c\xi_++c^*\xi_-),(i/2)(c\xi_+-c^*\xi_-),\eta,\zeta)\in\mathscr{L}$. ■

If **B** is **G** symmetric and there is a Jordan block of size 2 corresponding to the real eigenvalue λ, then two situations occur. Indeed, if (ξ_0,ξ_1) is a Jordan chain corresponding to this eigenvalue, then this Jordan chain may be replaced by $(c\xi_0,c\xi_1+cd\xi_0)$ for any pair of complex constants c and d where $c\neq 0$. Since

$$[c\xi_0,c\xi_1+cd\xi_0]=|c|^2[\xi_0,\xi_1] \tag{26}$$

the sign of the nonzero real number $[\xi_0,\xi_1]$ is independent of the choice of the Jordan chain (ξ_0,ξ_1). We then call the Jordan block a block of positive or negative sign[22] if $[\xi_0,\xi_1]$ is positive or negative, respectively. Replacing **B** with $\mathbf{U}^{-1}\mathbf{B}\mathbf{U}$ for some $\mathbf{U}\in\mathscr{L}$ does not change the sign of any Jordan block of size 2. Thus the sign of a Jordan block of size 2 is an invariant of the **G**-symmetric matrix **B**.

Let us compute $\mathbf{U}^{-1}\mathbf{B}\mathbf{U}$ in the four cases described by Theorem 2.1. Then we get the following:

(1) If **B** is diagonalizable with four real eigenvalues, we get

$$\mathbf{U}^{-1}\mathbf{B}\mathbf{U}=\operatorname{diag}(\lambda_0,\lambda_1,\lambda_2,\lambda_3). \tag{27}$$

(2) If **B** has a Jordan block of size 2 at the real eigenvalue λ and two more real eigenvalues μ and ν, then

$$\mathbf{U}^{-1}\mathbf{B}\mathbf{U}=\begin{pmatrix}\lambda+\epsilon & -\epsilon & 0 & 0\\ \epsilon & \lambda-\epsilon & 0 & 0\\ 0 & 0 & \mu & 0\\ 0 & 0 & 0 & \nu\end{pmatrix}, \tag{28}$$

where $\epsilon=[\xi_0,\xi_1]$. Note that in Eq. (28) the sign of the Jordan block of size 2 is exactly the sign of ϵ, since $\epsilon=[\xi_0,\xi_1]$.

(3) If **B** has a Jordan block of size 3 at the real eigenvalue λ and one more real eigenvalue μ, then, putting $\delta=(a-b)\alpha$ and $\epsilon = \sqrt{-\alpha}$

$$\mathbf{U}^{-1}\mathbf{B}\mathbf{U}=\begin{pmatrix}\lambda+\delta & \epsilon & -\delta & 0\\ -\epsilon & \lambda & \epsilon & 0\\ \delta & \epsilon & \lambda-\delta & 0\\ 0 & 0 & 0 & \mu\end{pmatrix}. \tag{29}$$

If $a=b$, we get $\delta=0$ in Eq. (29).

(4) If **B** has the two complex conjugate eigenvalues $x\pm iy$ and the two real eigenvalues μ and ν, then

$$\mathbf{U}^{-1}\mathbf{B}\mathbf{U}=\begin{pmatrix}x & y & 0 & 0\\ -y & x & 0 & 0\\ 0 & 0 & \mu & 0\\ 0 & 0 & 0 & \nu\end{pmatrix}. \tag{30}$$

III. THE STOKES CRITERION

In this section, Eqs. (27)–(30) are employed to find necessary and sufficient conditions for a real 4×4 matrix **M** to satisfy the Stokes criterion. If **M** is **G** symmetric, these conditions will be formulated directly in terms of the eigenvalues of **M**. On the other hand, if **M** is a general real 4×4 matrix, such conditions will be stated in terms of the matrix $\mathbf{A}=\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}$.

THEOREM 3.1: *Let* **M** *be a real* **G***-symmetric matrix. Then* **M** *satisfies the Stokes criterion if and only if one of the following two situations occurs*:

(1) **M** *has the one real eigenvalue λ_0 corresponding to a positive eigenvector and three real eigenvalues λ_1, λ_2, and λ_3 corresponding to negative eigenvectors, and $\lambda_0\geqslant\max(|\lambda_1|,|\lambda_2|,|\lambda_3|)$.*

(2) **M** *has the real eigenvalues λ, μ, and ν but is not diagonalizable. The eigenvectors corresponding to μ and ν are negative, whereas to the double eigenvalue λ there corresponds one Jordan block of size 2 with positive sign. Moreover, $\lambda\geqslant\max(|\mu|,|\nu|)$.*

Proof: Let us discuss the four cases of Theorem 2.1 and the corresponding four identities (27)–(30) [with **M** instead of **B**] in turn, and observe that **M** satisfies the Stokes criterion if and only if the left-hand side of the corresponding one of Eqs. (27)–(30) does.

In the first case, Eq. (27) immediately leads to the necessary and sufficient condition $\lambda_0\geqslant\max(|\lambda_1|,|\lambda_2|,|\lambda_3|)$ for **M** to satisfy the Stokes criterion.

In the second case, $\mathbf{U}^{-1}\mathbf{M}\mathbf{U}$ satisfying the Stokes criterion implies that $\lambda+\epsilon\geqslant\lambda-\epsilon$ and $\lambda+\epsilon\geqslant\epsilon$, so that one must have $\lambda\geqslant0$ and $\epsilon>0$ for the Stokes criterion to be satisfied. If $\lambda\geqslant0$ and $\epsilon>0$, we get for all real vectors $\xi=(1,q,u,v)$ satisfying Eq. (5)

$$[\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi,\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi]=(\lambda^2-\nu^2)(1-q^2)+2\epsilon\lambda(1-q)^2+(\nu^2-\mu^2)u^2$$
$$=(\lambda^2-\rho^2)(1-q^2)+2\epsilon\lambda(1-q)^2, \tag{31}$$

where $\rho=\max(|\mu|,|\nu|)$. This expression is non-negative if and only if $\lambda\geqslant\rho$ and $\epsilon>0$, as claimed.

