A complete characterization of pre-Mueller and Mueller matrices in polarization optics

B. N. Simon,[1] S. Simon,[2] N. Mukunda,[3] F. Gori,[4],* M. Santarsiero,[4] R. Borghi,[5] and R. Simon[6]

[1]*Department of Applied Mechanics, IIT Madras, Chennai 600 036, India*
[2]*Department of Computer Science, Women's Christian College, Chennai 600 006, India*
[3]*Centre for High Energy Physics, Indian Institute of Science, Bangalore 560 012, India*
[4]*Dipartimento di Fisica, Universitá Roma Tre and CNISM, Via della Vasca Navale 84, I-00146 Rome, Italy*
[5]*Dipartimento di Elettronica Applicata, Universitá Roma Tre and CNISM, Via della Vasca Navale 84, I-00146 Rome, Italy*
[6]*The Institute of Mathematical Science, Tharamani, Chennai 600 113, India*
*Corresponding author: gori@uniroma3.it

The Mueller–Stokes formalism that governs conventional polarization optics is formulated for plane waves, and thus the only qualification one could require of a 4×4 real matrix M in order that it qualify to be the Mueller matrix of some physical system would be that M map $\Omega^{(pol)}$, the positive solid light cone of Stokes vectors, into itself. In view of growing current interest in the characterization of partially coherent partially polarized electromagnetic beams, there is a need to extend this formalism to such beams wherein the polarization and spatial dependence are generically inseparably intertwined. This inseparability brings in additional constraints that a pre-Mueller matrix M mapping $\Omega^{(pol)}$ into itself needs to meet in order to be an acceptable physical Mueller matrix. These additional constraints are motivated and fully characterized. © 2010 Optical Society of America

OCIS codes: 260.5430, 270.5585, 260.2110, 120.5410, 230.5440.

1. INTRODUCTION

A paraxial beam propagating along the positive z axis is completely determined in terms of the transverse components of the electric field, specified throughout a transverse plane z = constant as functions of the transverse coordinates $(x, y) = \boldsymbol{\rho}$. If these components are independent of the transverse coordinates, then the situation corresponds to a plane wave propagating along the z axis. The traditional Mueller–Stokes formalism in terms of Stokes vector S and Mueller matrix M, describing, respectively, the beam and the optical system, presumes essentially this kind of situation wherein the spatial degree of freedom can be safely left out of consideration, the focus being on the polarization degree of freedom [1–3].

Recent years have witnessed an enormous interest in partially polarized partially coherent electromagnetic beams [4–20], and hence there is a need to extend the Mueller–Stokes formalism to such beams. Given a 4×4 real matrix M, it should necessarily map $\Omega^{(pol)}$, the positive cone of Stokes vectors, into itself in order that it could be the Mueller matrix of some physical system. Within the conventional formalism, this seems to be the only qualification that can be required of M. In a partially coherent partially polarized beam, polarization and spatial dependence happen to be inseparably intertwined. This inseparability brings in additional constraints that a 4×4 real matrix M mapping $\Omega^{(pol)}$ into itself needs to meet in order for it to be a physically acceptable Mueller matrix. The class of 4×4 real matrices that one would have hitherto believed to be Mueller matrices now must pass

additional tests before they can qualify to be physical Mueller matrices; and this will seem to suggest that these matrices should more properly be called pre-Mueller matrices, so that those pre-Mueller matrices that pass this additional physical requirement to be developed here can be called Mueller matrices.

We notice that the inseparability of polarization and spatial dependence can be seen as a classical analog of quantum entanglement, which is traditionally studied almost exclusively in the context of quantum systems. However, this notion is basically kinematic in nature, being a direct consequence of the superposition principle, and so it is bound to present itself whenever and wherever the state space of interest is the tensor product of two (or more) linear vector spaces. The vectors of the individual spaces, and hence (tensor) products of such vectors, will be expected to possess identifiable physical meaning. Polarization optics of paraxial electromagnetic beams happens to have precisely this kind of a setting, with a two-dimensional vector space describing the polarization degree of freedom and an infinite-dimensional vector space of square integrable functions describing the spatial degree of freedom. We could then refer to the inseparability of polarization and spatial dependence as a manifestation of non-quantum entanglement.

We hasten, however, to add a note by way of clarification. In quantum theory, entanglement conspires with the (truly nonclassical) measurement postulate, and the associated collapse of states, to produce dramatic and even seemingly paradoxical consequences, the Einstein–

Podolsky–Rosen paradox [21,22] being an early example. But classical linear optics does not suffer from the luxury or burden of this kind of measurement postulate, and hence only those consequences of entanglement that do not depend on the measurement postulate in an essential manner should be expected to present themselves in this context. This is exactly the case for the phenomena of our interest. However, since use of the phrase "non-quantum entanglement" could sound odd to some readers, we shall in short refer to the mingling of polarization and spatial variation (or modulation) as the "inseparability."

The principal purpose of the present work is to motivate the constraints arising from inseparability and characterize them fully. The next two sections of the paper act as preparation toward this end. We begin in Section 2 by recounting the conventional Mueller–Stokes formalism as it applies to plane waves. This is then extended in Section 3 to paraxial electromagnetic beams, and the role of inseparability (of polarization and spatial variation) is rendered transparent. These two sections equip us with all the tools needed to formulate in Section 4 the additional physical constraints on a pre-Mueller matrix M arising as consequence of inseparability. Our final result is formulated in the form of a necessary and sufficient condition, and a simple illustrative example is treated in some detail to clarify the nature of these further constraints. We conclude in Section 5 with some further remarks.

2. POLARIZATION OPTICS OF PLANE WAVES

For a plane wave whose propagation direction is along the (positive) z axis perpendicular to the (x,y) plane, the x and y components E_1, E_2 of the electric field are independent of the transverse-plane coordinates $\boldsymbol{\rho}$ and can be arranged into a (numerical) column vector:

$$\mathbf{E} \equiv \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \in \mathcal{C}^2. \qquad (2.1)$$

We have suppressed, for convenience, a space–time-dependent scalar factor of the form $e^{i(kz-\omega t)}$. While $\mathbf{E}^\dagger\mathbf{E} = |E_1|^2 + |E_2|^2$ is (a measure of) the intensity, the ratio $\gamma = E_1/E_2$ of the (complex) components, which ratio can be viewed as a point on the Riemann or Poincaré sphere S^2, specifies the polarization state. In particular, the signature of the imaginary part of γ describes the handedness of the (generally elliptic) polarization.

In the presence of fluctuations, $\mathbf{E}$ acquires some randomness, and in this case the state of polarization is effectively described by the 2×2 coherency or polarization matrix,

$$\Phi \equiv \langle \mathbf{EE}^\dagger \rangle = \begin{bmatrix} \langle E_1 E_1^* \rangle & \langle E_1 E_2^* \rangle \\ \langle E_2 E_1^* \rangle & \langle E_2 E_2^* \rangle \end{bmatrix}, \qquad (2.2)$$

where $\langle \cdots \rangle$ denotes ensemble average. The coherency matrix is Hermitian, $\Phi^\dagger = \Phi$, and positive semidefinite, $\mathbf{V}^\dagger\Phi\mathbf{V} = \mathrm{tr}(\Phi\mathbf{VV}^\dagger) \geqslant 0$, $\forall \mathbf{V} \in \mathcal{C}^2$. This positivity property may be denoted simply as $\Phi \geqslant 0$. Hermiticity and positivity are the defining properties of Φ: every 2×2 matrix obeying these two conditions is a valid coherency

matrix and represents some polarization state. Since Φ is a 2×2 matrix, the positivity condition takes the simple scalar form

$$\mathrm{tr}\,\Phi > 0,$$

$$\det \Phi \geqslant 0. \qquad (2.3)$$

It is clear that the intensity corresponds to $\mathrm{tr}\,\Phi$. Fully polarized light (pure states) corresponds to $\det \Phi = 0$ and partially polarized or mixed states to $\det \Phi > 0$.

Typical systems of interest in polarization optics are transversely homogeneous, in the sense that their action is independent of the coordinates (x,y) spanning the transverse plane in which the system lies. If such a system is deterministic and acts linearly at the field amplitude level, it is described by a complex 2×2 numerical matrix J called the Jones matrix of the system:

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}: \quad \mathbf{E} \to \mathbf{E}' = \begin{bmatrix} E_1' \\ E_2' \end{bmatrix} = J\mathbf{E} \Leftrightarrow \Phi \equiv \langle \mathbf{EE}^\dagger \rangle \to \Phi'$$

$$= \langle \mathbf{E}'\mathbf{E}'^\dagger \rangle = J\Phi J^\dagger. \qquad (2.4)$$

It is clear that Jones systems map pure states $(\det \Phi = 0)$ into pure states.

Since Φ is Hermitian, it can be conveniently described as *real* linear combination of the four orthogonal Hermitian matrices $\tau_0 = 1_{2\times 2}$, $\tau_1 = \sigma_3$, $\tau_2 = \sigma_1$, $\tau_3 = \sigma_2$:

$$\Phi = \frac{1}{2}\sum_{a=0}^{3} S_a \tau_a \Leftrightarrow S_a = \mathrm{tr}(\tau_a \Phi); \quad \mathrm{tr}\,\tau_a\tau_b = 2\delta_{ab}. \quad (2.5)$$

The reason for choosing the τ matrices, a permuted version of the Pauli matrices rather than the Pauli σ matrices themselves, is to be consistent with the optical convention that the circularly polarized states, the eigenstates of $\sigma_2 = \tau_3$, be along the "third" axis (polar axis) of the Poincaré sphere. The intensity equals $S_0 = \mathrm{tr}\,\Phi$. The expansion coefficients S_a are the components of the Stokes vector $S \in \mathbf{R}^4$. Note that $\tau_3^* = -\tau_3$, whereas $\tau_a^* = \tau_a$ if $a \neq 3$.

While Hermiticity of Φ is equivalent to reality of $S \in \mathbf{R}^4$, the positivity conditions $\mathrm{tr}\,\Phi > 0$, $\det \Phi \geqslant 0$ read, respectively,

$$S_0 > 0,$$

$$S_0^2 - S_1^2 - S_2^2 - S_3^2 \geqslant 0. \qquad (2.6)$$

Thus, permissible polarization states correspond to the positive light cone and its interior (solid cone). Pure states live on the surface of this cone. As suggested by this light cone structure, the proper orthochronous Lorentz group $SO(3,1)$ plays quite an important role in polarization optics [23,24].

