Chalmers Publication Library

(article starts on next page)

Four-dimensional Rotations in Coherent Optical Communications

Magnus Karlsson, *Senior Member, IEEE, Fellow, OSA*

Abstract—To model electromagnetic wave propagation for coherent communications without polarization dependent losses, the unitary 2×2 Jones transfer matrix formalism is typically used. In this study, we propose an alternative formalism to describe such transformations based on rotations in four-dimensional (4d) Euclidean space. This formalism is usually more attractive from a communication theoretical perspective, since decisions and symbol errors can be related to geometric concepts such as Euclidean distances between points and decision boundaries. Since 4d rotations is a richer description than the conventional Jones calculus, having six rather than four degrees of freedom (DOF), we propose an extension of the Jones calculus to handle all six DOF. In addition, we show that the two extra DOF in the 4d description represents transformations that are nonphysical for propagating photons, since they does not obey the fundamental quantum mechanical boson commutation relations. Finally, we exemplify on how the nonphysical rotations can change the polarization-phase degeneracy of well-known constellations such as single-polarization QPSK, polarization-multiplexed (PM-)QPSK and polarization-switched (PS-) QPSK. For example, we show how PM-QPSK, which is well known to consist of four polarization states each having four-fold phase degeneracy, can be represented as eight states of polarizations, each with binary phase degeneracy.

Index Terms—Coherent optical transmission, four-dimensional modulation, optical polarization.

I. INTRODUCTION

WITH coherent receivers being increasingly common in optical communications it is important to connect classical optics concepts such as Jones transfer matrices with more communications-relevant topics such as minimum distances, signal constellations, and Voronoi regions. The latter typically require description of the signals in an n-dimensional real Euclidean space, known as *signal space*, with additive noise perturbing the signal independently in the n orthogonal components.

The use of matrix methods to calculate the polarization evolution for electromagnetic waves was, given the earlier discoveries within optics, a relatively late invention, originally proposed by Jones in a series of papers [1]–[8] and Mueller [9], [10] in the 1940s. The reason was probably that before this time relatively few polarization-dependent devices were available, and there was no demand for what we now refer to as "Jones calculus"

and "Stokes–Mueller calculus." There are important differences between the two approaches; the Jones formalism is linear in the electromagnetic field, whereas the Mueller matrices model the transformation of *Stokes vectors*, that originate from optically observable intensities [11], which are proportional to the field squared. Since Stokes vectors are not linear in the field amplitude, and thus not uniform and Gaussian in the description of the additive noise, the formalism is less useful for making symbol decisions in coherent detection. However, a Stokes/Mueller approach may still be useful in the development of synchronization algorithms as, e.g., in [12], as well as for visualizing polarization states and constellations. For practical reasons it is important to distinguish between states of polarization (SOPs) and constellation points, since SOPs generally drift at a much slower rate (millliseconds) than absolute phases (microseconds) in coherent receivers. It is thus less demanding to track SOPs than absolute phases and constellation points.

The main purpose of this paper is to bridge the gap between the polarization calculii of classical optics and the four-dimensional (4d) signal space of communication theory. At first this might seem like a trivial reformulation of the Jones calculus—merely separating the real and imaginary parts of the complex (Jones-) field components into four real vector components. However, as will be shown, and which is somewhat surprising, this leads to a *more general description* than the classical Jones formalism.

In the simplified case of no polarization dependent losses (which this article deals with) the Jones transfer matrix is *unitary*, and the real 4×4 Mueller transfer matrix can be reduced to an *orthogonal* 3×3 matrix [13], [14].

The description of a polarized electromagnetic wave in terms of its four real quadrature components was pioneered by Betti *et al.* [15], [16] and Cusani *et al.* [17]. The signal is then described by a real fou-component vector, with a length squared proportional to the signal power. Lossless propagation of such a signal vector can then be described by the f4d real analog to the complex Jones matrix (see, e.g., Eq. (9) in [17]), which has four independent parameters, or equivalently, degrees of freedom (DOF). Obviously, such a 4×4 transfer matrix is a 4d rotation, since it will not change the length (or power) of the state vector. However, 4d rotations form a six-parameter group [18], implying that the 4 DOF Jones matrices do not represent the full set of possible transformations in real 4d space [19]. The question then emerges: these extra two DOF in the 4d rotations, what kind of transformations do they correspond to in Jones space? This paper will address that question.

As we will see the answer is very interesting, both from a mathematical and physical point of view. There is also a real

practical relevance from the application of matrix methods to fiber optic transmission. We will show that the two extra DOFs of the 4d rotations correspond to *nonphysical optical transformations*, i.e. polarization transformations that can be mathematically described by extensions of the Jones calculus, but which do not have any physical realization in electromagnetic wave propagation. Nevertheless, they are of interest for at least three reasons:

Extension: They enable us to extend the existing Jones calculus to cover a richer set of transformations.

Commutation: The nonphysical transformations commute with the allowed (conventional) ones. This is a unique property for these transformations that may simplify their practical implementation.

Synthesis: Even if the nonphysical transformations do not exist in wave propagation, they can be artificially created in digital signal processing or waveform generation, and can as such be valuable in e.g. novel kinds of coherent optical transmitters and receivers.

This article is organized as follows: In Sections II and III, we will for reference review the well-known Jones and Stokes-Mueller calculii for unitary media. In Section IV, we will describe the corresponding 4d rotational description. In Section V, we describe the general (6 DOF) set of 4d rotations including the nonphysical and allowed transformations, and their parameterization. Then in Section VI, we will connect back to the Jones calculus, and see how it can be extended to deal with the whole 6 DOF set of 4d rotations. We will also explicitly prove that the extra two DOF are nonphysical, in the sense that they do not (contrary to the other four DOF) satisfy the fundamental boson commutation relations. In Section VII, we will, as an example show how the nonphysical transformations can affect known constellations. Finally in Section VIII, we conclude and put the full 6 DOF description into the context of rotations and fiber transmission. A number of appendices provide supporting material, definitions and details not necessary for the main understanding. Parts of this paper, and some more details on the Stokes/Mueller modeling can be found in a related arXiv paper [20].

We employ the following notation conventions: Complex 2-vectors are denoted by boldface lowercase letters, as, e.g., e. Real 3-vectors (e.g. Stokes vectors or rotation vectors) are denoted with lowercase vector-arrow notation as $\vec{e}$. Unit vectors have a hat, as $\hat{e}$. Real four vectors are denoted with uppercase vector-arrow notation as $\vec{E}$. Matrices are usually uppercase, and their dimension should be clear from the context. For example the unity matrix is denoted I independent of dimensionality. A few specific matrices are denoted by lower case greek letters as, e.g., the Pauli matrices σ_j and the left-(right-) isoclinic generators ρ_j (λ_j), and sometimes three constant matrices forms a three vector as, e.g., $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. A scalar product like $\vec{h} \cdot \vec{\sigma}$ should then be interpreted as a linear combination of the three matrices $\sigma_1, \sigma_2, \sigma_3$.

II. JONES DESCRIPTION

The transverse electric field components of a plane electromagnetic wave, propagating in the $z-$direction, has only two components, and can thus be described as the column vector

$$e = \begin{pmatrix} e_x \\ e_y \end{pmatrix}. \quad (1)$$

We refer to this vector as the complex *Jones vector*. Often the Jones vector is given in the frequency domain, so we will allow it to be frequency dependent, although in this paper this dependence is not important, so we will not write it out explicitly. The complex vector elements describe the $x-$ and $y-$ components of the electric field, and their phase relative to some phase reference. For most purposes, e.g. in coherent optical signaling, it suffices to describe how the Jones vector is transformed when propagating in the system. A linear transmission medium can then be described by its complex transfer function, and for electromagnetic waves this generalizes to a complex 2×2 Jones matrix T, that relates the input and output Jones vectors via

$$e(z) = T(z)e(0) \quad (2)$$

where $e(0)$ denotes the input ($z = 0$) Jones vector to the system and z is an arbitrary output position. The transfer matrix T can be an arbitrary 2×2 matrix with complex coefficients, but in this analysis we will restrict ourselves to systems without any polarization-dependent loss (such as transmission fibers) where the signal power $P = e^{*t}e = e^{\dagger}e$ must be independent of the polarization state. This leads to $T^{-1} = T^{\dagger}$ so that T will be a *unitary* transformation matrix, belonging to the *unitary* group $U(2)$. The subgroup of unitary 2×2 matrices with determinant=+1 is called the *special unitary group*, or $SU(2)$.