In the third case, starting from Eq. (29) with $\delta=0$, we find for all real vectors $\xi=(1,q,u,v)$ satisfying Eq. (5)

$$[\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi,\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi]=(\lambda^2-\mu^2)(1-q^2-u^2)-\epsilon^2(1+u^2)+4\lambda\epsilon q(1-u), \tag{32}$$

which may become negative if we take $q=0$ and $u=1$. Hence, if **M** has a Jordan block of order 3 at some real eigenvalue, it cannot satisfy the Stokes criterion.

In the fourth case, starting from Eq. (30), we find for all real vectors $\xi=(1,q,u,v)$ satisfying Eq. (5)

$$[\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi,\mathbf{U}^{-1}\mathbf{M}\mathbf{U}\xi]=(x^2-y^2-\nu^2)+2qxy+(\nu^2-x^2+y^2)q^2+(\nu^2-\mu^2)u^2$$
$$\geqslant(x^2-y^2-\rho^2)+2qxy+(\rho^2-x^2+y^2)q^2, \tag{33}$$

which may become negative. Thus if **M** has complex eigenvalues, it cannot satisfy the Stokes criterion. ■

Let us derive necessary and sufficient conditions for an arbitrary real 4×4 matrix **M** to satisfy the Stokes criterion. Starting from a given real 4×4 matrix **M**, let us define **A** by

$$\mathbf{A}=\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}. \tag{34}$$

Then it is easily seen that **M** satisfies the Stokes criterion if and only

$$[\mathbf{A}\mathbf{I},\mathbf{I}]\geqslant0 \quad \text{and} \quad [\mathbf{M}\mathbf{I}]_1\geqslant0, \tag{35}$$

whenever **I** is a real vector satisfying $[\mathbf{I},\mathbf{I}]\geqslant0$ and $[\mathbf{I}]_1\geqslant0$. The second part of Eq. (35) is equivalent to the inequality [cf. Eq. (8)]

$$M_{11} \geqslant (M_{12}^2 + M_{13}^2 + M_{14}^2)^{1/2}. \tag{36}$$

THEOREM 3.2: *Let* **M** *be a real* 4×4 *matrix satisfying Eq. (36). Then* **M** *satisfies the Stokes criterion if and only if one of the following two situations occurs*:

(1) **A** *has the one real eigenvalue* λ_0 *corresponding to a positive eigenvector and three real eigenvalues* λ_1, λ_2, *and* λ_3 *corresponding to negative eigenvectors, and* $\lambda_0 \geqslant \max(0,\lambda_1,\lambda_2,\lambda_3)$.

(2) **A** *has the real eigenvalues* λ, μ, *and* ν *but is not diagonalizable. The eigenvectors corresponding to* μ *and* ν *are negative, whereas to the double eigenvalue* λ *there corresponds one Jordan block of size 2 with positive sign. Moreover,* $\lambda \geqslant \max(0,\mu,\nu)$.

Proof: Let us discuss the four cases of Theorem 2.1 and the corresponding four identities (27)–(30) [with **A** instead of **B**] in turn, and observe that one only has to verify that $[\mathbf{U}^{-1}\mathbf{AUI},\mathbf{I}] \geqslant 0$ for all real vectors **I** satisfying $[\mathbf{I},\mathbf{I}] \geqslant 0$ and $[\mathbf{I}]_1 \geqslant 0$, where $\mathbf{U}^{-1}\mathbf{AU}$ is the left-hand side of the corresponding one of Eqs. (27)–(30).

In the first case, Eq. (27) immediately leads to the necessary and sufficient condition $\lambda_0 \geqslant \max(0,\lambda_1,\lambda_2,\lambda_3)$, since $[\mathbf{U}^{-1}\mathbf{AUI},\mathbf{I}] = \lambda_0 I^2 - \lambda_1 Q^2 - \lambda_2 U^2 - \lambda_3 V^2$ for all real vectors $\mathbf{I} = \{I,Q,U,V\}$.

In the second case, as in the proof of Theorem 3.1, **M** [and hence **A**] satisfying the Stokes criterion implies that $\lambda \geqslant 0$ and $\epsilon > 0$ in Eq. (28). Thus if this is indeed the case, for real $\xi = \{1,q,u,v\}$ satisfying Eq. (5), we get for $\rho = \max(\nu,\mu)$

$$\begin{aligned}[\mathbf{U}^{-1}\mathbf{AU}\xi,\xi] &= (\lambda+\epsilon-\nu) - 2\epsilon q - (\lambda-\epsilon-\nu)q^2 + (\nu-\mu)u^2 \\ &\geqslant (\lambda+\epsilon-\rho) - 2\epsilon q - (\lambda-\epsilon-\rho)q^2 \\ &= (1-q)[(\lambda-\rho)(1+q) + \epsilon(1-q)],\end{aligned} \tag{37}$$

which is non-negative if $\lambda \geqslant \rho$ and $\epsilon > 0$, as claimed.

In the third case, starting from Eq. (29) [with **B** replaced by **A** and with $\delta = 0$] we obtain for real $\xi = \{1,q,u,v\}$ satisfying Eq. (5) and $\rho = \max(\nu,\mu)$

$$[\mathbf{U}^{-1}\mathbf{AU}\xi,\xi] = (\lambda-\mu)(1-q^2-u^2) + 2\epsilon q(1-u), \tag{38}$$

which may become negative. Hence, if **A** has a Jordan block of order 3 at some real eigenvalue, **M** cannot satisfy the Stokes criterion.