Under the action of a deterministic or Jones system J described in Eq. (2.4), the elements of the output coherency matrix Φ' are obviously linear in those of Φ. This, in view of the linear relation (2.5) between Φ and S, implies that under passage through such a system the output Stokes vector S' and the input S will be linearly related by a 4×4 real matrix $M(J)$ determined by J:

$$J : S \rightarrow S' = M(J)S. \tag{2.7}$$

We may call $M(J)$ the Mueller matrix of the Jones system J. It is known also as a Mueller–Jones matrix to emphasize the fact that it is constructed out of a Jones matrix. While $\Phi = \langle \mathbf{E}\mathbf{E}^\dagger \rangle$ is a 2×2 matrix, the tensor product $\tilde{\Phi} \equiv \langle \mathbf{E} \otimes \mathbf{E}^* \rangle$ is a four-dimensional column vector associated with Φ:

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\Phi}_0 \\ \tilde{\Phi}_1 \\ \tilde{\Phi}_2 \\ \tilde{\Phi}_3 \end{bmatrix} \equiv \begin{bmatrix} \langle E_1 E_1^* \rangle \\ \langle E_1 E_2^* \rangle \\ \langle E_2 E_1^* \rangle \\ \langle E_2 E_2^* \rangle \end{bmatrix} = \begin{bmatrix} \Phi_{11} \\ \Phi_{12} \\ \Phi_{21} \\ \Phi_{22} \end{bmatrix}. \tag{2.8}$$

This idea of going from a pair of indices, each running over 1 and 2, to a single index running over 0 to 3 and vice versa can often be used to advantage to associate with any 2×2 matrix K a corresponding column vector $\tilde{K}$ with $\tilde{K}_0 = K_{11}$, $\tilde{K}_1 = K_{12}$, $\tilde{K}_2 = K_{21}$, and $\tilde{K}_3 = K_{22}$. The tensor product $J \otimes J^*$ is a 4×4 matrix:

$$J \otimes J^* \equiv \begin{bmatrix} J_{11}J^* & J_{12}J^* \\ J_{21}J^* & J_{22}J^* \end{bmatrix}$$

$$= \begin{bmatrix} J_{11}J_{11}^* & J_{11}J_{12}^* & J_{12}J_{11}^* & J_{12}J_{12}^* \\ J_{11}J_{21}^* & J_{11}J_{22}^* & J_{12}J_{21}^* & J_{12}J_{22}^* \\ J_{21}J_{11}^* & J_{21}J_{12}^* & J_{22}J_{11}^* & J_{22}J_{12}^* \\ J_{21}J_{21}^* & J_{21}J_{22}^* & J_{22}J_{21}^* & J_{22}J_{22}^* \end{bmatrix}. \tag{2.9}$$

The transformation $\Phi \rightarrow \Phi' = J\Phi J^\dagger$ is thus equivalent to $\tilde{\Phi} \rightarrow \tilde{\Phi}' = J \otimes J^* \tilde{\Phi}$. Since $\tilde{\Phi}$ is related to the Stokes vector through

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Phi}_0 \\ \tilde{\Phi}_1 \\ \tilde{\Phi}_2 \\ \tilde{\Phi}_3 \end{bmatrix}, \tag{2.10}$$

it follows that

$$M(J) = A(J \otimes J^*)A^{-1}, \tag{2.11}$$

A being the numerical 4×4 matrix exhibited in Eq. (2.10); this matrix is essentially unitary: $A^{-1} = \frac{1}{2}A^\dagger$.

If $\det J$ is of unit magnitude, then $M(J)$ computed by this prescription is an element of $SO(3,1)$, the proper orthochronous group of Lorentz transformations; this was to be expected in view of the two-to-one homomorphism between $SL(2,C)$ and $SO(3,1)$ [24]. It follows that for any nonsingular J the associated Mueller–Jones matrix $M(J)$ is $|\det J|$ times an element of $SO(3,1)$. The prescription (2.11), though, applies to singular Jones matrices as well.

A. Mueller Matrices Arising as Convex Sums of Jones Systems

A nondeterministic (i.e., non-Jones) system is described directly by a Mueller matrix $M : S \rightarrow S' = MS$, and, by definition, such a Mueller matrix cannot equal $M(J)$ for any 2×2 (Jones) matrix J. Given a Mueller matrix M, how

does one test whether it is a Mueller–Jones matrix for some J or, equivalently, how does one test whether the system described by M is a deterministic (i.e., Jones) system? This question, which had received much attention [23,25–27], turned out to have a simple and elegant solution [28]. We go over in some detail the construction underlying this solution, for it plays a key role in our analysis to follow. Of central importance is the construction of a Hermitian matrix associated with each real matrix M; this construction has played a nontrivial role in subsequent developments of polarization optics. This Hermitian matrix was originally denoted N [28], but now we prefer to denote it $H^{(M)}$ to emphasize Hermiticity, its most important property.

The general real linear transformation $M : S \rightarrow S' = MS$ on the Stokes vectors S implies, in view of the linear relationship (2.5) or (2.10) between S and Φ, an associated linear transformation on Φ. Indeed, use of $S = A\tilde{\Phi}$, Eq. (2.10), in this transformation law immediately gives the 4×4 matrix $B^{(M)}$ transforming Φ linearly:

$$B^{(M)} : \Phi \rightarrow \Phi', \qquad \tilde{\Phi}' = B^{(M)}\tilde{\Phi},$$

$$\text{i.e.,} \quad \Phi'_{ij} = \sum_{k\ell=1}^{2} B^{(M)}_{ij,k\ell} \Phi_{k\ell};$$

$$B^{(M)} \equiv A^{-1}MA, \qquad M = AB^{(M)}A^{-1}. \tag{2.12}$$

Now define from $B^{(M)}$ a new matrix $H^{(M)}$ by permuting the indices of $B^{(M)}$:

$$H^{(M)}_{ik,j\ell} = B^{(M)}_{ij,k\ell}. \tag{2.13}$$

The transformation (2.12) now gets transcribed to

$$H^{(M)} : \Phi \rightarrow \Phi', \qquad \Phi'_{ij} = \sum_{k\ell} H^{(M)}_{ik,j\ell} \Phi_{k\ell}. \tag{2.14}$$

Note that $H^{(M)}$ is obtained from $B^{(M)}$ by simply interchanging $B^{(M)}_{02}$ with $B^{(M)}_{10}$, $B^{(M)}_{03}$ with $B^{(M)}_{11}$, $B^{(M)}_{22}$ with $B^{(M)}_{30}$, and $B^{(M)}_{23}$ with $B^{(M)}_{31}$. Further, while $\tilde{\Phi}' = B^{(M)}\tilde{\Phi}$ can be usefully viewed as a (column) vector equation, the corresponding transformation (2.14) involving $H^{(M)}$ cannot be so viewed. The fact that the output Φ' is Hermitian for all Hermitian input Φ shows that the map, or super-operator $H^{(M)}$, viewed as a 4×4 matrix with ik (going over 0 to 3) labeling the rows and $j\ell$ labeling the columns, is Hermitian. This correspondence between real matrices M and Hermitian matrices $H^{(M)}$ is clearly one-to-one. Elements of $H^{(M)}$ in terms of those of M have been presented as Eq. (8) of [28].

The sixteen 4×4 *Hermitian* matrices $U_{ab} = \frac{1}{2}\tau_a \otimes \tau_b^*$, with a, b independently running over the index set {0, 1, 2, 3}, form an orthonormal set or basis in the (vector) space of 4×4 matrices; these matrices are unitary and self-inverses:

$$U_{ab} = \frac{1}{2}\tau_a \otimes \tau_b^* = U_{ab}^\dagger = U_{ab}^{-1},$$

$$\mathrm{tr}(U_{ab}U_{cd}) = \delta_{ac}\delta_{bd}, \qquad a,b,c,d \in \{1,2,3,4\}. \tag{2.15}$$

[For the purpose on hand, complex conjugation of the second factor of the tensor product $\tau_a \otimes \tau_b^*$ is not optional, and this will become clear in a moment.] Thus every (Hermitian) 4×4 matrix can be written uniquely as a (real) linear combination of $\{\tau_a \otimes \tau_b^*\}$. As an important consequence of this fact we have

Proposition 1: There exists a natural one-to-one correspondence between the set of all 4×4 real matrices and the set of all 4×4 Hermitian matrices.

To appreciate the sense in which this correspondence is natural, let us write the hermitian matrix $H^{(M)}$ as a linear combination of the $\tau_a \otimes \tau_b^*$'s and provisionally denote the expansion coefficients by K_{ab}:

$$H^{(M)} = \frac{1}{2}\sum_{a,b=0}^{3} K_{ab}\tau_a \otimes \tau_b^*. \tag{2.16}$$

Equation (2.14) then reads

$$\Phi'_{ij} = \sum_{ab} K_{ab}\sum_{k\ell} (\tau_a)_{ij}(\tau_b)_{k\ell}^* \Phi_{k\ell}. \tag{2.17}$$

Recalling that $\Sigma_{k\ell}(\tau_b)_{k\ell}^* \Phi_{k\ell} = \mathrm{tr}(\tau_b\Phi) = S_b$, the above equation reduces after multiplication by $(\tau_c)_{ij}^*$, summing over i,j, and using Eq. (2.5), to $S' = KS$, showing that what we provisionally denoted K indeed equals M. We have thus proved that

$$H^{(M)} = \frac{1}{2}\sum_{a,b=0}^{3} M_{ab}\tau_a \otimes \tau_b^*. \tag{2.18}$$

Using the orthogonality relations in Eq. (2.15), this equation can be readily inverted, and we find that a Hermitian matrix H and the associated real matrix $M^{(H)}$ are related through

$$(M^{(H)})_{ab} = \frac{1}{2}\mathrm{tr}(H\tau_a \otimes \tau_b^*), \qquad a,b = 0,1,2,3. \tag{2.19}$$