Since our purpose is to relate the transfer matrix between different descriptions of the signal vector, we need to parametrize the unitary transformation, and to do that will find it useful to describe the *evolution* of the signal within the medium by the differential equation

$$i\frac{de}{dz} = He \quad (3)$$

where H is a matrix describing the transmission medium properties. Again, for lossless transmission we must have $d(e^{\dagger}e)/dz = ie^{\dagger}(H^{\dagger} - H)e = 0$ which implies that the matrix $H = H^{\dagger}$, i.e., it is *Hermitian*. A useful parameterization of Hermitian matrices is via the Pauli spin matrices σ_j (see definition, notation and general properties in Appendix A). This means that H can be written as a linear combination of the four Pauli spin matrices:

$$H = h_0\sigma_0 + h_1\sigma_1 + h_2\sigma_2 + h_3\sigma_3 = h_0 I + \vec{h} \cdot \vec{\sigma} \quad (4)$$

or explicitly

$$H = \begin{pmatrix} h_0 + h_1 & h_2 - ih_3 \\ h_2 + ih_3 & h_0 - h_1 \end{pmatrix}. \quad (5)$$

Here, h_k are four real coefficients describing the medium, which may be formed into a scalar h_0 and a real three vector $\vec{h} = (h_1, h_2, h_3)$. Assuming H to be constant enables the differential equation (3) to be solved using the matrix exponential as

$$e(z) = \exp[-iHz]e(0) = Te(0). \quad (6)$$

The matrix exponential can be expanded in a Taylor series to prove that

$$T = \exp[-iHz] = \exp(-ih_0 z)U(z\vec{h}) \qquad (7)$$

where

$$U(\vec{\alpha}) = \exp(-i\vec{\alpha} \cdot \vec{\sigma}) = [I \cos(\alpha) - i\frac{\vec{\alpha}}{\alpha} \cdot \vec{\sigma} \sin(\alpha)] \qquad (8)$$

denotes the generic special unitary transformation, or $SU(2)$ group member, parameterized by the real three-component vector $\vec{\alpha}$, the modulus of which is the scalar $\alpha = |\vec{\alpha}|$.

III. STOKES–MUELLER DESCRIPTION

The four-component Stokes vector corresponding to the Jones vector e is defined with four real components given by the scalar $P = e^\dagger \sigma_0 e = e^\dagger e$ (which is simply the optical power), and the real three vector $\vec{e} = e^\dagger \vec{\sigma} e$. However, for fully polarized light and in the absence of polarization-dependent losses, it is sufficient to consider the evolution of the vector $\vec{e}$ only.[1] As a consequence the transformation (Mueller) matrix will be a 3×3 rotation matrix. Explicitly, we can express $\vec{e}$ in the complex field components as

$$\vec{e} = \left[|e_x|^2 - |e_y|^2, \quad 2\mathrm{Re}(e_x e_y^*), -2\mathrm{Im}(e_x e_y^*)\right]^t \qquad (9)$$

from which we note that any common—or *absolute*—phase of e_x and e_y will *not* be modeled by the Stokes vector $\vec{e}$. The vector $\vec{e}$ is often described as a point on a sphere, called the *Poincaré sphere*, and polarization transformations can be described as rotations of this sphere. We can obtain the evolution equation for $\vec{e}$ by using the above definitions and Eqs. (3) and (4) to obtain

$$\frac{d\vec{e}}{dz} = e^\dagger i(H^\dagger \vec{\sigma} - \vec{\sigma}H)e = 2\vec{h} \times \vec{e} \qquad (10)$$

where $\vec{h}\times$ denotes the cross product operator

$$\vec{h}\times = \begin{pmatrix} 0 & -h_3 & h_2 \\ h_3 & 0 & -h_1 \\ -h_2 & h_1 & 0 \end{pmatrix}. \qquad (11)$$

The evolution equation (10) has the well-known geometrical interpretation of describing a vector $\vec{e}$ rotating around an axis directed along $\vec{h}$. If $\vec{h}$ is independent of z, (10) can be solved, again using the matrix exponential, with the result

$$\vec{e}(z) = \exp(z2\vec{h}\times)\vec{e}(0) = M(z2\vec{h})\vec{e}(0) \qquad (12)$$

where M denotes the Mueller matrix, which relates the input and output Stokes vectors. In this case when we have no polarization-dependent losses, the Mueller matrix is a rotation matrix, generally expressed as

$$M(\vec{\alpha}) = \exp[\vec{\alpha}\times] =$$
$$I + \sin(\alpha)\frac{\vec{\alpha}\times}{\alpha} + \frac{1 - \cos(\alpha)}{\alpha^2}(\vec{\alpha}\times)^2 =$$
$$I \cos(\alpha) + \sin(\alpha)(\hat{\alpha}\times) + (1 - \cos(\alpha))\hat{\alpha}(\hat{\alpha}\cdot) \qquad (13)$$

[1]We define the Stokes vectors to be instantaneous, in direct correspondence with the Jones vectors, without any temporal averaging.

which describes a rotation an angle $\alpha = |\vec{\alpha}|$ around the unit vector $\hat{\alpha} = \vec{\alpha}/\alpha$. The notation $\hat{\alpha}(\hat{\alpha}\cdot)$ is the projection operator, i.e., the matrix

$$\hat{\alpha}(\hat{\alpha}\cdot) = \frac{1}{\alpha^2} \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2 & \alpha_1\alpha_3 \\ \alpha_1\alpha_2 & \alpha_2^2 & \alpha_2\alpha_3 \\ \alpha_1\alpha_3 & \alpha_2\alpha_3 & \alpha_3^2 \end{pmatrix}. \qquad (14)$$

Therefore the Mueller matrix M is often described geometrically with the Poincaré sphere rotating around the birefringence vector $\vec{h}$. The set of real 3×3 matrices M preserving a real 3d vector length must fulfill $M^{-1} = M^t$, and forms the *orthogonal group* O_3. The subset of O_3 with determinant $= +1$ forms the O_3^+ group.

By comparing the Jones transfer matrix (8) with the Mueller matrix (13) we have a 3 DOF mapping between the two formalisms that illustrates the isomorphism between the O_3^+ matrices $R(\vec{\alpha})$ and the $SU(2)$ matrices $U(\vec{\alpha})$. However the Jones description (although not the $SU(2)$ group) is richer in that it may also contain the absolute phase change on the wave from the medium (the h_0 coefficient in Eq. (7)), which is absent from the Stokes–Mueller approach. On the other hand, the Stokes–Mueller approach facilitates the geometrically pleasing, and also intuitive, interpretation of rotating polarization vectors on the Poincaré sphere.

So far we have done nothing new, but the preceding relations between the Jones and Stokes–Mueller descriptions are well known, and can be found in numerous books and articles on the topic, e.g., [21]–[23]. We provide them for reference and to enable easy connections with a third and less known description, based on the 4d Euclidean space that will be introduced in the next section.

IV. FOUR-DIMENSIONAL SIGNAL SPACE DESCRIPTION

Within communication theory, signals are analyzed in a *signaling space* that is an n-dimensional (nd) vector space containing the set of transmitted signals (the signaling *constellation*), as a set of discrete points in nd. The received signals belong to this set, possibly perturbed by distortions, rotations, and noise. The geometric signal-space description is valuable, because the Euclidean distances between the constellation points will directly be a measure of the constellation's robustness to additive noise [24]. For coherent optical communications the signaling space is four-dimensional (4d), so a real 4d description of the electric field is in many cases preferred over the Jones or Stokes-Mueller formalisms in this context [16], [17], [24].

The Jones description can be taken to an equivalent 4d description by expressing the complex Jones vector e as the real 4-vector

$$\vec{E}(z) = \begin{pmatrix} \mathrm{Re}(e_x) \\ \mathrm{Im}(e_x) \\ \mathrm{Re}(e_y) \\ \mathrm{Im}(e_y) \end{pmatrix}. \qquad (15)$$

A Jones transformation such as (6) can now be written as

$$\vec{E}(z) = N(z)\vec{E}(0) \qquad (16)$$

where $N(z)$ is a real 4d transformation matrix. As we noted for the Jones and Stokes descriptions, the absence of polarization-dependent loss or gain restricts the power $P = \vec{E}^t \vec{E}$ to be conserved and thus $N^t N = I$, so the N is a 4d orthogonal matrix, belonging to the O_4 group. In order to identify the elements of N with those of the Jones matrix T in (6) we study the evolution equation for the vector $\vec{E}$, i.e.

$$\frac{d\vec{E}}{dz} = -K\vec{E} \tag{17}$$

for some matrix K describing the medium. Since P must be constant with respect to z, we can show that K should satisfy $K^t = -K$, i.e., K should be *skew-symmetric*. This means that only the 12 off-diagonal elements of K are non-zero. By separating (3) in real and imaginary parts, and using (4) we find that the matrix K equals

$$K = \begin{pmatrix} 0 & -(h_0 + h_1) & h_3 & -h_2 \\ h_0 + h_1 & 0 & h_2 & h_3 \\ -h_3 & -h_2 & 0 & -(h_0 - h_1) \\ h_2 & -h_3 & h_0 - h_1 & 0 \end{pmatrix}. \tag{18}$$

In analogy with the previous sections, we can integrate (17) if the h_k coefficients are constant, and describe the 4d rotation using the matrix exponential as

$$\vec{E}(z) = \exp(-zK)\vec{E}(0). \tag{19}$$

This description can be further simplified, but for that we need to describe the general properties of 4d rotation matrices, which will be done in the next section.