In the fourth case, starting from Eq. (30) [with **B** replaced by **A**] we obtain for real $\xi = \{1,q,u,v\}$ satisfying Eq. (5) and $\rho = \max(\nu,\mu)$

$$[\mathbf{U}^{-1}\mathbf{AU}\xi,\xi] = (x-\nu) + 2qy + (\nu-x)q^2 + (\nu-\mu)u^2 \geqslant (x-\rho) + 2qy + (\rho-x)q^2, \tag{39}$$

which becomes $\pm 2y$ for $q = \pm 1$ and hence may become negative. Thus if **A** has complex eigenvalues, **M** cannot satisfy the Stokes criterion. ■

If **M** is **G** symmetric, we have $\mathbf{A} = \mathbf{M}^2$, as a result of Eq. (16). Then the eigenvalues of **A** are the squares of the eigenvalues of **M**.

Apart from studying the Stokes criterion, it is useful to find, for a given constant $\delta \in (0,1)$, necessary and sufficient conditions for a real 4×4 matrix to satisfy the Stokes criterion in the following strong sense: If $\mathbf{I}_0$ is a real vector satisfying Eq. (3), then $\mathbf{I} = \{I,Q,U,V\} = \mathbf{MI}_0$ satisfies

$$\delta I \geqslant (Q^2 + U^2 + V^2)^{1/2}. \tag{40}$$

Any such real 4×4 matrix transforms the four-vector of Stokes parameters of an incident beam of light into a four-vector of Stokes parameters of an outgoing beam with degree of polarization not exceeding δ. Indeed, if $\mathbf{M}$ is $\mathbf{G}$ symmetric, it cannot have a Jordan block of size 2 with positive sign at the eigenvalue λ and satisfy the Stokes criterion in the above strong sense [cf. Eq. (31)]. Similarly, if $\mathbf{M}$ is a real 4×4 matrix satisfying Eq. (36), $\mathbf{A}$ defined by Eq. (34) cannot have a Jordan block of size 2 with positive sign at the eigenvalue λ while $\mathbf{M}$ satisfies the Stokes criterion in the above strong sense [cf. Eq. (37) applied to $\mathbf{M}\cdot\mathbf{G}\widetilde{\mathbf{M}}\mathbf{G}$ which has the same Jordan structure at $\lambda>0$ as $\mathbf{G}\widetilde{\mathbf{M}}\mathbf{G}\cdot\mathbf{M}$, including the sign of the Jordan block]. Hence, if $\mathbf{M}$ is to satisfy the Stokes criterion in the above strong sense for some $\delta\in(0,1)$, then $\mathbf{M}$ satisfies Condition 1 of Theorem 3.1 with $\lambda_0>\max(|\lambda_1|,|\lambda_2|,|\lambda_3|)$ if $\mathbf{M}$ is $\mathbf{G}$-symmetric, or Condition 1 of Theorem 3.2 with $\lambda_0>\max(0,\lambda_1,\lambda_2,\lambda_3)$.

Factorizations of a given real 4×4 matrix $\mathbf{M}$ of the form (11) with $\mathbf{U}_1$, $\mathbf{U}_2\in\mathcal{L}$ are not always possible, even if we allow $\mathbf{U}_1$ and $\mathbf{U}_2$ to be different and $\mathbf{M}$ is a Mueller matrix. A counterexample is provided by the matrix

$$\mathbf{M}=\begin{pmatrix}1 & 1 & 0 & 0\\ \epsilon & \epsilon & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}, \tag{41}$$

where $-1<\epsilon<1$. Indeed, writing $\mathbf{U}_1=\{u_{ij}\}_{i,j=1}^4$, $\mathbf{U}_2^{-1}=\{v_{ij}\}_{i,j=1}^4$, $\bar{u}_k=u_{k1}+u_{k2}$, and $\bar{v}_k=v_{1k}+\epsilon v_{2k}$ $(k=1,2,3,4)$, we obtain

$$\mathbf{U}_1\mathbf{M}\mathbf{U}_2^{-1}=\{\bar{u}_i\bar{v}_j\}_{i,j=1}^4. \tag{42}$$

Requiring $\mathbf{U}_1\mathbf{M}\mathbf{U}_2^{-1}$ to be $\mathbf{G}$ symmetric and using $\mathbf{U}_1$, $\mathbf{U}_2^{-1}\in\mathcal{L}$ imply for certain real x, y, z

$$(\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4)=\bar{u}_1(1,x,y,z),\quad (\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{v}_4)=\bar{v}_1(1,-x,-y,-z),$$

$$\bar{u}_1{}^2-\bar{u}_2{}^2-\bar{u}_3{}^2-\bar{u}_4{}^2=1-\epsilon^2,\quad \bar{v}_1{}^2-\bar{v}_2{}^2-\bar{v}_3{}^2-\bar{v}_4{}^2=0,\quad \bar{u}_1\bar{v}_1>0, \tag{43}$$

which is a contradiction. Hence, there do not exist matrices $\mathbf{U}_1$, $\mathbf{U}_2\in\mathcal{L}$ such that $\mathbf{U}_1\mathbf{M}\mathbf{U}_2^{-1}$ is $\mathbf{G}$ symmetric for $\mathbf{M}$ as in Eq. (41). Hence, a factorization of the form (11) as proposed by Xing[11] for Mueller matrices is not always possible.

IV. A CRITERION FOR MUELLER MATRICES

Barakat[12] pointed out the relationship between pure Mueller matrices and the orthochronous Lorentz group, but did not make their connection precise. This will be the contents of Theorem 4.1. We will use Theorem 4.1 primarily as a tool to prove Theorem 4.2.

THEOREM 4.1: *A real invertible 4×4 matrix* $\mathbf{M}$ *is a pure Mueller matrix if and only if it has the form* $\mathbf{M}=c\mathbf{U}$ *for some* $c>0$ *and* $\mathbf{U}\in\mathcal{L}$.