We write these in more detail for later use:

$$H^{(M)} = \frac{1}{2}\begin{bmatrix} M_{00}+M_{11}+M_{01}+M_{10} & M_{02}+M_{12}+i(M_{03}+M_{13}) & M_{20}+M_{21}-i(M_{30}+M_{31}) & M_{22}+M_{33}+i(M_{23}-M_{32}) \\ M_{02}+M_{12}-i(M_{03}+M_{13}) & M_{00}-M_{11}-M_{01}+M_{10} & M_{22}-M_{33}-i(M_{23}+M_{32}) & M_{20}-M_{21}-i(M_{30}-M_{31}) \\ M_{20}+M_{21}+i(M_{30}+M_{31}) & M_{22}-M_{33}+i(M_{23}+M_{32}) & M_{00}-M_{11}+M_{01}-M_{10} & M_{02}-M_{12}+i(M_{03}-M_{13}) \\ M_{22}+M_{33}-i(M_{23}-M_{32}) & M_{20}-M_{21}+i(M_{30}-M_{31}) & M_{02}-M_{12}-i(M_{03}-M_{13}) & M_{00}+M_{11}-M_{01}-M_{10} \end{bmatrix}. \tag{2.20}$$

This matrix in identical form was first presented in [28]. Of the sixteen 4×4 matrices U_{ab}, only U_{20}, U_{21}, U_{30}, and U_{31} have nonzero entries at the $_{13}$ location, and this explains the entry $M_{20}-M_{21}-i(M_{30}-M_{31})$ for $(H^{(M)})_{13}$. Written in detail, relation (2.19) has the form

$$M^{(H)} = \frac{1}{2}\begin{bmatrix} H_{00}+H_{11}+H_{22}+H_{33} & H_{00}-H_{11}+H_{22}-H_{33} & H_{01}+H_{10}+H_{23}+H_{32} & -i(H_{01}-H_{10})-i(H_{23}-H_{32}) \\ H_{00}+H_{11}-H_{22}-H_{33} & H_{00}-H_{11}-H_{22}+H_{33} & H_{01}+H_{10}-H_{23}-H_{32} & -i(H_{01}-H_{10})+i(H_{23}-H_{32}) \\ H_{02}+H_{20}+H_{13}+H_{31} & H_{02}+H_{20}-H_{13}-H_{31} & H_{03}+H_{30}+H_{12}+H_{21} & -i(H_{03}-H_{30})+i(H_{12}-H_{21}) \\ i(H_{02}-H_{20})+i(H_{13}-H_{31}) & i(H_{02}-H_{20})-i(H_{13}-H_{31}) & i(H_{03}-H_{30})+i(H_{12}-H_{21}) & H_{03}+H_{30}-H_{12}-H_{21} \end{bmatrix}. \tag{2.21}$$

The entry $-i(H_{01}-H_{10})+i(H_{23}-H_{32})$ for $(M^{(H)})_{13}$ is explained by the fact that the nonzero entries of U_{13} are at the $_{01}$, $_{10}$, $_{23}$, and $_{32}$ locations.

If the system described by M is a Jones system with Jones matrix J, it is clear from the transformation law $\Phi \to \Phi' = J\Phi J^\dagger$ given in Eq. (2.4) that $B^{(M)} = J\otimes J^*$ and, consequently, $H^{(M)} = \widetilde{J}\widetilde{J}^\dagger$, where $\widetilde{J}$ is the column vector associated with the 2×2 matrix J. The fact that the hermitian matrix associated with a Jones system is a one-dimensional projection $H^{(M)} = \widetilde{J}\widetilde{J}^\dagger$ immediately leads, in view of the one-to-one correspondence between M and $H^{(M)}$, to the fundamental result [28].

Proposition 2: Given a 4×4 real matrix M, it is a Mueller–Jones matrix iff the associated Hermitian matrix $H^{(M)}$ is a one-dimensional projection. That is, iff $H^{(M)} = \widetilde{J}\widetilde{J}^\dagger$ for some (complex) four-dimensional column vector $\widetilde{J}$. If $H^{(M)} = \widetilde{J}\widetilde{J}^\dagger$, then the 2×2 matrix J associated with $\widetilde{J}$ is the Jones matrix of the deterministic Jones system represented by M.

Consider now a transformation that is a convex sum of Jones systems:

$$\Phi \to \Phi' = \sum_{k=1}^{n} p_k J^{(k)}\Phi J^{(k)\dagger}, \tag{2.22}$$

$$p_k > 0, \qquad \sum_{k=1}^{n} p_k = 1.$$

This transformation may be realized by a set of n deterministic or Jones systems $J^{(1)}, J^{(2)}, \cdots, J^{(n)}$ arranged in parallel, with a fraction p_k of the light going through the kth Jones system $J^{(k)}$ and all the transformed beams com-

bined (incoherently) at the output. It can also be viewed as a fluctuating system that assumes the Jones form $J^{(k)}$ with probability p_k. In either case, it is clear that the Mueller matrix M of this nondeterministic system and its associated Hermitian matrix are corresponding convex sums:

$$M = \sum_{k=1}^{n} p_k M(J^k), \qquad H^{(M)} = \sum_{k=1}^{n} p_k \tilde{J}^{(k)} \tilde{J}^{(k)\dagger}. \quad (2.23)$$

It is useful to denote by $\{p_k, J^{(k)}\}$ the convex sum or ensemble realization represented by Eq. (2.22) or, equivalently, by Eq. (2.23). Obviously, such an ensemble or convex sum realization always leads to a positive semidefinite $H^{(M)}$. Further, it is an elementary fact that positive semidefinite H alone can be realized as a convex sum of projections. Thus as an immediate, and mathematically trivial, consequence of Proposition 2 we have [29]

Corollary: An optical system described by M is realizable as a convex sum or ensemble $\{p_k, J^{(k)}\}$ of Jones systems iff the associated $H^{(M)} \geqslant 0$. If $H^{(M)} \geqslant 0$, the number of Jones systems, n, needed for such a realization satisfies $n \geqslant r$, where r is the rank of $H^{(M)}$. There is no upper limit on n if $r \geqslant 2$.

This corollary is physically important and has attracted considerable attention [29–36].

B. Pre-Mueller Matrices and Their Classification
Given a 4×4 real matrix M, Proposition 2 gives the necessary and sufficient condition for M to arise as the Mueller matrix of some Jones system J. That still leaves open this more general question: how does one ascertain whether a given matrix M is a Mueller matrix? This question has an interesting history that is surprisingly recent.

In traditional polarization optics, which is formulated for plane waves and not for spatially varying beams, the state space $\Omega^{(\text{pol})}$ is the collection of all Stokes vectors:

$$\Omega^{(\text{pol})} = \{S \in \mathbf{R}^4 | S_0 > 0, \ S^T G S \geqslant 0\},$$

$$G = \text{diag}(1, -1, -1, -1),$$

$$S^T G S = S_0^2 - S_1^2 - S_2^2 - S_3^2. \quad (2.24)$$

Now G is the "Lorentz metric," and thus the state space $\Omega^{(\text{pol})}$ is the positive (solid) light cone in $\mathbf{R}^4$. Since a physical Mueller matrix should necessarily map states into states, our question reduces to one of effectively characterizing real linear transformations in $\mathbf{R}^4$ that map the positive (solid) light cone *into* itself. While $SO(3,1) \cup GSO(3,1)$, where $SO(3,1)$ is the proper orthochronous Lorentz group, is the obvious answer in the case of *onto* maps, the more general *into* case was raised in [37] as a serious issue in polarization optics. This issue was formulated as two simple conditions that M has to meet [Eqs. (2.29) and (2.31) of [37]], corresponding to the demand that the intensity and degree of polarization of the output be physical for every input pure state. Further, the measured Mueller matrices of Howell [38] were tested for these conditions, and violation was found in excess of 10%, a magnitude considerably larger than the tolerance

suggested by the reported measurements. It was thus concluded in [37] that the Howell system fails to map the positive light cone $\Omega^{(\text{pol})}$ into itself, and this is possibly the first time that a verdict of this kind was made on some published Mueller matrices.

Subsequent progress in respect of this issue was quite rapid. In a significant step forward Givens and Kostinski [39] derived, based on an impressive analysis of the spectrum of GM^TGM, what appeared to be a necessary and sufficient condition for M to map $\Omega^{(\text{pol})}$ into itself. They analyzed the Howell system based on their own condition, and concluded that their results were "in coincidence with the negative verdict on the Howell matrix delivered in [37]," p. 480. Soon after, Van der Mee [33] derived a more complete set of necessary and sufficient conditions for M to map $\Omega^{(\text{pol})}$ into itself; the analysis of Van der Mee, too, was based on the spectrum of GM^TGM.

Decomposition of a Mueller matrix M in various product forms, to gain insight into the physical effects M could have on the input polarization state, has been an activity of considerable interest [40–43]. The importance of obtaining the canonical or normal forms of Mueller matrices under the double-coset transformation $M \rightarrow L_\ell M L_r$, $L_\ell, L_r \in SO(3,1)$ was motivated in [24], and it was shown that the theorem of Givens and Kostinski [39] implied that the canonical form of every nonsingular (real) matrix M that maps $\Omega^{(\text{pol})}$ into itself is diagonal; i.e., $M = L_\ell M^{(1)} L_r$, where $M^{(1)} = \text{diag}(d_0, d_1, d_2, d_3)$, $d_0 \geqslant d_1 \geqslant d_2 \geqslant |d_3|$ and $L_\ell, L_r \in SO(3,1)$. It turned out that while the result of Van der Mee is essentially complete [44], that of Givens and Kostinski is incomplete. This means that the diagonal form $M^{(1)}$ of [24] noted above is not the only canonical form for a nonsingular M mapping $\Omega^{(\text{pol})}$ into itself; there exists another non-diagonal canonical form $M^{(2)}$, and this is essentially the case that was missed by the "theorem" of Givens and Kostinski quoted above. In a remarkably impressive and detailed study, Rao *et al.* [44,45] have further explored and completed the analysis of Van der Mee, leading to a complete solution to the question of canonical form for Mueller matrices under double-coseting by $SO(3,1)$ elements, raised in [24].