V. PARAMETRIZATION OF 4D ROTATIONS

The argument in the preceding section, that a rotation should preserve the vector length, can be generalized to show that an nd rotation matrix R can be expressed as the exponential of a skew-symmetric $n \times n$ matrix S. Such a matrix must satisfy $S^t = -S$ which forces the diagonal elements to vanish, and the $n(n-1)/2$ elements above the diagonal is enough to describe all degrees of freedom. In the special case of 4d we have 6 degrees of freedom, i.e., we need six parameters to describe an arbitrary 4d rotation. In Appendix B, we present a complementary discussion of 4d rotations and their nomenclature. Here, we will introduce the following six matrices as the 4d rotation basis (the notation will be motivated later):

$$\rho_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{20}$$

$$\rho_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{21}$$

$$\rho_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{22}$$

which forms the vector $\vec{\rho} = (\rho_1, \rho_2, \rho_3)$ and

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{23}$$

$$\lambda_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{24}$$

$$\lambda_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{25}$$

which forms the vector $\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$. An arbitrary real skew-symmetric 4d matrix L can thus be written as a linear combination of these matrices, and an arbitrary 4d rotation matrix is

$$R(\vec{\alpha}, \vec{\beta}) = \exp(L) = \exp(-\vec{\alpha} \cdot \vec{\rho} - \vec{\beta} \cdot \vec{\lambda}) \tag{26}$$

where $\vec{\alpha}, \vec{\beta}$ are two arbitrary, real, three-vectors parameterizing the 4d rotations 6 DOF. We can thus call the six matrices $\vec{\rho}, \vec{\lambda}$ the *generating matrices* for 4d rotations, and they satisfy the following identities:

$$\rho_k^2 = \lambda_k^2 = -I, \forall k,$$
$$\rho_1 \rho_2 \rho_3 = -\rho_2 \rho_1 \rho_3 = I,$$
$$\lambda_1 \lambda_2 \lambda_3 = -\lambda_2 \lambda_1 \lambda_3 = I \tag{27}$$

where I is the 4d unity matrix. These multiplication rules can be used to show that matrices formed by a linear combination of the ρ_k and the unity matrix form a multiplicative *subgroup* of the group of all real 4d skew-symmetric matrices. The same holds for the group of matrices formed by a linear combination of λ_k and the unity matrix. Quite surprisingly these two subgroubs *commute*, since

$$\rho_k \lambda_j = \lambda_j \rho_k \tag{28}$$

for all j and k. Note that multiplications *within* each subgroup is non-commuting, according to (27). The commuting property means that we can write the matrix exponential as the matrix product

$$R = R_R R_L = R_L R_R \tag{29}$$

where

$$R_R(\vec{\alpha}) = \exp(-\vec{\alpha} \cdot \vec{\rho}), \tag{30}$$
$$R_L(\vec{\beta}) = \exp(-\vec{\beta} \cdot \vec{\lambda}). \tag{31}$$

TABLE I
RELATION BETWEEN THE CONVENTIONAL JONES, STOKES–MUELLER AND 4D VECTOR DESCRIPTIONS OF ELECTROMAGNETIC TRANSFORMATIONS

	Signal Vector	Input/Output Transformation	Transformation Matrix
Jones	$\boldsymbol{e} = \begin{pmatrix} e_x \\ e_y \end{pmatrix}$	$\boldsymbol{e}_{\text{out}} = T\boldsymbol{e}_{\text{in}}$	$T = \exp[-i\alpha_0 - i\vec{\alpha} \cdot \vec{\sigma}] =$ $\exp(-i\alpha_0)[I\cos(\alpha) - i\hat{\alpha} \cdot \vec{\sigma}\sin(\alpha)]$
Stokes-Mueller	$\vec{e} = \begin{pmatrix} |e_x|^2 - |e_y|^2 \\ 2\text{Re}(e_x e_y^*) \\ -2\text{Im}(e_x e_y^*) \end{pmatrix}$	$\vec{e}_{\text{out}} = M\vec{e}_{\text{in}}$	$M = \exp[2\vec{\alpha}\times] =$ $I\cos(2\alpha) + \sin(2\alpha)\hat{\alpha} \times +(1-\cos(2\alpha))\hat{\alpha}(\hat{\alpha}\cdot)$
Four-dimensional	$\vec{E} = \begin{pmatrix} \text{Re}(e_x) \\ \text{Im}(e_x) \\ \text{Re}(e_y) \\ \text{Im}(e_y) \end{pmatrix}$	$\vec{E}_{\text{out}} = R\vec{E}_{\text{in}}$	$R = \exp[-(\alpha_0, 0, 0) \cdot \vec{\lambda} - \vec{\alpha} \cdot \vec{\rho}] =$ $\exp(-\alpha_0\lambda_1)[I\cos(\alpha) - \hat{\alpha} \cdot \vec{\rho}\sin(\alpha)]$

This implies that we can decompose an arbitrary 4d rotation into two commuting subgroups, which are often denoted *right-isoclinic*, R_R, and *left-isoclinic*, R_L, rotations, respectively [18]. An isoclinic rotation rotates all vectors the same angle. In Appendix B, we discuss isoclinic rotations further. Thus any 4d rotation can be written as one right-isoclinic rotation followed by a left-isoclinic rotation, or vice versa. The isoclinic rotations can be written in an explicit form by expanding the matrix exponential in a Taylor series and using

$$(\vec{\alpha} \cdot \vec{\rho})^2 = -I\alpha^2 \qquad (32)$$

to obtain

$$R_R = \exp[-\vec{\alpha} \cdot \vec{\rho}] = \exp[-\alpha\hat{\alpha} \cdot \vec{\rho}] =$$

$$\sum_{k=0}^{\infty} \frac{(-\alpha\hat{\alpha} \cdot \vec{\rho})^k}{k!} = I - (\hat{\alpha} \cdot \vec{\rho})\alpha - I\frac{\alpha^2}{2} + (\hat{\alpha} \cdot \vec{\rho})\frac{\alpha^3}{6} + \cdots$$

$$= I\cos(\alpha) - (\hat{\alpha} \cdot \vec{\rho})\sin(\alpha). \qquad (33)$$

Analogously, we find for the left-isoclinic rotations

$$R_L = I\cos(\beta) - (\hat{\beta} \cdot \vec{\lambda})\sin(\beta). \qquad (34)$$

Eqs. (29), (34), and (33) provides a full, 6 DOF parameterization of the 4d rotations that is straightforward to connect to the Jones and Mueller transformations. The order and sign of the generating matrices (20)–(25) are selected to comply with conventional definitions in Stokes and Jones calculus [20].

We can now return to Section IV and identify the elements $h_0, \vec{h}$ from the matrix K in (18) with the general 4d exponent L. This gives $\vec{\alpha} = z\vec{h}$ and $\vec{\beta} = z(h_0, 0, 0)$. We can thus express the 4d signal (19) in terms of 4d rotations as

$$\vec{E}(z) = \exp(-z\vec{h} \cdot \vec{\rho})\exp(-z(h_0, 0, 0) \cdot \vec{\lambda})\vec{E}(0). \qquad (35)$$

This completes the description of polarization transformations in terms of 4d rotations, and the result is that we now have three alternative descriptions of polarization transformations. These are the complex Jones calculus in Eq. (2), the real Stokes–Mueller description in Eq. (12) and the real 4d rotation description from Eq. (35). Those have, for reference. been summarized in Table I.

The above concludes a relatively straightforward extension of the Jones calculus to 4d, that in itself might be useful in, e.g., fiber communication analysis [17], [19], but it provides nothing really physically novel. However, the remainder of this paper will be devoted to the 4d rotations described by the left-isoclinic matrices λ_2 and λ_3, which seem to have no correspondence in

the classical Jones- or Stokes-vector descriptions. As we saw above, only the right-isoclinic rotations (33) of the 4d E-field vector have a physical realization in terms of the corresponding Jones and Mueller matrices. Obviously the λ_1-component gives a constant phase shift of the field vector, so that has a clear physical interpretation. However, the λ_2- and λ_3-components describe transformations that are not used within the classical Jones- or Stokes-Mueller calculus, and the reason is simply that they cannot be realized as physical transformations of the electromagnetic field. In fact, a propagating photon cannot undergo such a change, irrespective of the physical realization. The proof of this statement lies in that those state changes will not satisfy the bosonic commutation relations, as will be motivated in the next section and Appendix C. We will therefore refer to those transformations as *nonphysical rotations*. We should emphasize that the existence of the nonphysical rotations is not a shortcoming of the known polarization calculi. It is merely a richer property of the 4d real space that is not needed for the description of propagating waves, but which can be taken advantage of if we understand it.