Proof: The following six Jones and pure Mueller matrices correspond to each other [cf. Ref. 3]

$$\mathbf{J}_1=\begin{pmatrix}\cos(\alpha/2) & -\sin(\alpha/2)\\ \sin(\alpha/2) & \cos(\alpha/2)\end{pmatrix},\quad \mathbf{M}_1=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha & 0\\ 0 & \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 0 & 1\end{pmatrix}; \tag{44}$$

$$\mathbf{J}_2=\begin{pmatrix}\cos(\beta/2) & i\sin(\beta/2)\\ i\sin(\beta/2) & \cos(\beta/2)\end{pmatrix},\quad \mathbf{M}_2=\begin{pmatrix}1&0&0&0\\0&\cos\beta&0&-\sin\beta\\0&0&1&0\\0&\sin\beta&0&\cos\beta\end{pmatrix};\tag{45}$$

$$\mathbf{J}_3=\begin{pmatrix}e^{-i\gamma/2} & 0\\ 0 & e^{i\gamma/2}\end{pmatrix},\quad \mathbf{M}_3=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&\cos\gamma&-\sin\gamma\\0&0&\sin\gamma&\cos\gamma\end{pmatrix};\tag{46}$$

$$\mathbf{J}_4=\begin{pmatrix}e^{\chi/2} & 0\\ 0 & e^{-\chi/2}\end{pmatrix},\quad \mathbf{M}_4=\begin{pmatrix}\cosh\chi&\sinh\chi&0&0\\\sinh\chi&\cosh\chi&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix};\tag{47}$$

$$\mathbf{J}_5=\begin{pmatrix}\cosh(\omega/2) & \sinh(\omega/2)\\ \sinh(\omega/2) & \cosh(\omega/2)\end{pmatrix},\quad \mathbf{M}_5=\begin{pmatrix}\cosh\omega&0&\sinh\omega&0\\0&1&0&0\\\sinh\omega&0&\cosh\omega&0\\0&0&0&1\end{pmatrix};\tag{48}$$

$$\mathbf{J}_6=\begin{pmatrix}\cosh(\zeta/2) & -i\sinh(\zeta/2)\\ i\sinh(\zeta/2) & \cosh(\zeta/2)\end{pmatrix},\quad \mathbf{M}_6=\begin{pmatrix}\cosh\zeta&0&0&\sinh\zeta\\0&1&0&0\\0&0&1&0\\\sinh\zeta&0&0&\cosh\zeta\end{pmatrix}.\tag{49}$$

Any element of the group $\mathscr{L}$ can be written as the product of the six matrices $\mathbf{M}_1,\ldots,\mathbf{M}_6$ (in whatever order) for suitable α, β, γ, χ, ω, ζ.[23] For any of these six types of Lorentz transformation, one may find a Jones matrix from which it is derived. Hence, any $\mathbf{U}\in\mathscr{L}$ is a pure Mueller matrix. Indeed, writing

$$\mathbf{U}=\mathbf{M}_1(\alpha)\mathbf{M}_2(\beta)\mathbf{M}_3(\gamma)\mathbf{M}_4(\chi)\mathbf{M}_5(\omega)\mathbf{M}_6(\zeta)\tag{50}$$

we get $\mathbf{U}=\mathbf{M}_\mathbf{J}$ where for some arbitrary phase φ

$$\mathbf{J}=e^{i\varphi}\mathbf{J}_1(\alpha)\mathbf{J}_2(\beta)\mathbf{J}_3(\gamma)\mathbf{J}_4(\chi)\mathbf{J}_5(\omega)\mathbf{J}_6(\zeta).\tag{51}$$

Since a Jones matrix can always be written as a constant multiple of a matrix of determinant 1, a real invertible 4×4 matrix is a pure Mueller matrix if and only if it is a positive constant multiple of an element of the group $\mathscr{L}$. ■

If $\mathbf{M}$ is a pure Mueller matrix and singular, it corresponds to a singular Jones matrix $\mathbf{J}$. Since $\mathbf{J}$ is a limit of invertible Jones matrices, $\mathbf{M}$ is a limit of pure invertible Mueller matrices and hence a limit, as $n\to\infty$, of matrices $\mathbf{M}_n=c_n\mathbf{U}_n$ where $c_n>0$ and $\mathbf{U}_n\in\mathscr{L}$ for every $n\in\mathbf{N}$. We will not dwell on the structure of singular Jones and singular pure Mueller matrices.[24]

THEOREM 4.2: *Let* $\mathbf{M}$ *be a real* **G**-*symmetric matrix. Then* $\mathbf{M}$ *is a Mueller matrix if and only if one of the following two situations occurs*:

(1) $\mathbf{M}$ *has the one real eigenvalue* λ_0 *corresponding to a positive eigenvector and three real eigenvalues* λ_1, λ_2, *and* λ_3 *corresponding to negative eigenvectors, and* $\lambda_0\pm\lambda_1\geq|\lambda_2\pm\lambda_3|$.

(2) **M** *has the real eigenvalues* λ, μ, *and* ν *but is not diagonalizable. The eigenvectors corresponding to* μ *and* ν *are negative, whereas to the double eigenvalue* λ *there corresponds one Jordan block of size 2 with positive sign. Moreover,* $\lambda \geqslant \frac{1}{2}|\mu+\nu|$ *and* $\mu=\nu$.