Since these canonical forms play a key role in our analysis below, we list them here in a concise form. Matrices M that map the state space $\Omega^{(\text{pol})}$ into itself divide into two major and two minor families:

Type I: $M = L_\ell M^{(1)} L_r, \quad L_\ell, L_r \in SO(3,1),$

$$M^{(1)} = \text{diag}(d_0, d_1, d_2, d_3), \qquad d_0 \geqslant d_1 \geqslant d_2 \geqslant |d_3|;$$

Type II: $M = L_\ell M^{(2)} L_r, \quad L_\ell, L_r \in SO(3,1),$

$$M^{(2)} = \begin{bmatrix} d_0 & d_0 - d_1 & 0 & 0 \\ 0 & d_1 & 0 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix}, \qquad d_0 > d_1$$

$$> 0, \quad \sqrt{d_0 d_1} \geqslant d_2 \geqslant |d_3|;$$

Polarizer: $M = L_\ell M^{(\text{pol})} L_r, \quad L_\ell, L_r \in SO(3,1),$

$$M^{(\text{pol})} = \begin{bmatrix} d_0 & d_0 & 0 & 0 \\ d_0 & d_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad d_0 > 0;$$

Pin Map: $M = L_\ell M^{(\text{pin})} L_r, \quad L_\ell, L_r \in SO(3,1),$

$$M^{(\text{pin})} = \begin{bmatrix} d_0 & 0 & 0 & 0 \\ d_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad d_0 > 0. \qquad (2.25)$$

Since elements of $SO(3,1)$ have unit determinant, it follows that d_3 in the Type-I and Type-II cases is positive, negative, or zero according as det M is positive, negative, or zero. The M matrices in the third and fourth families are manifestly singular. The third family is a Jones system, the associated H matrix being a projection; indeed, $M^{(\text{pol})}$ corresponds to a Jones matrix J whose only nonvanishing element is $J_{11} = \sqrt{2d_0}$. Finally, the PinMap family is so named because $M^{(\text{pin})}$ produces a *fixed* output polarization state independent of the input, the intensity of the output being independent of the state of polarization of the input. This may be contrasted with $M^{(\text{pol})}$: while the output in the case of $M^{(\text{pol})}$ has an input-independent state of polarization, the intensity does depend on the state of polarization of the input.

The matrix $M^{(\text{pin})}$ is not a Jones system, but it is a convex sum of such systems. To see this, note that a perfect depolarizer represented by the Mueller matrix

$$M^{(\text{depol})} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (2.26)$$

is a convex sum of Jones systems; it can be realized, for instance, as an equal mixture of systems with Jones matrices $\tau_a, a = 0,1,2,3$. That $M^{(\text{pin})}$ is a convex sum of Jones systems follows from $M^{(\text{pin})} = M^{(\text{pol})} M^{(\text{depol})}$. Alternatively, it is readily seen that $H_{M^{(\text{pin})}}$ is an equal sum of two projections, and hence $M^{(\text{pin})}$ is an equal mixture of the Jones systems:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \qquad (2.27)$$

While $M^{(\text{pin})}$ is realized as convex sum of two Jones systems, $M^{(\text{depol})}$ cannot be so realized with less than four Jones matrices. This follows from the fact that $H_{M^{(\text{depol})}}$ is of full rank whereas $H_{M^{(\text{pin})}}$ is of rank two.

The classification of canonical forms for M matrices as given in Eq. (2.25) is complete in the following sense:

Proposition 3: Every M matrix that maps the state space $\Omega^{(\text{pol})}$ into itself falls uniquely in one of the four families described in Eq. (2.25).

That brings us to the main thesis of the present paper. A 4×4 real matrix M will have to map the state space $\Omega^{(\text{pol})}$ into itself for it to qualify to be the Mueller matrix of some physical system. This is certainly a necessary condition. And, within the conventional Mueller–Stokes formalism, no conceivable further requirement can be imposed on M. But the action of the transversely homogeneous system represented by the numerical matrix M can be extended from plane waves to paraxial beams; naturally, M will then affect only the polarization degree of freedom and act as identity on the (transverse) spatial degrees of freedom. If M indeed represented a physical system, even this extended action should map physical states into physical states. It turns out that this trivial-looking extension is not all that trivial: there are M matrices that appear physical at the level of the (restricted) state space $\Omega^{(\text{pol})}$ but fail to be physical on the extended state space. Our task in the rest of the paper is to identify precisely those M matrices whose action is physical even on the extended state space. Since only those M matrices that pass this further hurdle can be called physical Mueller matrices, and pending determination of the precise demand this hurdle places on M, the M matrices that map $\Omega^{(\text{pol})}$ into itself will be called pre-Mueller matrices. We may thus conclude this section by saying that Eq. (2.25) gives a complete classification of pre-Mueller matrices and their orbit structure under double-coseting by elements of $SO(3,1)$; the physical/nonphysical divide of pre-Mueller matrices remains to be accomplished. This divide will be presented in Section 4 after some further preparation in Section 3.

3. FROM PLANE WAVES TO BEAMS: THE BEAM CORRELATION MATRIX

We will now go beyond plane waves and consider paraxial electromagnetic beams. The simplest (quasi-) monochromatic beam field has, in a transverse plane $z = \text{constant}$ described by coordinates $(x,y) \equiv \boldsymbol{\rho}$, the form $\mathbf{E}(\boldsymbol{\rho}) = (E_1 \hat{\mathbf{x}} + E_2 \hat{\mathbf{y}}) \psi(\boldsymbol{\rho})$, where E_1, E_2 are complex constants, and the scalar-valued function $\psi(\boldsymbol{\rho})$ may be assumed to be square integrable over the transverse plane: $\psi(\boldsymbol{\rho}) \in L^2(\mathbf{R}^2)$. It is clear that the polarization part $(E_1 \hat{\mathbf{x}} + E_2 \hat{\mathbf{y}})$ and the spatial dependence or modulation part $\psi(\boldsymbol{\rho})$ of such a beam are well separated, allowing one to focus attention on one aspect at a time. When one is interested in only the modulation aspect, the part $(E_1 \hat{\mathbf{x}} + E_2 \hat{\mathbf{y}})$ may be suppressed, thus leading to "scalar optics:" this is the domain of traditional Fourier optics [46]. [Fourier optics for electromagnetic beams requires a more delicate formalism [47].] On the other hand, if the spatial part $\psi(\boldsymbol{\rho})$ is suppressed, we are led to the traditional polarization optics (of plane waves) described in Section 2.

Beams whose polarization and spatial modulation separate in the above manner will be called elementary beams. It is clear that elementary beams remain elementary under the action of transversely homogeneous anisotropic systems such as waveplates and polarizers. That they remain elementary under the action of isotropic or polarization-insensitive modulating systems such as free propagation, phase screens, and lenses is also clear.

Now suppose that we superpose or add two such elementary beam fields $(a\hat{\mathbf{x}}+b\hat{\mathbf{y}})\psi(\boldsymbol{\rho})$ and $(c\hat{\mathbf{x}}+d\hat{\mathbf{y}})\chi(\boldsymbol{\rho})$. The result is not of the elementary form $(e\hat{\mathbf{x}}+f\hat{\mathbf{y}})\phi(\boldsymbol{\rho})$, for any e, f, $\phi(\boldsymbol{\rho})$ unless either (a,b) is proportional to (c,d) so that one gets committed to a *common* polarization, or $\psi(\boldsymbol{\rho})$ and $\chi(\boldsymbol{\rho})$ are proportional so that one gets committed to a *fixed* spatial mode. In other words, the set of elementary fields is not closed under superposition.

Since one cannot possibly give up the superposition principle in optics, one needs to go beyond the set of elementary fields and pay attention to the consequences of inseparability of polarization and spatial variation (modulation). We are thus led to consider (in a transverse plane) beam fields of the more general form $\mathbf{E}(\boldsymbol{\rho})$ $=E_1(\boldsymbol{\rho})\hat{\mathbf{x}}+E_2(\boldsymbol{\rho})\hat{\mathbf{y}}$.

This form is obviously closed under superposition. We may write $\mathbf{E}(\boldsymbol{\rho})$ as a generalized Jones vector:

$$\mathbf{E}(\boldsymbol{\rho}) = E_1(\boldsymbol{\rho})\hat{\mathbf{x}} + E_2(\boldsymbol{\rho})\hat{\mathbf{y}} \Leftrightarrow \mathbf{E}(\boldsymbol{\rho}) = \begin{bmatrix} E_1(\boldsymbol{\rho}) \\ E_2(\boldsymbol{\rho}) \end{bmatrix},$$

$$E_1(\boldsymbol{\rho}), \quad E_2(\boldsymbol{\rho}) \in L^2(\mathbf{R}^2). \qquad (3.1)$$

The intensity at location $\boldsymbol{\rho}$ corresponds to $|E_1(\boldsymbol{\rho})|^2$ $+|E_2(\boldsymbol{\rho})|^2$. This field is of the elementary or separable form iff $E_1(\boldsymbol{\rho})$ and $E_2(\boldsymbol{\rho})$ are linearly dependent (proportional to one another). Otherwise, polarization and spatial modulation are inseparable.

The point is that the set of possible beam fields in a transverse plane constitutes the *tensor product* space $\mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$, whereas the set of all elementary fields constitutes just the *set product* $\mathcal{C}^2 \times L^2(\mathbf{R}^2)$ of $\mathcal{C}^2$ and $L^2(\mathbf{R}^2)$. Recall that the tensor product of two vector spaces is the closure, under superposition, of their set product. Thus the set product $\mathcal{C}^2 \times L^2(\mathbf{R}^2)$ forms a measure zero subset of the tensor product $\mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$. In other words, in a beam field represented by a generic element of $\mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$, polarization and spatial modulation should be expected to be inseparable: Thus inseparability is not an exception; it is the rule in $\mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$, the space of pure states appropriate for electromagnetic beams.

If a beam described by generalized Jones vector $E(\boldsymbol{\rho})$ with x,y components $E_1(\boldsymbol{\rho})$, $E_2(\boldsymbol{\rho})$ is passed through an x polarizer, it is not only that the output will be x polarized, it is certain to be in the spatial mode $E_1(\boldsymbol{\rho})$ as well. A similar conclusion holds if the beam is passed through a y polarizer. Thus a (transversely homogeneous) polarizer, whose action is $\boldsymbol{\rho}$ independent, not only chooses a polarization state but acts as a spatial mode selector as well. This is true even if $E_1(\boldsymbol{\rho})$ and $E_2(\boldsymbol{\rho})$ are not spatially orthogonal modes. In a similar manner, a spatial mode selector insensitive to polarization will end up acting also as a polarization discriminator. This is a consequence of inseparability between polarization and spatial modulation and can be seen as one classical analog of quantum entanglement.

