

PMD fundamentals: Polarization mode dispersion in optical fibers

J. P. Gordon and H. Kogelnik[‡]

Crawford Hill Laboratory, Bell Laboratories, Lucent Technologies, Box 400, Holmdel, NJ 07733

Contributed by H. Kogelnik, February 2, 2000

This paper reviews the fundamental concepts and basic theory of polarization mode dispersion (PMD) in optical fibers. It introduces a unified notation and methodology to link the various views and concepts in Jones space and Stokes space. The discussion includes the relation between Jones vectors and Stokes vectors, rotation matrices, the definition and representation of PMD vectors, the laws of infinitesimal rotation, and the rules for PMD vector concatenation.

1. Introduction

In the more than 15 years since the introduction of the early concepts (1, 2), the fundamentals of polarization mode dispersion (PMD) in optical fibers have become an important body of knowledge basic for the design of high-capacity optical communication systems. PMD effects are linear electromagnetic propagation phenomena occurring in so-called “single-mode” fibers. Despite their name, these fibers support two modes of propagation distinguished by their polarization. Because of optical birefringence in the fiber, the two modes travel with different group velocities, and the random change of this birefringence along the fiber length results in random coupling between the modes. With current practical transmission technology the resulting PMD phenomena lead to pulse distortion and system impairments that limit the transmission capacity of the fiber. Excellent reviews are available (3, 4), covering the practical aspects and applications of PMD concepts to fiber transmission systems and the effects of PMD on nonlinear fiber transmission (5). In this review we aim to complement these surveys and to collect and synthesize the fundamental concepts and theory of PMD, interweaving and linking the principal laws and key formulas that appear scattered in various places in the literature. We will explore the connection between frequency-domain and time-domain analyses and the isomorphic relation between the three-dimensional (3-D) view using real-valued 3-D Stokes vectors and the two-dimensional (2-D) view using complex-valued 2-D Jones vectors. Isomorphic pairings of operators such as these have been widely used elsewhere in physics such as in mechanics (6), in quantum mechanics (7), and even in the unification of quantum theory and general relativity (8). We borrow this methodology for our purposes.

As a preparation, we will first examine the description of polarization of light in the 3-D space of Stokes vectors (Stokes space) and the 2-D space of Jones vectors (Jones space, defined by the transverse coordinates x and y in the laboratory). Pauli spin matrices and spin vectors are the key to connect these two spaces and will be discussed in *Appendix A*, together with the necessary spin vector algebra. Because propagation through the fiber rotates the Stokes vectors of the light, we devote the following section to the rotation matrices in both Stokes and Jones space—i.e., the Lie groups $SU(2)$ and $SO(3)$ (9). Next, we examine the PMD vectors and principal states of polarization (PSPs) used to describe light propagation in randomly birefrin-

gent fibers and the associated differential group delay (DGD). We do this for a variety of representations and establish their connections. Several laws of infinitesimal rotation have been very valuable for visualizing PMD effects in Stokes space. Their connection to PMD vectors will be discussed. Finally, we review the various rules for PMD vector concatenation valuable for analyzing PMD of concatenated pieces of fiber. These rules appear in sum, integral, and differential form.

2. Notation

The unified view to which we aspire requires a unified notation. We attempt to keep our notation simple and transparent while linking to the notation already established as much as possible. The following is an abbreviated listing of our notation:

x, y, z : fiber coordinates; z is the direction of propagation; x and y are the transverse coordinates—i.e., those of Jones space.

$e^{j(\omega_0 t - \beta z)}$: continuous wave traveling in the z direction; j is the imaginary unit, ω_0 the angular carrier frequency, t time, and β the propagation constant.

E, \tilde{E} : electric field vectors; $\tilde{E}(\omega)$ is the Fourier transform of the complex transverse (x, y) electric field vector $E(t)$ and has a complex amplitude e such that

$$\tilde{E} = e|s\rangle. \quad [2.1]$$

The vector of the real electric field is $\text{Re}(Ee^{j\omega_0 t})$.

ω : deviation from the angular carrier frequency ω_0 of the light. The optical frequency is $(\omega_0 + \omega)$.

$|s\rangle$: 2-D complex Jones (column) ket vector, $|s\rangle = \begin{pmatrix} s_x \\ s_y \end{pmatrix}$. The bra $\langle s|$ indicates the corresponding complex conjugate row vector—i.e., $\langle s| = (s_x^*, s_y^*)$. The bra-ket notation is used to distinguish Jones vectors from Stokes vectors. Our Jones vectors are all of unit magnitude—i.e., $\langle s|s\rangle = s_x^* s_x + s_y^* s_y = 1$.

\hat{s} : 3-D Stokes vector of unit length indicating the polarization of the field, and corresponding to $|s\rangle$. The component of \hat{s} are the Stokes parameters.

Abbreviations: PMD, polarization mode dispersion; 3-D and 2-D, three- and two-dimensional; PSP, principal state of polarization; DGD, differential group delay.

[‡]To whom reprint requests should be addressed. E-mail: herwig@lucent.com.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

$$\begin{aligned}
s_1 &= s_x s_x^* - s_y s_y^* \\
s_2 &= s_x s_y^* + s_x^* s_y \\
s_3 &= j(s_x s_y^* - s_x^* s_y).
\end{aligned}
\tag{2.2}$$

By this definition, $s_3 = 1$ for right-circular polarized light ($s_y = js_x$) conforming with the traditional optics definition (10). However, left-circular definitions are also used in the literature (11). We always use the same letter symbols for corresponding Jones and Stokes vectors. Note that a common phase shift of both components of $|s\rangle$ does not change \hat{s} .