Proof: Since a Mueller matrix satisfies the Stokes criterion and every $\mathbf{U}\in\mathscr{L}$ is a pure Mueller matrix [cf. Theorem 4.1], we only have to consider the two cases of Theorem 3.1. In either case, we compute the coherency matrix **T** from the right-hand sides of Eqs. (27) and (28), respectively. From Eq. (27) we get [cf. Eqs. (A12)–(A15)]

$$\mathbf{T}=\operatorname{diag}(t_0,t_1,t_2,t_3), \tag{52}$$

where $t_0=(\lambda_0+\lambda_1+\lambda_2+\lambda_3)/4$, $t_1=(\lambda_0+\lambda_1-\lambda_2-\lambda_3)/4$, $t_2=(\lambda_0-\lambda_1+\lambda_2-\lambda_3)/4$, and $t_3=(\lambda_0-\lambda_1-\lambda_2+\lambda_3)/4$. Since its eigenvalues must be non-negative, we obtain the first part of the theorem. On the other hand, from Eq. (28) we get

$$\mathbf{T}=\frac{1}{4}\begin{pmatrix} 2\lambda+\mu+\nu & 0 & 0 & 0 \\ 0 & 2\lambda-\mu-\nu & 0 & 0 \\ 0 & 0 & 2\epsilon+\mu-\nu & -2i\epsilon \\ 0 & 0 & 2i\epsilon & 2\epsilon-\mu+\nu \end{pmatrix}, \tag{53}$$

whose eigenvalues are $(2\lambda+\mu+\nu)/4$, $(2\lambda-\mu-\nu)/4$, and $\{2\epsilon\pm[(\mu-\nu)^2+4\epsilon^2]^{1/2}\}/4$. Requiring these eigenvalues to be non-negative, one gets the second part of the theorem. ■

If one computes the eigenvalues of the coherency matrix from the right-hand side of Eq. (30), one gets $(2x+\mu+\nu)/4$, $(2x-\mu-\nu)/4$, and $\pm[(\mu-\nu)^2+y^2]^{1/2}/4$. However, starting from the right-hand side of Eq. (29) with $\delta=0$, one gets $(3\lambda+\mu)/4$ and the three real roots of $z^3-pz^2-(p^2+2q^2)z+p^3=0$ with $p=(\lambda-\mu)/2$ and $q=(\epsilon/2)$. Since the three real roots add up to p and have $-p^3$ as their product, not all three roots can be non-negative, unless $p=0$, which leads to $z=0,\pm\epsilon/\sqrt{2}$. Thus in neither case the coherency matrix is non-negative and **M** is a Mueller matrix.

V. CONCLUSIONS

In this article we have given necessary and sufficient conditions for a real 4×4 matrix **M** to satisfy the Stokes criterion, in terms of the eigenvalues of the matrix $\mathbf{A}=\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}$ (Theorem 3.2). For the physically relevant class of real 4×4 matrices **M** for which **GM** is symmetric, criteria have been provided, in terms of the eigenvalue structure of **M**, to satisfy the Stokes criterion (Theorem 3.1) and to be a Mueller matrix (Theorem 4.2). As a result, a real matrix **M** for which **A** has complex eigenvalues, or a real matrix **M** for which **GM** is symmetric and **M** has complex eigenvalues does not satisfy the Stokes criterion. Moreover, the result of each of these theorems has been derived in one of two cases: the first case in which there are four real eigenvalues λ_0, λ_1, λ_2, λ_3 one of which (λ_0) is dominant, and the second case in which there are three eigenvalues λ, μ, ν where the one of multiplicity two (λ) is dominant.

It is not obvious how, in analogy to Theorem 3.2, to give necessary and sufficient conditions for a real 4×4 matrix **M** to be a Mueller matrix in terms of the eigenvalue structure of $\mathbf{A}=\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}$. If **M** is a real 4×4 matrix and the two situations described in the statement of Theorem 3.2 are considered, then $\lambda_0\pm\lambda_1\geqslant|\lambda_2\pm\lambda_3|$ in the first situation and $\lambda\geqslant\frac{1}{2}|\mu+\nu|$ and $\mu=\nu$ in the second situation, i.e., one finds the same eigenvalue inequalities for **A** as obtained for **M** in Theorem 4.2. The reason is that **A** is a Mueller matrix if **M** is a Mueller matrix [cf. Eq. (A30)]. The converse is not necessarily true even if **M** were to satisfy the Stokes criterion, as exemplified by the case in which $\mathbf{M}=\mathbf{G}\mathbf{U}$ with $\mathbf{U}\in\mathscr{L}$. One strategy to solve the problem is to give necessary and sufficient conditions for a real 4×4 matrix **M** to be representable in the form $\mathbf{M}=\mathbf{U}\mathbf{D}$, where $\mathbf{U}\in\mathscr{L}$ and **D** is **G** symmetric, and then to relate the eigenvalue structure of $\mathbf{D}$ to the eigenvalue structure of $\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}$, observing that $\tilde{\mathbf{M}}\mathbf{G}\mathbf{M}=\mathbf{D}^2$. Since Theorem 4.2 applies to $\mathbf{D}$, a result analogous to Theorem 3.2 would result for Mueller matrices. Research on the polar decomposition problem (writing $\mathbf{M}=\mathbf{U}\mathbf{D}$ with $\mathbf{U}\in\mathscr{L}$ and $\mathbf{D}$ being $\mathbf{G}$ symmetric) is in progress.

Cloude's criterion for a real 4×4 matrix to be a Mueller matrix can be applied to any such matrix, contrary to Theorem 4.2 which only applies to $\mathbf{G}$-symmetric matrices. However, the eigenvalue structure can be used to determine if a given real 4×4 matrix satisfies the Stokes criterion, and here Cloude's criterion cannot be applied. Moreover, checking a somewhat different eigenvalue inequality would tell one immediately if the matrix already found to satisfy the Stokes criterion is also a Mueller matrix. For $\mathbf{G}$-symmetric matrices we can actually do it; for general real 4×4 matrices some work still is to be done.

ACKNOWLEDGMENTS

The author is greatly indebted to S. R. Cloude, J. W. Hovenier, and A. C. M. Ran for stimulating discussions and to J. W. Hovenier for his comments on various drafts of the article.

APPENDIX: THE COHERENCY MATRIX

In this appendix, we introduce the coherency matrix[13,25] and give an alternative proof of Cloude's criterion for a real 4×4 matrix to be a Mueller matrix. We do not use group theory. A different matrix was introduced by Simon[18] in a study of pure Mueller matrices, because apart from a positive constant factor this matrix is unitarily equivalent to Cloude's coherency matrix.