Now, to handle fluctuating beams, we pass on to the beam correlation (BC) matrix $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}') \equiv \langle \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^{\dagger} \rangle$, defined as the ensemble average of an outer product of (generalized) Jones vectors [4–7]. Such a matrix describes both the coherence and the polarization properties of the beam under consideration. It is a generalization of the nu-

merical coherency matrix of plane waves considered in the previous section, Eq. (2.2), now to the case of beam fields. It can equally well be viewed as a generalization of the mutual coherence function of scalar statistical optics to include polarization. For our present purpose, there is no need to make any finer distinction between the space–time and space–frequency descriptions. The two-point functions appearing in the BC matrix may be viewed either as equal-time coherence functions (in which case one speaks of the beam coherence-polarization matrix [5,6]) or correlation functions at a particular frequency (cross-spectral-density matrix [4,7]). We are free to view the BC matrix either as the 2×2 matrix of two-point functions,

$$\Phi(\boldsymbol{\rho};\boldsymbol{\rho}') = \begin{bmatrix} \langle E_1(\boldsymbol{\rho})E_1(\boldsymbol{\rho}')^* \rangle & \langle E_1(\boldsymbol{\rho})E_2(\boldsymbol{\rho}')^* \rangle \\ \langle E_2(\boldsymbol{\rho})E_1(\boldsymbol{\rho}')^* \rangle & \langle E_2(\boldsymbol{\rho})E_2(\boldsymbol{\rho}')^* \rangle \end{bmatrix}, \qquad (3.2)$$

or as the associated column vector $\widetilde{\Phi}(\boldsymbol{\rho};\boldsymbol{\rho}')$ of two-point functions: $\widetilde{\Phi}(\boldsymbol{\rho};\boldsymbol{\rho}') = \langle \mathbf{E}(\boldsymbol{\rho}) \otimes \mathbf{E}(\boldsymbol{\rho}')^* \rangle$.

It is clear from the very definition (3.2) of the BC matrix that this matrix kernel, viewed as an operator from $\mathcal{C}^2 \otimes L^2(\mathbf{R}^2) \to \mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$, is Hermitian and positive semidefinite:

$$\Phi_{jk}(\boldsymbol{\rho};\boldsymbol{\rho}') = \Phi_{kj}(\boldsymbol{\rho}';\boldsymbol{\rho})^*, \qquad j,k = 1,2;$$

$$\int d^2\boldsymbol{\rho} d^2\boldsymbol{\rho}' E(\boldsymbol{\rho})^{\dagger} \Phi(\boldsymbol{\rho};\boldsymbol{\rho}') \mathbf{E}(\boldsymbol{\rho}') \geqslant 0,$$

i.e., $\quad \displaystyle\sum_{jk} \int d^2\boldsymbol{\rho} d^2\boldsymbol{\rho}' E_j(\boldsymbol{\rho})^* \Phi_{jk}(\boldsymbol{\rho};\boldsymbol{\rho}') E_k(\boldsymbol{\rho}') \geqslant 0,$

$$\forall \mathbf{E}(\boldsymbol{\rho}) \in \mathcal{C}^2 \otimes L^2(\mathbf{R}^2). \qquad (3.3)$$

The positivity requirement thus demands that the expectation value of $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')$ be nonnegative for every Jones vector $\mathbf{E}(\boldsymbol{\rho})$. Hermiticity and positivity are the defining properties of the BC matrix: every 2×2 matrix of two-point functions $\Phi_{jk}(\boldsymbol{\rho};\boldsymbol{\rho}')$ meeting just these two conditions is a valid BC matrix of some beam of light.

We can use the BC matrix to define the generalized Stokes vector $S(\boldsymbol{\rho};\boldsymbol{\rho}')$ [14,15]:

$$\Phi(\boldsymbol{\rho};\boldsymbol{\rho}') = \frac{1}{2}\sum_{a=0}^{3} S_a(\boldsymbol{\rho};\boldsymbol{\rho}')\tau_a \Leftrightarrow S_a(\boldsymbol{\rho};\boldsymbol{\rho}') = \mathrm{tr}(\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')\tau_a).$$

$$(3.4)$$

That this is an invertible relation shows that $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')$ and $S(\boldsymbol{\rho};\boldsymbol{\rho}')$ carry identical information: action of an optical system on one defines a unique equivalent action on the other. The Hermiticity and positivity requirement on the BC matrix can be easily transcribed into corresponding requirements on $S(\boldsymbol{\rho};\boldsymbol{\rho}')$. Hermiticity reads

$$S_a(\boldsymbol{\rho};\boldsymbol{\rho}') = S_a(\boldsymbol{\rho}';\boldsymbol{\rho})^*, \qquad a = 0,1,2,3, \qquad (3.5)$$

whereas positivity reads

$$\sum_{a=0}^{3} g_a \int d^2\boldsymbol{\rho} d^2\boldsymbol{\rho}' S_a(\boldsymbol{\rho};\boldsymbol{\rho}')\hat{S}_a(\boldsymbol{\rho}';\boldsymbol{\rho}) \geqslant 0, \qquad (3.6)$$

for every Stokes vector $\hat{S}(\boldsymbol{\rho};\boldsymbol{\rho}')$ arising from Jones vectors of the form $\mathbf{E}(\boldsymbol{\rho}) \in \mathcal{C}^2 \otimes L^2(\mathbf{R}^2)$. The signatures g_a correspond to the Lorentz metric: $g_0 = 1$, $g_a = -1$ for $a \neq 0$.

4. FROM PRE-MUELLER MATRICES TO MUELLER MATRICES: THE ROLE OF INSEPARABILITY

We now have at our disposal all the tools necessary to determine whether or not a given pre-Mueller matrix is a physical Mueller matrix. Let us consider the transformation of the generalized Stokes vector $S(\boldsymbol{\rho};\boldsymbol{\rho}')$ and the associated BC matrix $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')$ under the action of a transversely homogeneous optical system described by numerical pre-Mueller matrix M. We begin our analysis with pre-Mueller matrices of Type I.

A. Type-I pre-Mueller Matrices
We will first study pre-Mueller matrices presented in the canonical form $M^{(1)} = \mathrm{diag}(d_0, d_1, d_2, d_3)$. Extension of the conclusions to pre-Mueller matrices not in the canonical form will turn out to be quite straightforward. In view of the system's homogeneity, the action of $M^{(1)}$ is necessarily independent of $\boldsymbol{\rho}, \boldsymbol{\rho}'$, and we have

$$M^{(1)} = \mathrm{diag}(d_0, d_1, d_2, d_3) : \begin{bmatrix} S_0(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_1(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_2(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_3(\boldsymbol{\rho};\boldsymbol{\rho}') \end{bmatrix} \rightarrow \begin{bmatrix} S_0'(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_1'(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_2'(\boldsymbol{\rho};\boldsymbol{\rho}') \\ S_3'(\boldsymbol{\rho};\boldsymbol{\rho}') \end{bmatrix}$$

$$= \begin{bmatrix} d_0 S_0(\boldsymbol{\rho};\boldsymbol{\rho}') \\ d_1 S_1(\boldsymbol{\rho};\boldsymbol{\rho}') \\ d_2 S_2(\boldsymbol{\rho};\boldsymbol{\rho}') \\ d_3 S_3(\boldsymbol{\rho};\boldsymbol{\rho}') \end{bmatrix}. \qquad (4.1)$$

The elements of the output BC matrix $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ associated with the output Stokes vector $S'(\boldsymbol{\rho};\boldsymbol{\rho}')$ resulting from the action of $M^{(1)}$ on $S(\boldsymbol{\rho};\boldsymbol{\rho}')$, are easily computed using Eq. (3.4):

$$\Phi_{11}'(\boldsymbol{\rho};\boldsymbol{\rho}') = [(d_0 + d_1)\Phi_{11}(\boldsymbol{\rho};\boldsymbol{\rho}') + (d_0 - d_1)\Phi_{22}(\boldsymbol{\rho};\boldsymbol{\rho}')]/2,$$

$$\Phi_{22}'(\boldsymbol{\rho};\boldsymbol{\rho}') = [(d_0 + d_1)\Phi_{22}(\boldsymbol{\rho};\boldsymbol{\rho}') + (d_0 - d_1)\Phi_{11}(\boldsymbol{\rho};\boldsymbol{\rho}')]/2,$$

$$\Phi_{12}'(\boldsymbol{\rho};\boldsymbol{\rho}') = [(d_2 + d_3)\Phi_{12}(\boldsymbol{\rho};\boldsymbol{\rho}') + (d_2 - d_3)\Phi_{21}(\boldsymbol{\rho};\boldsymbol{\rho}')]/2,$$

$$\Phi_{21}'(\boldsymbol{\rho};\boldsymbol{\rho}') = [(d_2 + d_3)\Phi_{21}(\boldsymbol{\rho};\boldsymbol{\rho}') + (d_2 - d_3)\Phi_{12}(\boldsymbol{\rho};\boldsymbol{\rho}')]/2. \qquad (4.2)$$

Clearly, a necessary condition for the pre-Mueller matrix $M^{(1)} = \mathrm{diag}(d_0, d_1, d_2, d_3)$ to be a physical Mueller matrix is that the output $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ in Eq. (4.2) be a valid BC matrix, for every valid input BC matrix $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')$. Hermiticity of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ is manifest in view of that of $\Phi(\boldsymbol{\rho};\boldsymbol{\rho}')$ and reality of the parameters d_a. Thus what remains to be checked is the positivity of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$. While testing the positivity of a generic matrix kernel could be a formidable

task in general, it turns out that this test can be carried out fairly easily in the present case.