A first clue to an understanding of the nonphysical rotations will be to go back to the Jones formalism, and attempt to express these transformations there.

VI. THE NONPHYSICAL ROTATIONS IN JONES SPACE

The left-isoclinic transformation acting on the 4d vector

$$\vec{A} = \begin{pmatrix} \text{Re}(a_x) \\ \text{Im}(a_x) \\ \text{Re}(a_y) \\ \text{Im}(a_y) \end{pmatrix} \qquad (36)$$

with the corresponding Jones vector $\boldsymbol{a} = (a_x\ a_y)^t$ is

$$\vec{A}_{\text{out}} = \exp(-\vec{\beta} \cdot \vec{\lambda})\vec{A}_{\text{in}} \qquad (37)$$

where the subscripts denote input/output states. If we write this in complex form, by collecting the proper real and imaginary parts, it can be expressed as

$$\begin{pmatrix} a_x \\ a_y^* \end{pmatrix}_{\text{out}} = U^*(\vec{\beta})\begin{pmatrix} a_x \\ a_y^* \end{pmatrix}_{\text{in}} \qquad (38)$$

where U^* is the unitary transformation defined in (8) with conjugated elements. This is still a unitary transformation but now acting on the vector $(a_x\ a_y^*)^t$, i.e. the Jones vector with the y-component conjugated. Such transformation vectors (with one

TABLE II
EXTENSION OF THE CONVENTIONAL JONES CALCULUS TO ACCOUNT FOR THE FULL 4D ROTATIONS

	Signal State	Transformation	Transformation Matrix
Jones	$\boldsymbol{E} = \begin{pmatrix} e_x & e_y^* \\ e_y & -e_x^* \end{pmatrix}$	$\boldsymbol{E}_{\text{out}} = U(\vec{\alpha})\,\boldsymbol{E}_{\text{in}}U^\dagger(\vec{\beta})$	$U(\vec{\alpha}) = \exp[-i\vec{\alpha}\cdot\vec{\sigma}]$ $U(\vec{\beta}) = \exp[-i\vec{\beta}\cdot\vec{\sigma}]$
Four-dim.	$\vec{E} = \begin{pmatrix} \text{Re}(e_x) \\ \text{Im}(e_x) \\ \text{Re}(e_y) \\ \text{Im}(e_y) \end{pmatrix}$	$\vec{E}_{\text{out}} = R\vec{E}_{\text{in}}$	$R = \exp[-\vec{\beta}\cdot\vec{\lambda} - \vec{\alpha}\cdot\vec{\rho}]$

element conjugated) arise, e.g., in the context of parametric amplifiers [25], but for that case the transformation matrices are not unitary. In fact, in Appendix C we show that the transformation (38) is only consistent with the quantum mechanical formulation of polarization operators if the second and third components vanish, i.e., $\vec{\beta} = (\beta_1, 0, 0)$. In that special case, (38) reduces to

$$\begin{pmatrix} a_x \\ a_y^* \end{pmatrix}_{\text{out}} = \begin{pmatrix} \exp(-i\beta_1) & 0 \\ 0 & \exp(i\beta_1) \end{pmatrix}\begin{pmatrix} a_x \\ a_y^* \end{pmatrix}_{\text{in}} \quad (39)$$

which is the same phase shift applied to both field components, i.e., an absolute phase change. This rotation is not unphysical, but rather the most rapidly changing degree of freedom of all in coherent systems. On the other hand, the changes described by the λ_2 and λ_3 rotation generators does not obey the boson commutation relations, as shown in Appendix C, and we conclude that they are indeed nonphysical for propagating photons, and we will refer to them as such. The only exception occurs for the special cases where the isoclinic rotation angle is a multiple of π, for which the rotation matrix is plus or minus the unity matrix, irrespective of the rotation generator, cf. (33) and (34). However, this does not preclude a mathematical description of the nonphysical rotations, nor the fact that they can be synthesized artificially, e.g. by DSP. In the following discussion we will refer to the right (left) isoclinic rotations with the dimensionless parameters $\vec{\alpha}$ ($\vec{\beta}$).

As we saw above, the transformation (38) is unitary, but with a *redefined Jones state vector*. This suggests the following unified description of left- and right isoclinic rotations in Jones space. We replace the Jones vector with a *Jones state matrix* defined as

$$E = \begin{pmatrix} e_x & e_y^* \\ e_y & -e_x^* \end{pmatrix}. \quad (40)$$

With this definition a right-isoclinic (i.e. the conventional) transformation is performed by applying a unitary transformation to the two *column* vectors. Similarly, a left-isoclinic (comprising the phase and nonphysical) transformation is realized by applying a unitary transformation to the *row* vectors. The first column vector is the conventional Jones vector and the second is the corresponding orthogonal state of polarization, and obviously these two must transform with the same unitary matrix, so the transformation UE with U unitary, is just a trivial extension of the standard Jones analysis. However, the left-isoclinic rotation (38) can now be expressed as $U^*E^t = (EU^\dagger)^t$, i.e., as a *multiplication from the right* with the *conjugate transposed* unitary matrix $U^\dagger$. To summarize, the full input–output relation for the

general, 6-DOF, 4d rotation

$$\vec{E}_{\text{out}} = R(\vec{\alpha}, \vec{\beta})\vec{E}_{\text{in}} \quad (41)$$

where R is given by (26), can thus be compactly expressed by the Jones state matrix as

$$E_{\text{out}} = U(\vec{\alpha})E_{\text{in}}U^\dagger(\vec{\beta}) \quad (42)$$

where again U denotes the unitary matrix as defined in (8). In this way, we can recover the full six-parameter rotation family by left and right multiplying the Jones state matrix with unitary matrices.

It is noteworthy that the Jones state matrix is of unitary form, but with determinant $-P$, which means that its unitary structure (40) is preserved by the transformation (42), as expected. The commutation between right- and left-isoclinic rotations is also consistent with (42), due to the associativity of matrix products. This "generalized" Jones calculus is summarized in Table II.

Finally we should state, for completeness, that a similar generalization of Jones calculus as given by (42) is possible also for the Stokes–Mueller calculus. In that case, the Stokes vector is generalized to a *Stokes state matrix*, which can be related to the Jones state matrix (40), cf. [20, eq. (F5)]. The right- and left-isoclinic transformations are also there described by rightwards and leftwards matrix multiplications. In [20] we also show that there is a strong connection with the Stokes state matrix and the "paddled-Stokes vector" introduced by Frigo *et al.* [26] and recently used for system analysis in [27]. In fact Frigo's "paddle" emerges as a column vector of said Stokes state matrix.

VII. EXAMPLES: QPSK, PS-QPSK, AND PM-QPSK

We have thus derived a useful representation of 4d rotations as $R(\vec{\alpha}, \vec{\beta})$ where the 3-vector α models the polarization changes, the first component, β_1, of $\vec{\beta}$ models the absolute phase shift, and the remaining two parameters, $\beta_{2,3}$ are nonphysical. It is instructive to give a few examples on how the nonphysical transformations affect the constellations of some well-known modulation formats in 4d coherent communications. We will consider three examples. The presented constellations are intended more as examples than propositions for new, promising formats. The latter would require more careful evaluation, and could be an promising topic for future research.

A. QPSK

The simplest non-trivial example of an nonphysical rotation acting on a constellation is probably single-polarization

QPSK, which can be represented as the 4d constellation (or point set) $\mathcal{C}_{QPSK1} = \{(\pm 1, \pm 1, 0, 0)^t\}/\sqrt{2}$, taken with all four sign combinations. If we, as an example, apply the nonphysical 4d rotation given by $\vec{\alpha} = (0, 0, 0)$ and $\vec{\beta} = (0, 0, \pi/4)$ we obtain the constellation $\mathcal{C}_{QPSK2} = \pm\{(-1, 1, 1, 1)^t, (1, 1, -1, 1)^t\}/2$, which occupies two orthogonal polarization states (left- and right-hand circular) with binary PSK modulation in each. The QPSK1 constellation points $\pm(1, 1, 0, 0)^t$ moves to the right-hand circular SOP, whereas the $\pm(1, -1, 0, 0)^t$ constellation points moves to the left-hand circular SOP. Such a transformation cannot be realized within conventional passive optics, but is possible to realize in the electronic domain using DSP. We emphasize that this is merely a rotation, so the $\mathcal{C}_{QPSK2}$ constellation has the same properties (e.g. sensitivity and spectral efficiency) as standard QPSK, but it has one important practical benefit: The local oscillator (LO) phase noise resilience will be improved since we have, after appropriate polarization demux, two orthogonally polarized BPSK streams rather than one single-polarization QPSK stream. The prize paid is that one occupies two polarization states rather than one. This also shows that single-polarization QPSK can be geometrically equivalent to dual-polarization BPSK, which may not be widely known.