Consider the identity 2×2 matrix σ_0 and the three Pauli matrices σ_1, σ_2, and σ_3 defined by

$$\sigma_0=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \sigma_1=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \sigma_2=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_3=\begin{pmatrix}0&-i\\i&0\end{pmatrix}. \tag{A1}$$

These matrices are Hermitian, form a basis of the vector space of complex 2×2 matrices, and satisfy $\sigma_0^2=\sigma_1^2=\sigma_2^2=\sigma_3^2=\text{identity}$; $\sigma_1\sigma_2=i\sigma_3$, $\sigma_2\sigma_3=i\sigma_1$, $\sigma_3\sigma_1=i\sigma_2$, $\sigma_2\sigma_1=-i\sigma_3$, $\sigma_3\sigma_2=-i\sigma_1$, $\sigma_1\sigma_3=-i\sigma_2$; $\frac{1}{2}\operatorname{Tr}(\sigma_j)=\delta_{0,j}$.

For any intensity vector $\mathbf{I}=\{I,Q,U,V\}$ we introduce the polarization density matrix[26,27]

$$\mathbf{W}=\sum_{r=0}^{3}[\mathbf{I}]_r\sigma_r=\begin{pmatrix}I+Q&U-iV\\U+iV&I-Q\end{pmatrix} \tag{A2}$$

so that

$$[\mathbf{I}]_r=\tfrac{1}{2}\operatorname{Tr}(\mathbf{W}\sigma_r)=\tfrac{1}{2}\operatorname{Tr}(\sigma_r\mathbf{W}). \tag{A3}$$

Let us derive an expression for the pure Mueller matrix $\mathbf{M}$ defined by

$$\mathbf{I}=\mathbf{M}\mathbf{I}_0, \tag{A4}$$

where we have the dyadic products

$$\mathbf{W}=\mathbf{E}\mathbf{E}^\dagger,\quad \mathbf{W}_0=\mathbf{E}_0\mathbf{E}_0{}^\dagger \tag{A5}$$

and † denotes the Hermitian conjugate. The E fields are connected by the Jones matrix $\mathbf{J}$ defined by

$$\mathbf{E}=\mathbf{J}\mathbf{E}_0,\quad \mathbf{E}^\dagger=\mathbf{E}_0{}^\dagger\mathbf{J}^\dagger. \tag{A6}$$

Using Eqs. (A2), (A3), (A5), and (A6) we get

$$[\mathbf{I}]_r=\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{W}\sigma_r)=\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{J}^\dagger\sigma_r\,\mathbf{J}\mathbf{E}_0\mathbf{E}_0{}^\dagger)=\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{J}^\dagger\sigma_r\,\mathbf{J}\mathbf{W}_0)=\sum_{s=0}^{3}\,[\mathbf{I}_0]_s\cdot\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{J}^\dagger\sigma_r\,\mathbf{J}\sigma_s) \tag{A7}$$

so that [cf. Eq. (A4)]

$$[\mathbf{M}]_{r+1,s+1}=\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{J}^\dagger\sigma_r\,\mathbf{J}\sigma_s). \tag{A8}$$

Next, let us expand the matrix $\mathbf{J}$ in σ_j as follows:

$$\mathbf{J}=\begin{pmatrix} A_2 & A_3 \\ A_4 & A_1 \end{pmatrix}=\sum_{t=0}^{3} k_t\sigma_t. \tag{A9}$$

Substituting Eq. (A9) and its Hermitian conjugate equation into Eq. (A8) we obtain

$$[\mathbf{M}]_{r+1,s+1}=\sum_{t=0}^{3}\sum_{u=0}^{3}\tfrac{1}{2}\,\mathrm{Tr}(\sigma_r\,\sigma_t\,\sigma_s\,\sigma_u)\,[\mathbf{T}]_{tu}, \tag{A10}$$

where

$$[\mathbf{T}]_{rs}=k_r\,k_s^*. \tag{A11}$$

It is immediate from Eq. (A11) that $\mathbf{T}$ is a Hermitian matrix with eigenvalues $\{\lambda,0,0,0\}$ where $\lambda=\Sigma_{r=0}^{3}|k_r|^2$. The matrix $\mathbf{T}$ is called the *coherency matrix.*[13,25]

Equation (A10) is equivalent to a system of 16 linear equations with 16 unknowns, which reduces to four systems of four equations with four unknowns. Calculating those 256 coefficients with the help of the properties of the Pauli matrices and inverting each of the four ensuing systems of equations we get

$$\left.\begin{aligned} T_{00}&=\tfrac{1}{4}(M_{11}+M_{22}+M_{33}+M_{44})\\ T_{11}&=\tfrac{1}{4}(M_{11}+M_{22}-M_{33}-M_{44})\\ T_{22}&=\tfrac{1}{4}(M_{11}-M_{22}+M_{33}-M_{44})\\ T_{33}&=\tfrac{1}{4}(M_{11}-M_{22}-M_{33}+M_{44}) \end{aligned}\right\}, \tag{A12}$$

$$\left.\begin{aligned} T_{03}&=\tfrac{1}{4}(M_{14}-iM_{23}+iM_{32}+M_{41})\\ T_{12}&=\tfrac{1}{4}(iM_{14}+M_{23}+M_{32}-iM_{41})\\ T_{21}&=\tfrac{1}{4}(-iM_{14}+M_{23}+M_{32}+iM_{41})\\ T_{30}&=\tfrac{1}{4}(M_{14}+iM_{23}-iM_{32}+M_{41}) \end{aligned}\right\}, \tag{A13}$$

$$\left.\begin{aligned} T_{01}&=\tfrac{1}{4}(M_{12}+M_{21}-iM_{34}+iM_{43})\\ T_{10}&=\tfrac{1}{4}(M_{12}+M_{21}+iM_{34}-iM_{43})\\ T_{23}&=\tfrac{1}{4}(iM_{12}-iM_{21}+M_{34}+M_{43})\\ T_{32}&=\tfrac{1}{4}(-iM_{12}+iM_{21}+M_{34}+M_{43}) \end{aligned}\right\}, \tag{A14}$$

$$\left.\begin{aligned} T_{02}&=\tfrac{1}{4}(M_{13}+M_{31}+iM_{24}-iM_{42})\\ T_{20}&=\tfrac{1}{4}(M_{13}+M_{31}-iM_{24}+iM_{42})\\ T_{13}&=\tfrac{1}{4}(-iM_{13}+iM_{31}+M_{24}+M_{42})\\ T_{31}&=\tfrac{1}{4}(iM_{13}-iM_{31}+M_{24}+M_{42}) \end{aligned}\right\}. \tag{A15}$$

It turns out that the coefficients of the equations expressing **T** in **M** are exactly one quarter of the coefficients of the equations expressing **M** in **T**.