Let us take as the input a special pure-state BC matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^\dagger$, corresponding to the generalized Jones vector $\mathbf{E}(\boldsymbol{\rho})$ that is an equal superposition of an x-polarized mode and a y-polarized mode, the two modes being spatially orthogonal:

$$\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^\dagger,$$

$$\mathbf{E}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_1(\boldsymbol{\rho}) \\ \psi_2(\boldsymbol{\rho}) \end{bmatrix}, \qquad \int \psi_j(\boldsymbol{\rho})\psi_k(\boldsymbol{\rho})^* d^2\boldsymbol{\rho} = \delta_{jk}. \qquad (4.3)$$

[$\psi_1(\boldsymbol{\rho})$ and $\psi_2(\boldsymbol{\rho})$ could, for instance, be two distinct Hermite–Gaussian modes.] This means that the entries of the input BC matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ have the deterministic form $\Phi_{jk}^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \psi_j(\boldsymbol{\rho})\psi_k(\boldsymbol{\rho}')^*$. A consequence of this simple (product) form is that the four entries of the BC matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ form an orthonormal set of (two-point) functions:

$$\int d^2\boldsymbol{\rho} d^2\boldsymbol{\rho}' \Phi_{ij}^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')\Phi_{kl}^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')^* = \delta_{ik}\delta_{jl}. \qquad (4.4)$$

This fact will prove to be of much value in our analysis below.

To test positivity of the output BC matrix $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$, given in Eq. (4.2) and resulting from input $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}') = \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^\dagger$, let us define four (generalized) Jones vectors:

$$\mathbf{E}^{(\pm)}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_1(\boldsymbol{\rho}) \\ \pm\psi_2(\boldsymbol{\rho}) \end{bmatrix}, \qquad \mathbf{F}^{(\pm)}(\boldsymbol{\rho}) = \begin{bmatrix} \psi_2(\boldsymbol{\rho}) \\ \pm\psi_1(\boldsymbol{\rho}) \end{bmatrix}. \qquad (4.5)$$

[The input Jones vector $\mathbf{E}(\boldsymbol{\rho})$ happens to coincide with $\mathbf{E}^{(+)}(\boldsymbol{\rho})$.] Expectation values of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ for the four Jones vectors $\mathbf{E}^{(\pm)}(\boldsymbol{\rho}), \mathbf{F}^{(\pm)}(\boldsymbol{\rho})$ are easily computed using Eqs. (4.2)–(4.5) in Eq. (3.3). These expectation values are $(d_0 + d_1) \pm (d_2 + d_3)$ for $\mathbf{E}^{(\pm)}(\boldsymbol{\rho})$ and $(d_0 - d_1) \pm (d_2 - d_3)$ for $\mathbf{F}^{(\pm)} \times (\boldsymbol{\rho})$.

Now positivity of $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ requires, as a necessary condition, that these four expectation values be nonnegative, and this requirement places on the parameters d_a the constraints

$$-d_1 - d_2 - d_3 \leqslant d_0,$$

$$-d_1 + d_2 + d_3 \leqslant d_0;$$

$$d_1 + d_2 - d_3 \leqslant d_0,$$

$$d_1 - d_2 + d_3 \leqslant d_0. \qquad (4.6)$$

Violation of any one of these four conditions will render the output $\Phi'(\boldsymbol{\rho};\boldsymbol{\rho}')$ nonphysical as BC matrix. Since the input BC matrix $\Phi^{(0)}(\boldsymbol{\rho};\boldsymbol{\rho}')$ is obviously physical, this will in turn render $M^{(1)}$ nonphysical as Mueller matrix: Eq. (4.6) is thus a set of necessary conditions for the pre-Mueller matrix $M^{(1)}$ to be a Mueller matrix.

Suppose these four inequalities are met. Can we then conclude that the pre-Mueller matrix $M^{(1)}$ is a physically acceptable Mueller matrix? To answer this question in the

affirmative, we write in detail the associated Hermitian matrix $H_{M^{(1)}} = \frac{1}{2}\Sigma_a d_a \tau_a \otimes \tau_a^* = \Sigma_a d_a U_{aa}$:

$$H^{(M)} = \frac{1}{2}\begin{bmatrix} d_0+d_1 & 0 & 0 & d_2+d_3 \\ 0 & d_0-d_1 & d_2-d_3 & 0 \\ 0 & d_2-d_3 & d_0-d_1 & 0 \\ d_2+d_3 & 0 & 0 & d_0+d_1 \end{bmatrix}. \quad (4.7)$$

Validity of the four inequalities in Eq. (4.6) implies, fortunately, that this matrix is positive semidefinite. This in turn implies that the given diagonal system $M^{(1)}$ is a convex sum of Jones systems and therefore takes every BC matrix into a BC matrix, showing that Eq. (4.6) is a sufficient condition for $M^{(1)}$ to be a Mueller matrix. We have thus proved

Proposition 4: The pre-Mueller matrix $M^{(1)} = \mathrm{diag}(d_0, d_1, d_2, d_3)$ is a Mueller matrix iff the associated Hermitian matrix $H_{M^{(1)}} \geq 0$, that is, iff $M^{(1)}$ can be realized as a convex sum of Jones systems, or, equivalently, iff the entries of $M^{(1)}$ respect the inequalities in Eq. (4.6).

Having settled the diagonal case, we now go beyond and consider the more general Type-I pre-Mueller matrices. As noted in Eq. (2.25), these are necessarily of the general form $M = L_\ell M^{(1)} L_r$, where $L_\ell, L_r \in SO(3,1)$ and

$M^{(1)}$ is diagonal. We have already noted that L_ℓ, L_r are physical Mueller matrices: indeed, they correspond to deterministic systems with respective Jones matrices $J_\ell, J_r \in SL(2,C)$. Thus if $M^{(1)}$ is a Mueller matrix, then it has a convex sum realization $\{p_k, J^{(k)}\}$. This implies that $M = L_\ell M^{(1)} L_r$ has the convex sum realization $\{p_k, J_\ell J^{(k)} J_r\}$ and hence is a valid Mueller matrix. The converse follows by virtue of the invertibility of J_ℓ, J_r, and we have

Proposition 5: A Type-I pre-Mueller matrix, which is necessarily of the form $M = L_\ell M^{(1)} L_r$ with $L_\ell, L_r \in SO(3,1)$ and $M^{(1)}$ diagonal, is a physical Mueller matrix iff $M^{(1)}$ is.

B. Type-II pre-Mueller Matrices

Having fully classified Type-I pre-Mueller matrices into Mueller and non-Mueller matrices, we now turn our attention to Type-II pre-Mueller matrices. The analysis turns out to be quite parallel to the one in Subsection 4.A

Recall from Section 2 that a Type-II pre-Mueller matrix in its canonical form $M^{(2)}$ has only one nonvanishing off-diagonal element whose value is fixed by the diagonals, namely, $M_{01} = d_0 - d_1$, where d_0, d_1, d_2, d_3 are the diagonals. The action of $M^{(2)}$ on $S(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$ can be computed as before. The (generalized) Stokes vector has this simple transformation law:

$$M^{(2)} : S(\boldsymbol{\rho}; \boldsymbol{\rho}') \to S'(\boldsymbol{\rho}; \boldsymbol{\rho}') = \begin{bmatrix} S_0'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_1'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_2'(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ S_3'(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix} = \begin{bmatrix} d_0 S_0(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0-d_1) S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_1 S_1(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_2 S_2(\boldsymbol{\rho}; \boldsymbol{\rho}') \\ d_3 S_3(\boldsymbol{\rho}; \boldsymbol{\rho}') \end{bmatrix}. \quad (4.8)$$

The elements of the output BC matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ associated with $S'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ and computed from Eq. (3.4) are

$$\Phi_{11}'(\boldsymbol{\rho}; \boldsymbol{\rho}') = d_0 \Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}'),$$

$$\Phi_{22}'(\boldsymbol{\rho}; \boldsymbol{\rho}') = d_1 \Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_0-d_1)\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}'),$$

$$\Phi_{12}'(\boldsymbol{\rho}; \boldsymbol{\rho}') = [(d_2+d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2-d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}')]/2,$$

$$\Phi_{21}'(\boldsymbol{\rho}; \boldsymbol{\rho}') = [(d_2+d_3)\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}') + (d_2-d_3)\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}')]/2.$$

$$(4.9)$$

As in the case of $M^{(1)}$, the canonical form pre-Mueller matrix $M^{(2)}$ does not couple the pair $\Phi_{11}(\boldsymbol{\rho}; \boldsymbol{\rho}')$, $\Phi_{22}(\boldsymbol{\rho}; \boldsymbol{\rho}')$ with $\Phi_{12}(\boldsymbol{\rho}; \boldsymbol{\rho}')$, $\Phi_{21}(\boldsymbol{\rho}; \boldsymbol{\rho}')$.

Again, a necessary condition for the pre-Mueller matrix $M^{(2)}$ to be a physically acceptable Mueller matrix is that the output $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ in Eq. (4.9) be a valid BC matrix for every valid input BC matrix $\Phi(\boldsymbol{\rho}; \boldsymbol{\rho}')$. As in the case of $M^{(1)}$, let us take as input the pure-state BC matrix $\Phi^{(0)} \times (\boldsymbol{\rho}; \boldsymbol{\rho}') = \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^\dagger$, with $\mathbf{E}(\boldsymbol{\rho})$ as described in Eq. (4.3). To test positivity of the output BC matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$, given in Eq. (4.9) and resulting from input $\Phi^{(0)}(\boldsymbol{\rho}; \boldsymbol{\rho}')$

$= \mathbf{E}(\boldsymbol{\rho})\mathbf{E}(\boldsymbol{\rho}')^\dagger$, we use in place of $\mathbf{E}^{(\pm)}(\boldsymbol{\rho})$ and $\mathbf{F}^{(\pm)}(\boldsymbol{\rho})$ slightly modified (generalized) Jones vectors $\mathbf{E}^{(\theta)}(\boldsymbol{\rho})$ and $\mathbf{F}^{(\theta)}(\boldsymbol{\rho})$:

$$\mathbf{E}^{(\theta)}(\boldsymbol{\rho}) = \begin{bmatrix} \cos\theta\psi_1(\boldsymbol{\rho}) \\ \sin\theta\psi_2(\boldsymbol{\rho}) \end{bmatrix}, \qquad \mathbf{F}^{(\theta)}(\boldsymbol{\rho}) = \begin{bmatrix} \cos\theta\psi_2(\boldsymbol{\rho}) \\ \sin\theta\psi_1(\boldsymbol{\rho}) \end{bmatrix}.$$

$$(4.10)$$

Expectation values of the output BC matrix $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ for these two families of Jones vectors can be computed as before. These expectation values are $d_0\cos^2\theta + d_1\sin^2\theta + (d_2+d_3)\cos\theta\sin\theta$ for $\mathbf{E}^\theta(\boldsymbol{\rho})$ and $(d_0-d_1)\sin^2\theta + (d_2-d_3)\cos\theta\sin\theta$ for $\mathbf{F}^\theta(\boldsymbol{\rho})$.