B. PS-QPSK

The polarization-switched (PS)-QPSK format is interesting since its asymptotic power efficiency is the best among all un-coded 4d formats (1.76 dB better than QPSK) [28], but with a reduced spectral efficiency relative to QPSK. It has led to a number of theoretical [24] and experimental [29], [30] investigations. The constellation for PS-QPSK has in the literature been described by three different representations:

1) the 4d cross-polytope

$$\mathcal{C}_1 = \{(\pm 1, 0, 0, 0)^t, (0, \pm 1, 0, 0)^t,$$
$$(0, 0, \pm 1, 0)^t, (0, 0, 0, \pm 1)^t\}; \quad (43)$$

2) QPSK in either the x- or y-polarization, which consists of all sign selections of

$$\mathcal{C}_2 = \frac{1}{\sqrt{2}}\{(\pm 1, \pm 1, 0, 0)^t, (0, 0, \pm 1, \pm 1)^t\}; \quad (44)$$

3) the hypercube subset with even parity,

$$\mathcal{C}_3 = \frac{1}{2}\{\pm(1, 1, 1, 1)^t, \pm(1, 1, -1, -1)^t.$$
$$\pm(1, -1, 1, -1)^t, \pm(1, -1, -1, 1)^t\}. \quad (45)$$

These constellations can be transformed to each other by simple physical rotations; for example $\mathcal{C}_1$ is taken to $\mathcal{C}_2$ by a 45° phase rotation ($\vec{\alpha} = (0, 0, 0), \vec{\beta} = (\pi/4, 0, 0)$). Similarly $\mathcal{C}_3$ is obtained from $\mathcal{C}_2$ by a 45° polarization rotation ($\vec{\alpha} = (0, 0, \pi/4), \vec{\beta} = (0, 0, 0)$).

All these representations have in common that they describe PS-QPSK as two polarization states that each has a fourfold absolute phase degeneracy. However, this degeneracy can be broken by applying a nonphysical rotation. As an example, apply the 4d rotation $\vec{\alpha} = (0, 0, 0), \vec{\beta} = (0, 0, \beta_3)$ to $\mathcal{C}_1$ to obtain the
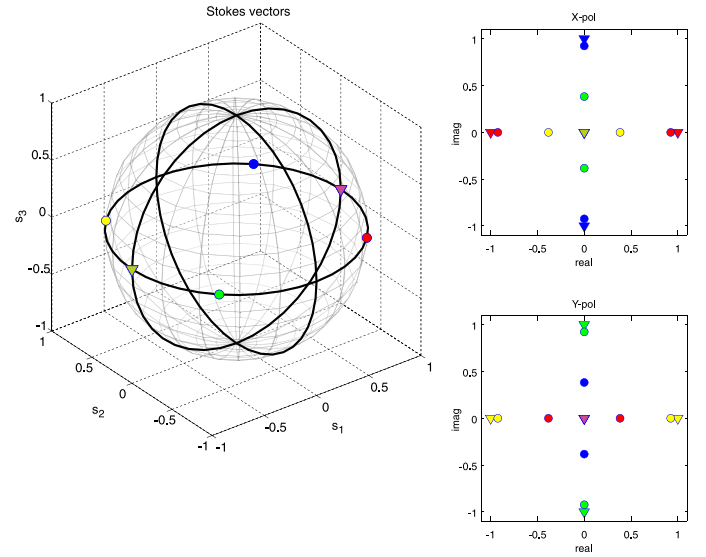

Fig. 1. Stokes vectors and constellations for PS-QPSK in configuration $\mathcal{C}_1$ (triangles) and $\mathcal{C}_4$ (circles), which is obtained from $\mathcal{C}_1$ by applying the nonphysical rotation $R = \exp(-\vec{\lambda} \cdot (0, 0, \pi/8))$. The points are color coded to show correspondence between Stokes space and constellation diagrams. Magenta triangles denotes coinciding blue and red points, and dito for green-yellow-ish triangles.

constellation

$$\mathcal{C}_4 = \{\pm(c, 0, s, 0)^t, \pm(0, c, 0, -s)^t,$$
$$\pm(-s, 0, c, 0)^t, \pm(0, s, 0, c)^t\} \quad (46)$$

where $c = \cos(\beta_3)$ and $s = \sin(\beta_3)$. This PS-QPSK representation is unique in that it can have four polarization states, each of which has twofold phase degeneracy. In the special case $\beta_3 = \pi/8$ they are symmetric and separated by 90° on the Poincaré sphere as illustrated in Fig. 1. The four subsets in (43) are color coded, red, blue, yellow and green, respectively. We emphasize that $\mathcal{C}_4$ is only a 4d-rotated version of the conventional PS-QPSK constellation $\mathcal{C}_1$, so both representations has the same sensitivity and spectral efficiency. However, $\mathcal{C}_4$ might potentially be easier to detect in a coherent receiver, since it puts more (fourfold) degeneracy in the polarization DOFs which requires slower tracking in the receiver, and less (twofold) degeneracy in the absolute phase DOF. In other words, representations $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{C}_3$ consists of two polarization states with four absolute phases each, in contrast with $\mathcal{C}_4$ which consists of four polarization states with two phases each. The corresponding phase tracking DSP for $\mathcal{C}_4$ needs to deal with only a binary rather than quaternary phase modulation, which is known to have less strict linewidth requirements. A drawback is that the polarization tracking algorithm will be more complex, and the conventional constant modulus algorithm will likely not work without modifications. Also the ambiguities in polarization and phase of these new representations would require further investigation, as well as how well such algorithms can deal with nonlinear distortions.

To suggest a detailed realization for such a coherent receiver is beyond the scope of this article, but the main point is that the nonphysical rotations can be used to move degeneracies
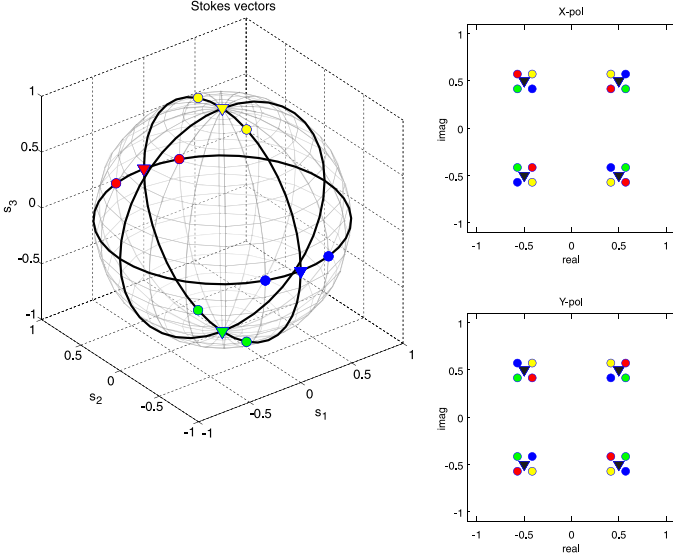
Fig. 2. Stokes vectors and constellations for PM-QPSK in configuration $\mathcal{C}_5$ (triangles), and $\mathcal{C}_5$ after the nonphysical rotation $R = \exp\left(-\vec{\lambda} \cdot (0, \pi/20, 0)\right)$ (circles). The points are color coded to show correspondence between Stokes space and constellation diagrams. Black triangles denotes coinciding blue, red, green and yellow points.

between the polarization and the phase degree of freedoms of coherent modulation formats in a hitherto unknown way.