The following result is due to Cloude.[13,14]

THEOREM A.1: *Let* **M** *be a real* 4×4 *matrix, and define* **T** *by Eqs.* (A12)–(A15). *Then* **T** *is a Hermitian* 4×4 *matrix, i.e.* $\mathbf{T}^\dagger=\mathbf{T}$. *Moreover,*

(1) **M** *is a Mueller matrix if and only if the eigenvalues of* **T** *are nonnegative*;

(2) **M** *is a pure Mueller matrix if and only if the eigenvalues of* **T** *are* $\{\lambda,0,0,0\}$ *with* $\lambda\geqslant 0$.

Proof: If **M** is a pure Mueller matrix and is given by Eq. (A8) for some Jones matrix **J**, then **T** is given by Eq. (A11) and the eigenvalues of **T** are $\{\lambda,0,0,0\}$ where $\lambda=\mathbf{k}^\dagger\mathbf{k}\geqslant 0$. Conversely, if the eigenvalues of **T** are $\{\lambda,0,0,0\}$ with $\lambda\geqslant 0$, let **k** be a nontrivial complex four-vector such that $\mathbf{T}\mathbf{k}=\lambda\mathbf{k}$. Then $\lambda\geqslant 0$ being the only nonzero eigenvalue of **T** means that $[\mathbf{T}]_{rs} = k_r k_s^*$ and $\lambda=\Sigma_{r=0}^3|k_r|^2$. If $\mathbf{k}=\{k_0,k_1,k_2,k_3\}$, define the Jones matrix **J** by Eq. (A9). Then [cf. Eq. (A11)]

$$\tfrac{1}{2}\,\mathrm{Tr}(\mathbf{J}^\dagger\sigma_r\,\mathbf{J}\sigma_s)=\sum_{t=0}^{3}\sum_{u=0}^{3}\tfrac{1}{2}\,\mathrm{Tr}(\sigma_r\,\sigma_t\,\sigma_s\,\sigma_u)\,[\mathbf{T}]_{tu}=[\mathbf{M}]_{r+1,s+1}, \tag{A16}$$

which settles the pure Mueller part of the theorem.

Next, let **M** be an arbitrary matrix of the form

$$\mathbf{M}=\sum_{l=1}^{N}\gamma_l\mathbf{M}_{\mathbf{J}_l}, \tag{A17}$$

where $\gamma_1,\gamma_2,...,\gamma_N\geqslant 0$ and $\mathbf{M}_{\mathbf{J}_1},...,\mathbf{M}_{\mathbf{J}_N}$ are the pure Mueller matrices obtained from the Jones matrices $\mathbf{J}_1,...,\mathbf{J}_N$. Applying Eqs. (A12)–(A15)

$$\mathbf{T}=\sum_{l=1}^{N}\gamma_l\mathbf{T}_{\mathbf{J}_l}, \tag{A18}$$

where $\mathbf{T}_{\mathbf{J}_1},...,\mathbf{T}_{\mathbf{J}_N}$ are defined in terms of $\mathbf{M}_{\mathbf{J}_1},...,\mathbf{M}_{\mathbf{J}_N}$ by Eqs. (A12)–(A15). Since every matrix $\mathbf{T}_{\mathbf{J}_1},...,\mathbf{T}_{\mathbf{J}_N}$ is positive semidefinite and $\gamma_1,...,\gamma_N\geqslant 0$, **T** must also be positive semidefinite.

Conversely, suppose the matrix **T** obtained from **M** using Eqs. (A12)–(A15) is positive semidefinite, then there exists a unitary 4×4 matrix **U** (i.e., $\mathbf{U}^\dagger=\mathbf{U}^{-1}$) such that

$$\mathbf{T}=\mathbf{U}\Lambda\mathbf{U}^\dagger,\quad \Lambda=\mathrm{diag}(\lambda_0,\lambda_1,\lambda_2,\lambda_3), \tag{A19}$$

where λ_0, λ_1, λ_2, $\lambda_3\geqslant 0$. Let us denote the four columns of **U** by $\mathbf{k}_0$, $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$. Then $\mathbf{k}_0$, $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$ form an orthonormal basis and $\mathbf{T}\mathbf{k}_r=\lambda_r\,\mathbf{k}_r$ for $r=0,1,2,3$. We then readily derive

$$\mathbf{T}=\sum_{r=0}^{3}\lambda_r\,\mathbf{k}_r\,\mathbf{k}_r^{\dagger}. \tag{A20}$$

Writing

$$\mathbf{J}_r = \sum_{s=0}^{3} [\mathbf{k}_r]_s \sigma_s \tag{A21}$$

we get

$$\mathbf{T}_{\mathbf{J}_r} = \mathbf{k}_r \mathbf{k}_r^\dagger, \quad r=0,1,2,3. \tag{A22}$$

As a result

$$\mathbf{M} = \sum_{r=0}^{3} \lambda_r \mathbf{M}_{\mathbf{J}_r}, \tag{A23}$$

with $\mathbf{M}_{\mathbf{J}_r}$ the pure Mueller matrix obtained from the Jones matrix $\mathbf{J}_r$. ■