Now positivity of $\Phi'(\boldsymbol{\rho}; \boldsymbol{\rho}')$ requires, as a necessary condition, that these expectation values be nonnegative for all $0 \leq \theta < \pi$, and this requirement is seen to be equivalent to the pair of conditions $d_0 d_1 \geq (d_2+d_3)^2/4$, $d_2-d_3 = 0$; these arise, respectively, from the $\mathbf{E}^{(\theta)}$ and $\mathbf{F}^{(\theta)}$ families. We may rewrite these as

$$d_3 = d_2, \qquad (d_2)^2 \leq d_0 d_1. \quad (4.11)$$

This is a pair of necessary conditions for the pre-Mueller matrix $M^{(2)}$ to be a Mueller matrix. The condition $(d_2)^2 \leq d_0 d_1$ is already part of the definition of $M^{(2)}$, and thus $d_3 = d_2$ is the new requirement arising from consideration

of the action of $M^{(2)}$ on BC matrices, i.e., from consideration of inseparability.

Our next task is to show that these conditions are sufficient as well. To this end we proceed as in the case of $M^{(1)}$ and compute the hermitian matrix $H_{M^{(2)}}$ associated with $M^{(2)}$:

$$H_{M^{(2)}} = \begin{bmatrix} d_0 & 0 & 0 & \frac{1}{2}(d_2+d_3) \\ 0 & 0 & \frac{1}{2}(d_2-d_3) & 0 \\ 0 & \frac{1}{2}(d_2-d_3) & d_0-d_1 & 0 \\ \frac{1}{2}(d_2+d_3) & 0 & 0 & d_1 \end{bmatrix}. \tag{4.12}$$

The inequalities in Eq. (4.11) are precisely the conditions under which $H_{M^{(2)}}$ is positive semidefinite. This in turn implies that $M^{(2)}$ satisfying Eq. (4.11) is a convex sum of Jones systems and hence is a physical Mueller matrix. We have thus proved

Proposition 6: The pre-Mueller matrix $M^{(2)}$ is a Mueller matrix iff the associated Hermitian matrix $H_{M^{(2)}} \geqslant 0$, that is, iff $M^{(2)}$ can be realized as a convex sum of Jones systems or, equivalently, iff the entries of $M^{(2)}$ respect the inequalities in Eq. (4.11).

We can now proceed to consider Type-II pre-Mueller matrices that are not of the canonical form $M^{(2)}$. We know from Section 2 that any such matrix has the form $M = L_\ell M^{(2)} L_r$, where $L_\ell, L_r \in SO(3,1)$. By considerations similar to the ones leading to Proposition 6 in the Type-I case, we arrive at

Proposition 7: A type-II pre-Mueller matrix, which is necessarily of the form $M = L_\ell M^{(2)} L_r$ with $L_\ell, L_r \in SO(3,1)$, is a Mueller matrix iff $M^{(2)}$ is.

Having completed classification of the pre-Mueller matrices in the Type-I and Type-II families into physical and nonphysical ones, we are now left with two minor families to handle. As noted following Eq. (2.25), $M^{(\text{pol})}$ is a Jones system. Let J be the Jones matrix representing this system (polarizer). In view of the two-to-one homomorphism between $SL(2,C)$ and $SO(3,1)$ alluded to earlier, $L_\ell, L_r \in SO(3,1)$ define respective Jones matrices J_ℓ, J_r of unit determinant; these Jones matrices are unique except for multiplicative factor ± 1, and, as is well known, this signature ambiguity is of nontrivial origin. Thus $M = L_\ell M^{(\text{pol})} L_r$ is a Jones system with Jones matrix $\pm J_\ell J J_r$ and hence is physical. A similar argument will show that the last family, namely the PinMap family, also has no nonphysical M matrix. For completeness, we state the situation in with respect to these two minor families as the following.

Proposition 8: Pre-Mueller matrices belonging to the polarizer and pin map families are, respectively, Jones systems and convex sums of Jones systems. Their associated H matrices are positive semidefinite, and all pre-Mueller matrices in these two families are physical Mueller matrices.

C. Complete Characterization of Mueller Matrices
In the last two subsections we carried out a complete classification of pre-Mueller matrices into physical and nonphysical ones. Double-coseting under the $SO(3,1)$ group has played such an important role in this process that we capture this role as a separate result.

Proposition 9: Given two 4×4 real matrices M and M' that are in the same double-coset orbit under $SO(3,1)$, i.e., $M' = L_\ell M L_r$ for some $L_\ell, L_r \in SO(3,1)$, M' is a convex sum of Jones systems iff M is. And $H_{M'} \geqslant 0$ iff $H_M \geqslant 0$. In other words, M' is a Mueller matrix iff M is.

Proof: Suppose M has the convex sum realization $\{p_k, J^{(k)}\}$, i.e., $M = \Sigma_k p_k M(J^{(k)})$. Then, clearly, M' has the convex sum realization $\{p_k, J_\ell J^{(k)} J_r\}$. Conversely, if M' has the convex sum realization $\{p_k', J'\}$, then M has the convex sum realization $\{p_k', (J_\ell)^{-1} J'^{(k)} (J_r)^{-1}\}$.

Now suppose $H^{(M)} \geqslant 0$. This means $H^{(M)} = \Sigma_k p_k \widetilde{J}^{(k)} \widetilde{J}^{(k)\dagger}, p_k > 0$. This immediately implies $H_{M'} = \Sigma_k p_k (J_\ell \widetilde{J}^{(k)} J_r)(J_\ell \widetilde{J}^{(k)} J_r)^\dagger$, which proves its positivity. Here $(J_\ell \widetilde{J}^{(k)} J_r)$, as usual, denotes the column vector associated with the 2×2 matrix $J_\ell J^{(k)} J_r$. The converse follows from the invertibility of J_ℓ, J_r, completing proof of the proposition.

With this result, proof of the principal conclusion of this paper is complete. Our main theorem may thus be stated as follows.

Main Theorem: A 4×4 real matrix M is a Mueller matrix iff the associated Hermitian matrix $H^{(M)} \geqslant 0$. Every physically acceptable Mueller matrix is a convex sum of Mueller–Jones matrices.

D. The Role of Inseparability: An Illustrative Example
We present a simple example to illustrate the kind of restrictions on M matrices brought in by consideration of inseparability. Let us restrict attention to M matrices of the special simple three-parameter form

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & d_1 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix}. \tag{4.13}$$

We are obviously in the Type-I situation, but we are not considering here the $SO(3,1)$ orbit under double-coseting.

It is clear that M will map Stokes vectors into Stokes vectors if and only if M satisfies the following three conditions:

$$M : \Omega^{(\text{pol})} \to \Omega^{(\text{pol})} \Leftrightarrow -1 \leqslant d_k \leqslant 1, \qquad k = 1,2,3. \tag{4.14}$$

In the absence of considerations of inseparability these would have been the only conditions M will need to satisfy. Thus the allowed region in the Euclidean space $\mathbf{R}^3$ spanned by the parameters (d_1, d_2, d_3) would have been a cube with vertices at $(\pm 1, \pm 1, \pm 1)$. Now each one of the four conditions in Eq. (4.6) with $d_0 = 1$, arising out of consideration of inseparability, forbids the (open) half-space on one side of a plane. Thus the allowed region is the intersection of the four allowed half-spaces. This region is
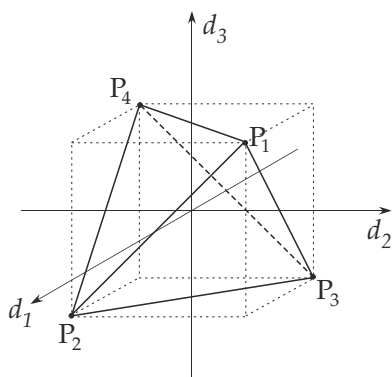
Fig. 1.   The tetrahedron that corresponds to Mueller matrices.

clearly the tetrahedron with vertices at $P_1=(1,1,1)$, $P_2=(1,-1,-1)$, $P_3=(-1,1,-1)$ and $P_4=(-1,-1,1)$.

Thus, it is not the entire cubic region (4.14) permitted by conventional wisdom, but only the solid tetrahedron, with one-third volume of the cube, that withstands the closer scrutiny presented by consideration of inseparability. This is illustrated in Fig. 1. The region outside the tetrahedron is nonphysical and does not correspond to Mueller matrices: points in the cube but exterior to the tetrahedron correspond to pre-Mueller matrices that are not physical Mueller matrices; those outside the cube correspond to M matrices that are not even pre-Mueller matrices.

5. CONCLUDING REMARKS

We conclude with some further observations. First, in the mathematics literature and in the literature of quantum information theory, what we have called pre-Mueller matrices go by the name "positive maps," and the subset of pre-Mueller matrices that prove to be physically acceptable in the sense of our main theorem corresponds to what are called "completely positive maps." But we have endeavored here to arrive at a physical characterization of Mueller matrices entirely within the framework of BC matrices familiar to the classical optics community, without resorting to the mathematical theory of these maps.

Second, a pre-Mueller matrix that fails our main theorem *will not* produce any nonphysical effect acting on the coherency matrix of plane waves or on the BC matrix of elementary (or polarization-modulation separable) beams. It follows that BC matrices that are convex sums of elementary beams will not be able to witness the failure of a pre-Mueller matrix M whose associated $H^{(M)}$ is not positive semidefinite. Only BC matrices that are not convex sums of elementary beams can expose the nonphysical nature of a pre-Mueller matrix that violates our main theorem. In other words, pre-Mueller matrices cannot be further divided into physical and nonphysical subsets without consideration of inseparability.