C. PM-QPSK

The standard PM-QPSK is a 16-ary format with the points located on the vertices of the 4D hypercube:

$$\mathcal{C}_{4\text{cube}} = \frac{1}{2}\{(\pm1, \pm1, \pm1, \pm1)^t\} \tag{47}$$

taken with all sign combinations. In Stokes space these 16 levels lie on 4 different polarization states ($\pm45°$, right- and left-hand circular), each with a fourfold absolute-phase degeneracy. The constellation and Stokes representation is shown as triangles in Fig. 2. The 16 points are color coded in the four subsets given by the 4 polarization states of $\mathcal{C}_{4\text{cube}}$. To apply any of the physical rotations ($\vec{\alpha}$) would correspond to rotating the constellation in Stokes space around the vector $\vec{\alpha}$, an angle $2|\vec{\alpha}|$. An absolute phase rotation β_1 would not affect the Stokes representation, but it would rotate the constellation diagrams in the x- and y-pol. The nonphysical rotations are not so easy to interpret, however. An example with $\vec{\beta} = (0, \pi/20, 0)$ is shown as the circles in Fig 2. This corresponds to a $20°$ movement on the Poincaré sphere, but the rotation axis depend on the absolute phase of the constellation point. As a result, we can see (circles in Fig 2) the four-fold phase degeneracy of each polarization state in $\mathcal{C}_{4\text{cube}}$ to be broken into two polarization states, each with two-fold phase degeneracy. This reflects the sign which is ambiguous for Stokes vectors, just as in the previous examples for QPSK and PS-QPSK. The constellation points (shown in Fig 2(b)) are seen to split in 4 points each, which should be interpreted so that every point in the x-pol constellation diagram has one (and only one) partner in the y-pol diagram. Again, this representation of PM-QPSK might be more robust to e.g. local oscillator phase

noise as it has only a 2-fold phase degeneracy, but to detect it and benefit from it would require a novel coherent receiver structure and DSP.
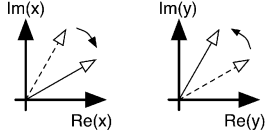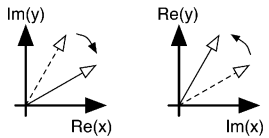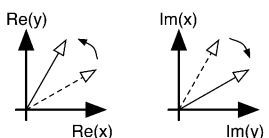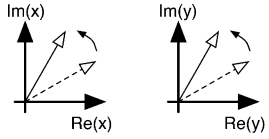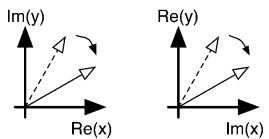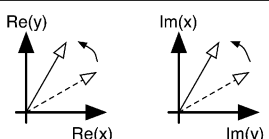
VIII. DISCUSSION AND CONCLUSION

In this paper we have presented the 6 DOF four-dimensional rotation group, and for the first time related and classified these rotations in terms of the more well-known Jones calculus used in polarization optics. The two extra DOFs that have no counterpart in classical Jones analysis were shown to be nonphysical, i.e. not arising for propagating photons. The parametrization of 4d rotations in terms of well-known optical transformations is summarized in Table III.

These findings are relevant for optical communications, where a multidimensional real formalism in 4d space is often preferred over the complex analysis. The fact that we can extend the Jones analysis to model also the nonphysical rotations may be of value in applications where such non-physical rotations are synthesized, e.g. by using DSP. In Table III we list the isoclinic rotations and their mathematical (and physical) interpretations. Two commuting classes of rotations, left- and right-isoclinic, are interpreted schematically and physically, and the typical drift times in coherent fiber links are given. These drift times will naturally vary depending on the specific link realization. However, the polarization (right-isoclinic) drifts are usually in the millisecond regime or longer, originating from physical movements or temperature changes along the fiber cable. The absolute phase (left-isoclinic, β_1) drift times depend on the signal and local oscillator linewidths, and are are typically around microseconds. The entries in the "schematic" column and the underlying axis change between each entry in this column are further explained in Appendix B.

The indicated drift times are the tracking times that coherent receivers need to satisfy, and clearly it is more favorable to modulate data in the polarization (right-isoclininc) degrees of freedom which are less demanding to track. The non-physical rotations cannot arise during propagation of light, but can be synthesized in DSP. A potential use of the such rotations (left-isoclininc with $\beta_2 \neq 0$ and/or $\beta_3 \neq 0$) are to move modulation format degrees of freedom between polarization and absolute phase. We gave a few examples of this; the simplest being QPSK that can be transmitted either as single-polarization QPSK or dual-polarization BPSK after applying an nonphysical rotation. Similarly, PM-QPSK which is often seen as QPSK in 4 different polarization states, can after applying an nonphysical rotation actually be transmitted as BPSK in 8 different polarization states. The idea of moving signaling degrees of freedom from absolute phase to polarization will likely be even more attractive for multi-level QAM formats such as e.g. 16-QAM, since those are often limited by phase noise. This will however require development of novel DSP algorithms, but there may well be a potential use of the nonphysical rotations in receiver synchronization algorithms, although to give specific examples of such receiver implementations are beyond the scope of this paper.

TABLE III
CLASSIFICATION OF THE 6 DOF 4D ROTATIONS, AND THEIR SCHEMATIC AND TYPICAL DRIFT TIMES IN COHERENT TRANSMISSION SYSTEMS

Rotation class	Rotation Matrix	Physical interpretation	Schematic	Typical drift time
Right-isoclinic (two counter-directed simple rotations)	$\exp[-\alpha_1 \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}]$	Linear birefringence. Phase retardation between $\hat{x}$ and $\hat{y}$ polarization components; Im(x) rotates towards Re(x) while Re(y) rotates towards Im(y).		milliseconds
	$\exp[-\alpha_2 \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}]$	Linear birefringence. Phase retardation between $\hat{x} + \hat{y}$ and $\hat{x} - \hat{y}$ polarization components; Im(y) rotates towards Re(x) while Im(x) rotates towards Re(y).		milliseconds
	$\exp[-\alpha_3 \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}]$	Circular birefringence. Phase retardation between $\hat{x} + i\hat{y}$ and $\hat{x} - i\hat{y}$ polarization components; rotation of Re(x) towards Re(y) and Im(x) towards Im(y).		milliseconds
Left-isoclinic (two co-directed simple rotations)	$\exp[-\beta_1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}]$	Common phase change for $\hat{x}$- and $\hat{y}$-polarizations		microseconds
	$\exp[-\beta_2 \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}]$	Nonphysical, Re(x) - Im(y) rotates in opposite direction of Re(y) - Im(x)		–
	$\exp[-\beta_3 \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}]$	Nonphysical, Re(x) - Re(y) rotates in opposite direction of Im(x) - Im(y)		–

To summarize, we have presented an extension of the conventional polarization analysis to cover the full degrees of freedom of real four-dimensional rotations. We have identified physical as well as nonphysical degrees of freedom in the process. Future work on this theory could be to extend it to a full polarization calculus, accounting also for polarizing elements with polarization dependent loss or gain, as well as looking in to applications for the nonphysical transformations.

APPENDIX A

THE PAULI MATRICES

As is common in polarization work [21], [23] we define the Pauli matrices in, for physics literature, nonstandard order (cf. [31, Ch. 4.5]), which complies with the Stokes vector definition used in the classic optical texts as e.g. [32]. Thus the Pauli spin matrices are defined by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{48}$$

which can be formed into a 3-vector with 2×2 matrices as elements $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, as suggested by Fano [33]. The Pauli matrices are Hermitian and have zero trace (except for $Tr(\sigma_0) = 2$). They satisfy the multiplication rules

$$\sigma_0 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

$$\sigma_1 \sigma_2 = -\sigma_2 \sigma_1 = i\sigma_3 \tag{49}$$

where the last equation allow for cyclic permutation of indices. This also shows that each Pauli matrix (excluding σ_0) anticommutes with the other 2 Pauli matrices, and commutes with itself and σ_0. The Pauli matrices are linearly independent and can be used as a basis for all complex 2×2 matrices. In other words, an arbitrary complex 2×2 matrix C can be written as

$$C = \sum_{k=0}^{3} c_k \sigma_k \tag{50}$$

where c_k are complex coefficients given by

$$c_k = \text{Tr}(C\sigma_k)/2. \tag{51}$$

If we now take C to be the *coherency matrix*

$$C = \boldsymbol{e}\boldsymbol{e}^\dagger = \begin{pmatrix} e_x e_x^* & e_x e_y^* \\ e_x^* e_y & e_y e_y^* \end{pmatrix} \tag{52}$$

the coefficients c_k are given by

$$c_k = \mathrm{Tr}(\boldsymbol{e}\boldsymbol{e}^\dagger \sigma_k)/2 = (\boldsymbol{e}^\dagger \sigma_k \boldsymbol{e})/2 \tag{53}$$

where the last expression for $k = 1, 2, 3$ are the components of the Stokes vector $\vec{e}$. Thus we obtain the useful relation between Jones and Stokes vectors

$$\boldsymbol{e}\boldsymbol{e}^\dagger = \frac{P + \vec{e} \cdot \vec{\sigma}}{2} \tag{54}$$

as originally pointed out in [33], [34].

APPENDIX B

THE 4D ROTATIONS

Much of the material in this appendix can be found in the Wikipedia article [18], but it is given here for reference and as an introduction.

To classify the 4d rotations we start by defining the *simple rotations*, that rotate only two coordinate axes while leaving the plane spanned by the other two invariant. For example a simple rotation in the 1,2-plane is

$$T_{12}(\phi) = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 & 0 \\ -\sin(\phi) & \cos(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{55}$$

which leaves the 3-4-plane invariant. Another one is

$$T_{24}(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & 0 & \sin(\phi) \\ 0 & 0 & 1 & 0 \\ 0 & -\sin(\phi) & 0 & \cos(\phi) \end{pmatrix} \tag{56}$$

which leaves the 1,3-plane invariant. In this way we can define six simple rotations $T_{12}, T_{13}, T_{14}, T_{23}, T_{24}, T_{34}$ that span the full 6-DOF space of 4d rotations.