COROLLARY A.2.[28] *A block-diagonal real* 4×4 *matrix* $\mathbf{M}$ *of the form*

$$\mathbf{M} = \begin{pmatrix} a_1 & b_1 & 0 & 0 \\ c_1 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & b_2 \\ 0 & 0 & c_2 & a_4 \end{pmatrix} \tag{A24}$$

is a Mueller matrix if and only if

$$[(a_3-a_4)^2+(b_2+c_2)^2+(b_1-c_1)^2]^{1/2} \leqslant a_1-a_2, \tag{A25}$$

$$[(a_3+a_4)^2+(b_2-c_2)^2+(b_1+c_1)^2]^{1/2} \leqslant a_1+a_2. \tag{A26}$$

Further, $\mathbf{M}$ *is a pure Mueller matrix if and only if*

$$a_1=a_2>0, \quad a_3=a_4, \quad b_1=c_1, \quad b_2=-c_2, \quad a_1=[a_3^2+b_1^2+b_2^2]^{1/2}. \tag{A27}$$

Proof: We readily compute the eigenvalues of the coherency matrix. We find

$$\lambda_{0,1} = \tfrac{1}{4}(a_1+a_2) \pm \tfrac{1}{4}[(a_3+a_4)^2+(b_1+c_1)^2+(b_2-c_2)^2]^{1/2}, \tag{A28}$$

$$\lambda_{2,3} = \tfrac{1}{4}(a_1-a_2) \pm \tfrac{1}{4}[(a_3-a_4)^2+(b_1-c_1)^2+(b_2+c_2)^2]^{1/2}. \tag{A29}$$

Requiring all four eigenvalues to be non-negative, we get the result. ■

Using $\mathbf{G}=\mathrm{diag}(1,-1,-1,-1)$ and writing $\mathbf{T}(\mathbf{M})$ for the coherency matrix associated with the real 4×4 matrix $\mathbf{M}$ through Eqs. (A12)–(A15), we easily find the following identity:

$$\mathbf{T}(\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}) = \mathbf{G}\cdot\mathbf{T}(\mathbf{M})\cdot\mathbf{G}. \tag{A30}$$

As a result, $\mathbf{M}$ is $\mathbf{G}$ symmetric whenever $\mathbf{T}(\mathbf{M})$ commutes with $\mathbf{G}$. Moreover, since $\mathbf{T}(\mathbf{M})$ and $\mathbf{G}\cdot\mathbf{T}(\mathbf{M})\cdot\mathbf{G}$ are obviously unitarily equivalent, $\mathbf{G}\tilde{\mathbf{M}}\mathbf{G}$ is a Mueller matrix whenever $\mathbf{M}$ is a Mueller matrix.

[1] In this article we use the term "Stokes vector" for a four-vector $\{I,Q,U,V\}$ as above, where $I(I\geqslant 0)$ may be an intensity, source function or flux, and Q/I, U/I, and V/I describe the state of polarization.
[2] S. Chandrasekhar, *Radiative Transfer* (Oxford University, London, New York, 1950); also (Dover, New York, 1960).
[3] H. C. van de Hulst, *Light Scattering by Small Particles* (Wiley, New York, 1957); also (Dover, New York, 1981).
[4] J. W. Hovenier and C. V. M. van der Mee, Astron. Astrophys. **128**, 1 (1983).

[5] W. A. Shurcliff, *Polarized Light. Production and Use* (Harvard University, Cambridge, MA, and Oxford University, London, 1962).
[6] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).
[7] N. V. Konovalov, *Polarization Matrices Corresponding to Transformations in the Stokes Cone*, Preprint **171**, Keldysh Inst. Appl. Math. [Acad. Sci. USSR, Moscow, 1985 (Russian)].
[8] C. V. M. van der Mee and J. W. Hovenier, J. Math. Phys. **33**, 3574 (1992).
[9] D. I. Nagirner, Astron. Astrophys. **275**, 318 (1993).
[10] K. A. O'Donnell and M. E. Knotts, J. Opt. Soc. Am. A **8**, 1126 (1991).
[11] Zhang-Fan Xing, J. Mod. Opt. **39**, 461 (1992).
[12] R. Barakat, Opt. Commun. **38**, 159 (1981).
[13] S. R. Cloude, Optik **75**, 26 (1986).
[14] S. R. Cloude, *Conditions for the Physical Realisability of Matrix Operators in Polarimetry*, in Polarization Considerations for Optical Systems, II, edited by R. A. Chipman (SPIE, 1989), Vol. 1166.
[15] R. M. A. Azzam, J. Opt. Soc. Am. **68**, 1756 (1978).
[16] E. S. Fry and G. W. Kattawar, Appl. Opt. **20**, 2811 (1981).
[17] J. W. Hovenier, H. C. van de Hulst, and C. V. M. van der Mee, Astron. Astrophys. **157**, 301 (1986).
[18] R. Simon, Opt. Commun. **42**, 293 (1982).
[19] R. Simon, J. Mod. Opt. **34**, 569 (1987).
[20] J. W. Hovenier, *Structure of Pure Mueller Matrices*, Appl. Opt. (submitted, 1993).
[21] J. Bognár, *Indefinite Inner Product Spaces* (Springer, Berlin, Heidelberg, 1974).
[22] I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products* (Birkhäuser OT 8, Basel, Boston, 1983).
[23] F. D. Murnaghan, *The Theory of Group Representations* (Dover, New York, 1938).
[24] R. Simon, Opt. Commun. **77**, 349 (1990).
[25] J. R. Huynen, Ph.D. thesis, Delft Technical University, The Netherlands, 1970.
[26] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields. Part III. Quantum Mechanics* (Pergamon, Oxford, 1965).
[27] M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, Oxford, 1989).
[28] Necessary and sufficient conditions for the matrix (A24) to satisfy the Stokes criterion were given in Ref. 8. In the special case where $b_1=c_1$ and $b_2=-c_2$, such conditions also appear in Refs. 7 and 9.