Further, ever since it was proved that every Jones system corresponds to a Mueller matrix whose associated H matrix is a projection [28], it has been clear that ensembles of Jones systems necessarily correspond to positive semidefinite H matrices, and conversely. It has thus been occasionally suggested by various authors, begin-

ning with [29], that considerations of Mueller matrices might be restricted only to such ensembles. But it has remained only a suggestion, and one without any physical basis, and hence could not set aside as nonphysical an experimentally measured Mueller matrix whose associated H matrix has a negative eigenvalue, particularly when the reported Mueller system was not realized by the experimenter specifically as an ensemble of Jones systems. For instance, the symmetric Mueller matrix of Van Zyl [48] was analyzed in [24] and was found to be a matrix of Type III, with canonical-form parameters $d_0=0.9735$, $d_1=0.9112$, $d_2=0.4640$, $d_3=-0.3838$. This clearly violates [the second constraint in] Eq. (4.6) to a substantial extent. Equivalently, the eigenvalues of $H^{(M)}$ are 1.0906, 0.8393, 0.4526, −0.3825.

That $H^{(M)}$ in this case is not positive, and hence the Van Zyl system is not a convex sum of Jones systems, was always known. However, one did not hitherto have a physical basis on which this M could be judged as nonphysical. Inseparability has now given us such a physical basis.

Finally, the issue that has been settled in the present work can be viewed as one of lifting the linear input–output relation $S'=MS$ at the level of Stokes vectors to an input–output relation at the level of field amplitudes. We have seen that a linear relation among Stokes vectors effected by a Mueller–Jones matrix can be lifted to the linear relation

$$E_1' = aE_1 + bE_2,$$

$$E_2' = cE_1 + dE_2, \tag{5.1}$$

at the field amplitude level. Here a, b, c, d are deterministic scalars. Indeed, these are precisely the entries of the Jones matrix of the system. It is also known [29] that for Mueller matrices whose associated Hermitian matrix $H^{(M)} \geq 0$, the relation $S'=MS$ can be lifted to a relation of the form (5.1), a, b, c, d being appropriately correlated random variables this time. Now there exist M matrices that map valid Stokes vectors into valid Stokes vectors but whose associated Hermitian matrix $H^{(M)}$ is indefinite, and in this last case understanding the relation $S'=MS$ at the field amplitude level remained a puzzle. The present work resolves this puzzle by simply showing that there are no Mueller matrices of this last type: Every physical Mueller matrix can be understood at the field amplitude level, either as a deterministic or as a stochastic linear input–output relation.

ACKNOWLEDGMENTS

REFERENCES

1.   L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, 1995), Chap. 6.
2.   C. Brosseau, *Fundamentals of Polarized Light: A Statistical Approach* (Wiley, 1998).

3.   J. J. Gil, "Polarimetric characterization of light and media," Eur. Phys. J.: Appl. Phys. **40**, 1–47 (2007).
4.   D. F. V. James, "Change of polarization of light-beams on propagation through free space," J. Opt. Soc. Am. A **11**, 1641–1643 (1994).
5.   F. Gori, "Matrix treatment of partially polarized, partially coherent beams," Opt. Lett. **23**, 241–243 (1998).
6.   F. Gori, M. Santarsiero, S. Vicalvi, R. Borghi, and G. Guattari, "Beam coherence-polarization matrix," J. Opt. A, Pure Appl. Opt. **7**, 941–951 (1998).
7.   E. Wolf, "Unified theory of coherence and polarization of random electromagnetic beams," Phys. Lett. A **312**, 263–267 (2003).
8.   F. Gori, M. Santarsiero, G. Piquero, R. Borghi, A. Mondello, and R. Simon, "Partially polarized Gaussian Schell-model beams," J. Opt. A, Pure Appl. Opt. **3**, 1–9 (2001).
9.   F. Gori, M. Santarsiero, R. Simon, G. Piquero, R. Borghi, and G. Guattari, "Coherent-mode decomposition of partially coherent partially polarized beams," J. Opt. Soc. Am. A **20**, 78–84 (2003).
10. J. Tervo, T. Setälä, and A. T. Friberg, "Degree of coherence of electromagnetic fields," Opt. Express **11**, 1137–1142 (2003).
11. T. Setälä, J. Tervo, and A. T. Friberg, "Complete electromagnetic coherence in the space frequency domain," Opt. Lett. **29**, 328–330 (2004).
12. M. R. Dennis, "A three-dimensional degree of polarization based on Rayleigh scattering," J. Opt. Soc. Am. A **24**, 2065–2069 (2007).
13. P. Refrégiér and F. Goudail, "Invariant degrees of coherence of partially polarized light," Opt. Express **13**, 1051–1060 (2005).
14. O. Korotkova and E. Wolf, "Generalized Stokes parameters of random electromagnetic beams," Opt. Lett. **30**, 198–200 (2005).
15. T. Setälä, J. Tervo, and A. T. Friberg, "Stokes parameters and polarization contrasts in Young's interference experiment," Opt. Lett. **31**, 2208–2210 (2006).
16. A. Aiello and J. P. Woerdman, "Role of spatial coherence in polarization tomography," Opt. Lett. **30**, 1599–1601 (2005).
17. F. Gori, M. Santarsiero, and R. Borghi, "Maximizing Young's fringe visibility through reversible optical transformations," Opt. Lett. **32**, 588–590 (2007).
18. A. Luis, "Degree of coherence for vectorial electromagnetic fields as a distance between correlation matrices," J. Opt. Soc. Am. A **24**, 1063–1068 (2007).
19. R. Martínez-Herrero and P. M. Mejías, "Maximum visibility under unitary transformations in two-pinhole interference for electromagnetic fields," Opt. Lett. **32**, 1471–1473 (2007).
20. R. Martínez-Herrero and P. M. Mejías, "Relation between degrees of coherence for electromagnetic fields," Opt. Lett. **32**, 1504–1506 (2007).
21. A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" Phys. Rev. **47**, 777–780 (1935).
22. N. Bohr, "Can quantum-mechanical description of physical reality be considered complete?" (Ed. note: commentary on [21]) Phys. Rev. **48**, 696–702 (1935).
23. R. Barakat, "Bilinear constraints between elements of the 4×4 Mueller-Jones transfer matrix of polarization theory," Opt. Commun. **38**, 159–161 (1981).
24. R. Sridhar and R. Simon, "Normal form for Mueller matrices in polarization optics," J. Mod. Opt. **41**, 1903–1915 (1994).
25. G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," Trans. Cambridge Philos. Soc. **9**, 399–416 (1852).
26. K. D. Abhyankar and A. L. Fymat, "Relationships between the elements of the phase matrix for scattering," J. Math. Phys. **10**, 1935–1938 (1969).
27. E. S. Fry and E. W. Kattawar, "Relationships between elements of the Stokes matrix," Appl. Opt. **20**, 2811–2814 (1981).
28. R. Simon, "The connection between Mueller and Jones matrices of polarization optics," Opt. Commun. **42**, 293–297 (1982).
29. K. Kim, L. Mandel, and E. Wolf, "Relationship between Jones and Mueller matrices for random media," J. Opt. Soc. Am. A **4**, 433–437 (1987).
30. J. J. Gil and E. Bernabeu, "A depolarization criterion in Mueller matrices," Opt. Acta **32**, 259–261 (1985).
31. J. J. Gil and E. Bernabeu, "Obtainment of the polarizing and retardation parameters of a nondepolarizing optical system from the decomposition of its Mueller matrix," Optik (Stuttgart) **76**, 67–71 (1987).
32. S. R. Cloude, "Group-theory and polarization algebra," Optik (Stuttgart) **75**, 26–36 (1986).
33. V. M. van der Mee, "An eigenvalue criterion for matrices transforming Stokes parameters," J. Math. Phys. **34**, 5072–5088 (1993).
34. D. G. M. Anderson and R. Barakat, "Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix," J. Opt. Soc. Am. A **11**, 2305–2319 (1994).
35. A. Aiello, G. Puentes, and J. P. Woerdman, "Linear optics and quantum maps," Phys. Rev. A **76**, 032323 (2007).
36. Sudha, A. V. G. Rao, A. R. U. Devi, and A. K. Rajagopal, "Positive-operator-valued measure view of the ensemble approach to polarization optics," J. Opt. Soc. Am. A **25**, 874–880 (2008).
37. M. Sanjay Kumar and R. Simon, "Characterization of Mueller matrices in polarization optics," Opt. Commun. **88**, 464–470 (1992).
38. B. J. Howell, "Measurement of the polarization effects of an instrument using partially polarized light," Appl. Opt. **18**, 809–812 (1979).
39. C. R. Givens and B. Kostinski, "A simple necessary and sufficient condition on physically realizable Mueller matrices," J. Mod. Opt. **40**, 471–481 (1993).
40. S. Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," J. Opt. Soc. Am. A **13**, 1106–1113 (1996).
41. R. Ossikovski, A. De Martino, and S. Guyot, "Forward and reverse product decompositions of depolarizing Mueller matrices," Opt. Lett. **32**, 689–691 (2007).
42. V. Devlaminck and P. Terrier, "Definition of a parameter form for non-singular Mueller matrices," J. Opt. Soc. Am. A **25**, 2636–2643 (2008).
43. F. Boulvert, G. Le Brun, B. Le Jeune, J. Cariou, and L. Martin, "Decomposition algorithm for an experimental Mueller matrix," Opt. Commun. **282**, 692–704 (2009).
44. A. V. Gopala Rao, K. S. Mallesh, and Sudha, "On the algebraic characterization of a Mueller matrix in polarization optics: I. Identifying a Mueller matrix from its *N* matrix," J. Mod. Opt. **45**, 955–987 (1998).
45. A. V. Gopala Rao, K. S. Mallesh, and Sudha, "On the algebraic characterization of a Mueller matrix in polarization optics: II. Necessary and sufficient condition for Jones-derived Mueller matrix," J. Mod. Opt. **45**, 989–999 (1998).
46. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1968).
47. N. Mukunda, R. Simon, and E. C. G. Sudarshan, "Fourier optics for the Maxwell field: formalism and applications," J. Opt. Soc. Am. A **2**, 416–426 (1985).
48. J. J. van Zyl, C. H. Papas, and C. Elachi, "On the optimum polarization of incoherently reflected waves," IEEE Trans. Antennas Propag. **AP-35**, 818–825 (1987).