The most general rotation in 4d is the *double rotation*, which leaves only the origin invariant. For example we can realize a double rotation as the product of two simple rotations as

$$D_a(\phi_1, \phi_2) = T_{12}(\phi_1)T_{34}(\phi_2)$$

$$= \begin{pmatrix} \cos(\phi_1) & \sin(\phi_1) & 0 & 0 \\ -\sin(\phi_1) & \cos(\phi_1) & 0 & 0 \\ 0 & 0 & \cos(\phi_2) & \sin(\phi_2) \\ 0 & 0 & -\sin(\phi_2) & \cos(\phi_2) \end{pmatrix}$$

And we may similarly define $D_b(\phi_1, \phi_2)$ and $D_c(\phi_1, \phi_2)$ in an analogous way as

$$D_b(\phi_1, \phi_2) = T_{14}(\phi_1)T_{23}(\phi_2)$$

$$= \begin{pmatrix} \cos(\phi_1) & 0 & 0 & -\sin(\phi_1) \\ 0 & \cos(\phi_2) & -\sin(\phi_2) & 0 \\ 0 & \sin(\phi_2) & \cos(\phi_2) & 0 \\ \sin(\phi_1) & 0 & 0 & \cos(\phi_1) \end{pmatrix}$$

and

$$D_c(\phi_1, \phi_2) = T_{13}(\phi_1)T_{24}(\phi_2)$$

$$= \begin{pmatrix} \cos(\phi_1) & 0 & -\sin(\phi_1) & 0 \\ 0 & \cos(\phi_2) & 0 & \sin(\phi_2) \\ \sin(\phi_1) & 0 & \cos(\phi_1) & 0 \\ 0 & -\sin(\phi_2) & 0 & \cos(\phi_2) \end{pmatrix}$$

to get a set of 3 double rotations that span all six DOF.

A double rotation for which the two rotation angles are equal is called *isoclinic*, and will rotate any vector the same angle. We will distinguish between the *left-isoclinic*, which have the two double rotations going in the same direction, and *right-isoclinic* where the rotation angles are in opposite directions. Thus $D_{a,b,c}(\phi, \phi)$ are the left- and $D_{a,b,c}(\phi, -\phi)$ are the right-isoclinic rotations. It is noteworthy that the relative "directions" of two single rotations depends on the chosen ordering of the 4 basis axes, e.g. 1234. If two neighboring axis are swapped, e.g. to the system 2134, the right-isoclinic rotations becomes left-isoclinic and vice versa. It follows that the iso-clinicity is unaffected by two such swappings, which includes a cyclic permutations of any chosen triplet of the coordinate axes. The coordinate axes shown in the "schematic" column of Table III runs through such triple-cyclic permutations that leaves the Re(x)-axis (and the isoclinicity) unaffected as we go down the column.

Any simple rotation can then be described as the product of a right- and a left-isoclinic rotation with the same angles.

The left- and right-isoclinic rotations form two 3-DOF *subgroups*, i.e. any sequence of left-isoclinic rotations will always remain left-isoclinic, and vice versa. The really interesting property of the isoclinic parameterization, however, is that any left-isoclinic rotation *commutes* with any right-isoclinic. Rotations within each subgroup do not commute, however. This enables any 4d rotation to be expressed as a (commuting) product of one right- and one left-isoclinic rotation. The underlying group theoretical reason for this is that the 4d rotation group, $SO(4)$ is isomorphic with the product of two 3d rotation groups $O_3^+ \times O_3^+$.

APPENDIX C

LEFT-ISOCLINIC PHOTON TRANSFORMATIONS ARE (MOSTLY) NONPHYSICAL

Dr. Colin McKinstrie suggested the derivation in this appendix to me.

We make the customary extension (similarly to e.g. in [35], [36]) to quantum mechanics where the electromagnetic field amplitudes e_x and e_y correspond to the quantum mechanical

creation operators $\hat{\psi}_x$ and $\hat{\psi}_y$, and annihilation operators $\hat{\psi}_x^\dagger$ and $\hat{\psi}_y^\dagger$ for the respective polarization components. Examples on use and for these operators in a polarization context can be found in e.g. [35], [37]. As postulated by quantum mechanics (e.g. [38, Ch. 10.3]) these operators obey the boson commutation relations

$$[\hat{\psi}_j, \hat{\psi}_k^\dagger] = \hat{\psi}_j \hat{\psi}_k^\dagger - \psi_k^\dagger \hat{\psi}_j = \delta_{jk} \qquad (57)$$

for $j, k \in \{x, y\}$. We will now study input-output transformations for these operators, and the requirements the commutation relations put on the transformation matrices. We consider left- and right-isoclinic transformations separately.

Right-isoclinic transformations: This is the conventional input-output unitary transformation

$$\begin{pmatrix} \hat{\psi}_{xo} \\ \hat{\psi}_{yo} \end{pmatrix} = \begin{pmatrix} \gamma & \eta \\ -\eta^* & \gamma^* \end{pmatrix} \begin{pmatrix} \hat{\psi}_{xi} \\ \hat{\psi}_{yi} \end{pmatrix} \qquad (58)$$

where subscripts i, o correspond to input, output, and where we know from the unitary property that the coefficients γ and η satisfy

$$|\gamma|^2 + |\eta|^2 = 1. \qquad (59)$$

If we now apply the commutator relations (57) to the output operators, we find e.g. for the $x-$component

$$[\hat{\psi}_{xo}, \hat{\psi}_{xo}^\dagger] = (\gamma \hat{\psi}_{xi} + \eta \hat{\psi}_{yi})(\gamma^* \hat{\psi}_{xi}^\dagger + \eta^* \hat{\psi}_{yi}^\dagger)$$
$$- (\gamma^* \hat{\psi}_{xi}^\dagger + \eta^* \hat{\psi}_{yi}^\dagger)(\gamma \hat{\psi}_{xi} + \eta \hat{\psi}_{yi}) \qquad (60)$$

and by using (57) we can reduce this to

$$1 = [\hat{\psi}_{xo}, \hat{\psi}_{xo}^\dagger] = |\gamma|^2 [\hat{\psi}_{xi}, \hat{\psi}_{xi}^\dagger] + |\eta|^2 [\hat{\psi}_{yi}, \hat{\psi}_{yi}^\dagger] =$$
$$= |\gamma|^2 + |\eta|^2 \qquad (61)$$

which equals the unitarity condition. The same can be for the y-component. This shows that the commutator relations are satisfied for all right-isoclinic transformations of the polarization operators.

Left-isoclinic transformations: This transformation is

$$\begin{pmatrix} \hat{\psi}_{xo} \\ \hat{\psi}_{yo}^\dagger \end{pmatrix} = \begin{pmatrix} \gamma & \eta \\ -\eta^* & \gamma^* \end{pmatrix} \begin{pmatrix} \hat{\psi}_{xi} \\ \hat{\psi}_{yi}^\dagger \end{pmatrix} \qquad (62)$$

where again unitarity requires (59), but the lower-row vector components are conjugated, in contrast with the reight-isoclinic counterpart. If we now apply the commutator relations (57) to the output operators of (62), we find for the $x-$component

$$[\hat{\psi}_{xo}, \hat{\psi}_{xo}^\dagger] = (\gamma \hat{\psi}_{xi} + \eta \hat{\psi}_{yi}^\dagger)(\gamma^* \hat{\psi}_{xi}^\dagger + \eta^* \hat{\psi}_{yi}) -$$
$$(\gamma^* \hat{\psi}_{xi}^\dagger + \eta^* \hat{\psi}_{yi})(\gamma \hat{\psi}_{xi} + \eta \hat{\psi}_{yi}^\dagger) \qquad (63)$$

and by using $[\hat{\psi}_{xo}, \hat{\psi}_{yo}] = [\hat{\psi}_{xo}^\dagger, \hat{\psi}_{yo}^\dagger] = 0$, this reduces to

$$1 = [\hat{\psi}_{xo}, \hat{\psi}_{xo}^\dagger] = |\gamma|^2 [\hat{\psi}_{xi}, \hat{\psi}_{xi}^\dagger] - |\eta|^2 [\hat{\psi}_{yi}, \hat{\psi}_{yi}^\dagger] =$$
$$= |\gamma|^2 - |\eta|^2. \qquad (64)$$

The same expression is obtained from the $y-$component. Together with (59) this leads to

$$|\gamma| = 1 \qquad (65)$$

$$|\eta| = 0. \qquad (66)$$

This means that the left-isoclinic transformation is only physically possible if $\eta = 0$. Since, from (38) and (8) we identify $\eta = (i\beta_2 - \beta_3) \sin(\beta)/\beta$, we can see that the condition implies either $\beta_2 = \beta_3 = 0$ or $\sin(\beta) = 0$. The latter are degenerate special cases when the transformation equals the unity matrix or the negative unity matrix.

We can thus conclude the only left-isoclinic rotations that are possible for propagating photons are those where $\beta_2 = \beta_3 = 0$, i.e., the left-isoclinic transformations are "mostly" (for two out of three DOF) nonphysical.

REFERENCES

[1] R. C. Jones, "A new calculus for the treatment of optical systems—I: Description and discussion of the calculus," *J. Opt. Soc. Amer.*, vol. 31, no. 7, pp. 488–493, 1941.
[2] H. Hurwitz Jr. and R. C. Jones, "A new calculus for the treatment of optical systems—II: Proof of three equivalence theorems," *J. Opt. Soc. Amer.*, vol. 31, no. 7, pp. 493–495, 1941.
[3] R. C. Jones, "A new calculus for the treatment of optical systems—III: The Sohncke theory of optical activity," *J. Opt. Soc. Amer.*, vol. 31, no. 7, pp. 500–503, 1941.
[4] R. C. Jones, "A new calculus for the treatment of optical systems—IV.," *J. Opt. Soc. Amer.*, vol. 32, no. 8, pp. 486–493, 1942.
[5] R. C. Jones, "A new calculus for the treatment of optical systems—V. A more general formulation, and description of another calculus," *J. Opt. Soc. Amer.*, vol. 37, no. 2, pp. 107–110, 1947.
[6] R. C. Jones, "A new calculus for the treatment of optical systems—VI. Experimental determination of the matrix," *J. Opt. Soc. Amer.*, vol. 37, no. 2, pp. 110–112, 1947.
[7] R. C. Jones, "A new calculus for the treatment of optical systems—VII. Properties of the N-matrices," *J. Opt. Soc. Amer.*, vol. 38, no. 8, pp. 671–683, 1948.
[8] R. C. Jones, "New calculus for the treatment of optical systems—VIII. Electromagnetic theory," *J. Opt. Soc. Amer.*, vol. 46, no. 2, pp. 126–131, 1956.
[9] H. Mueller, "Memorandumon the polarization optics of the photo-elastic shutter, NDRC project OEMsr-576," National Defence Research Committee, Tech. Rep. no. 2, Nov. 15, 1943.
[10] N. G. Parke III, "Matrix optics," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, NA, USA, 1948.
[11] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," *Trans. Cambridge Phil. Soc.*, vol. 9, pp. 399–416, 1852.
[12] P. Johannisson, M. Sjödin, M. Karlsson, H. Wymeersch, E. Agrell, and P. A. Andrekson, "Modified constant modulus algorithm for polarization-switched QPSK," *Opt. Exp.*, vol. 19, no. 8, pp. 7734–7741, 2011.
[13] H. Takenaka, "A unified formalism for polarization optics by using group theory," *Nouvelle revue d'optique*, vol. 4, 1973.
[14] S. R. Cloude, "Group theory and polarisation algebra," *Optik (Stuttgart)*, vol. 75, no. 1, pp. 26–36, 1986.

[15] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Multilevel coherent optical system based on Stokes parameters modulation," *J. Lightw. Tech.*, vol. 8, no. 7, pp. 1127–1136, Jul. 1990.

[16] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "A novel multilevel coherent optical system: 4-quadrature signaling," *J. Lightw. Tech.*, vol. 9, no. 4, pp. 514–523, Apr. 1991.

[17] R. Cusani, E. Iannone, A. Salonico, and M. Todaro, "An efficient multilevel coherent optical system: M-4Q-QAM," *J. Lightw. Tech.*, vol. 10, no. 6, pp. 777–786, 1992.

[18] (2013). "Rotations in 4-dimensional euclidean space," [Online]. Available: http://en.wikipedia.org/wiki/Rotations_in_4-dimensional_Euclidean_space

[19] M. Karlsson and E. Agrell, "Four-dimensional optimized constellations for coherent optical transmission systems," in *Proc. Europ. Conf. Opt. Comm.*, 2010, Paper WeC3.

[20] M. Karlsson. (2013). "The connection between polarization calculus and four-dimensional rotations," [Online]. Available: http://arxiv.org/abs/1303.1836

[21] J. N. Damask, *Polarization Optics in Telecommunications*. New York, NY, USA: Springer Verlag, 2005.

[22] N. J. Frigo, "A generalized geometrical representation of coupled mode theory," *IEEE J. Quantum Electron.*, vol. 22, no. 11, pp. 2131–2140, 1986.

[23] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Nat. Acad. Sci. USA*, vol. 97, no. 9, pp. 4541–4550, 2000.

[24] E. Agrell and M. Karlsson, "Power-Efficient modulation formats in coherent transmission systems," *J. Lightw. Tech.*, vol. 27, pp. 5115–5126, Nov. 2009.

[25] W. H. Louisell, A. Yariv, and A. E. Siegman, "Quantum fluctuations and noise in parametric processes—I.," *Phys. Rev.*, vol. 124, no. 6, p. 1646, 1961.

[26] N. J. Frigo and F. Bucholtz, "Geometrical representation of optical propagation phase," *J. Lightw. Tech.*, vol. 27, no. 15, pp. 3283–3293, Jan. 2009.

[27] N. J. Frigo, F. Bucholtz, and C. V. McLaughlin, "Polarization in phase modulated optical links: Jones- and generalized stokes-space analysis," *J. Lightw. Technol.*, vol. 31, no. 9, pp. 1503–1511, May 2013. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-31-9-1503

[28] M. Karlsson and E. Agrell, "Which is the most power-efficient modulation format in optical links?" *Opt. Exp.*, vol. 17, no. 13, pp. 10 814–10 819, 2009.

[29] M. Sjödin, P. Johannisson, H. Wymeersch, P. Andrekson, and M. Karlsson, "Comparison of polarization-switched QPSK and polarization-multiplexed QPSK at 30 Gbit/s," *Opt. Exp.*, vol. 19, no. 8, pp. 7839–7846, 2011.

[30] J. Renaudier, P. Serena, A. Bononi, M. Salsi, O. Bertran-Pardo, H. Mardoyan, P. Tran, E. Dutisseuil, G. Charlet, and S. Bigo, "Generation and detection of 28 Gbaud polarization switched-QPSK in WDM long-haul transmission systems," *J. Lightw. Tech.*, vol. 30, no. 9, pp. 1312–1318, 2012.

[31] G. Arfken, *Mathematical Methods for Physicists, Third Edition*. New York, NY, USA: Academic, 1985.

[32] M. Born and E. Wolf, *Principles of optics, 7th ed*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[33] U. Fano, "A Stokes-parameter technique for the treatment of polarization in quantum mechanics," *Phys. Rev.*, vol. 93, pp. 121–123, 1954.

[34] D. L. Falkoff and J. E. MacDonald, "On the Stokes parameters for polarized radiation," *J. Opt. Soc. Amer.*, vol. 41, no. 11, pp. 861–862, 1951.

[35] W. P. Bowen, N. Treps, R. Schnabel, and P. K. Lam, "Experimental demonstration of continuous variable polarization entanglement," *Phys. Rev. Lett.*, vol. 89, no. 25, p. 253601, Dec. 2002.

[36] C. McKinstrie, M. Raymer, S. Radic, and M. Vasilyev, "Quantum mechanics of phase-sensitive amplification in a fiber," *Opt. Comm.*, vol. 257, pp. 146–163, 2006.

[37] N. Korolkova, G. Leuchs, R. Loudon, T. C. Ralph, and C. Silberhorn, "Polarization squeezing and continuous-variable polarization entanglement," *Phys. Rev. A*, vol. 65, p. 052306, Apr. 2002. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevA.65.052306

[38] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*. Cambridge, U.K.: Cambridge Univ. Press, 1995.

Magnus Karlsson received the Ph. D in 1994 from Chalmers University of Technology, Gothenburg, Sweden. Since 1995, he has been with the Photonics Laboratory at Chalmers, first as Assistant Professor and since 2003 as Professor in photonics. He has authored or coauthored more than 280 scientific journal and conference contributions in the areas of nonlinear optics and fiber optic transmission, and cofounded the Chalmers fiber-optic communication research center FORCE in 2010. His research interests include nonlinear optics, fiber optic transmission, coherent communications, and polarization effects.

Dr. Karlsson has served in the technical committee for the Optical Fiber Communication Conference (OFC), and currently serves in the technical program committees for the European Conference of Optical Communication (ECOC), and the Asia Communications and Photonics Conference (ACP). He is an Associate Editor for *Optics Express* since 2010, and was appointed Fellow of the Optical Society of America in 2012.