I: 2×2 or 3×3 identity matrix. The distinction should be clear from context.

T: 2×2 unitary transmission matrix in Jones space. Relates output to input via

$$|t\rangle = T|s\rangle. \tag{2.3}$$

We use the symbols s and t when necessary for clarity to designate respective input and output quantities, as illustrated in Scheme 1.



Scheme 1. Block diagram of optical fiber under test.

U: 2×2 Jones matrix, with $\det(U) = 1$. Related to *T* by

$$T = e^{-j\phi_0} U, \tag{2.4}$$

where ϕ_0 is the common phase.

R: 3×3 rotation matrix in Stokes space isomorphic to *U*. Relates output to input via

$$\hat{t} = R\hat{s}. \tag{2.5}$$

$\sigma_1, \sigma_2, \sigma_3$: 2×2 Pauli spin matrices, for our purposes defined as

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}. \tag{2.6}$$

$\vec{\sigma}$: Pauli spin vector in Stokes space, $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$.

$\vec{\beta} \cdot \vec{\sigma}$: 2×2 matrix in Jones space, $\vec{\beta} \cdot \vec{\sigma} = \beta_1 \sigma_1 + \beta_2 \sigma_2 + \beta_3 \sigma_3$.

$\vec{\beta}$: 3-D birefringence vector in Stokes space describing local fiber properties.

$\vec{\tau}$: our output PMD vector in Stokes space. Its length τ is the DGD, and its direction is that of the Stokes vector of the slow principal state. In a commonly used notation introduced by Poole (2, 3), the output PMD vector is labeled $\vec{\Omega}$; its length, also the DGD, is labeled $\Delta\tau$. The two PMD vectors $\vec{\tau}$ and $\vec{\Omega}$ are related by inversion through the s_3 axis, for reasons discussed in *Appendix B* in the supplemental data at www.pnas.org.

$\hat{p}, \hat{q}, \hat{r}$: unit Stokes vectors; \hat{p} and \hat{q} are sometimes used to describe the polarization of the slow and fast

principal states, respectively, while \hat{r} is used for a rotation axis.

Subscript ω : indicates differentiation—i.e., $ds/d\omega = s_\omega$.

3. Jones and Stokes Vectors

This preparatory section deals with the representation of polarization in Jones and Stokes spaces and the connection between the two. Throughout this paper we assume that there is no fiber nonlinearity and no polarization-dependent loss, and that the usual loss term of the fiber has been factored out so that we can deal with unitary transmission matrices, *T* and *U*, and 3-D rotation matrices, *R*,

$$TT^\dagger = I; \quad UU^\dagger = I; \quad RR^\dagger = I, \tag{3.1}$$

where the dagger denotes the Hermitian conjugate. Note that *R* has real elements, hence R^\dagger is simply the transpose of *R*. The input and output field components $\vec{E}_s(\omega)$ and $\vec{E}_t(\omega)$ at frequency ω are related by

$$\vec{E}_t = T(\omega)\vec{E}_s. \tag{3.2}$$

After dropping the field magnitudes and using the unit Jones vectors (12) to describe the state of polarization this becomes

$$|t\rangle = T|s\rangle, \tag{3.3}$$

where $|s\rangle$ and $|t\rangle$ include phase and polarization information.

The Pauli spin matrices allow us to write the components s_i of the Stokes vector corresponding to $|s\rangle$ in the compact form (11)

$$s_i = \langle s | \sigma_i | s \rangle. \tag{3.4}$$

With the spin vector $\vec{\sigma}$ (9, 13) the Stokes vector is, simply,

$$\hat{s} = \langle s | \vec{\sigma} | s \rangle. \tag{3.5}$$

In the following we discuss several useful connections between the Jones and Stokes vectors, particularly the projection operator, dot products, and vector superpositions.

As a preparation, recall that the scalar product of two Jones vectors $|p\rangle$ and $|q\rangle$ is

$$\langle q | p \rangle = q_x^* p_x + q_y^* p_y, \tag{3.6}$$

and that their dyadic operator is

$$|p\rangle\langle q| = \begin{pmatrix} p_x q_x^* & p_x q_y^* \\ p_y q_x^* & p_y q_y^* \end{pmatrix}. \tag{3.7}$$

From Eqs. 3.6 and 3.7 one can see that

$$\text{Tr}(|p\rangle\langle q|) = \langle q | p \rangle, \tag{3.8}$$

where Tr stands for the trace operator.

Projection Operator. As shown in *Appendix A*, any complex 2×2 matrix can be expanded in terms of the unit matrix *I* and the three spin matrices (7, 11). For Hermitian matrices, the coefficients are real. We will have several occasions to use this technique, perhaps the most useful being an examination of the projector $|s\rangle\langle s|$ of a Jones vector $|s\rangle$. The result of such an expansion is the surprisingly simple relation

$$|s\rangle\langle s| = \frac{1}{2}(I + \hat{s} \cdot \vec{\sigma}), \tag{3.9}$$

where \hat{s} is the Stokes vector corresponding to $|s\rangle$ and where we use Eqs. A.8, A.9, and 3.8 in the form $\text{Tr}(\sigma_i |s\rangle\langle s|) = \langle s | \sigma_i | s \rangle = s_i$. After multiplication with $|s\rangle$, Eq. 3.9 becomes

$$|s\rangle = \hat{s} \cdot \vec{\sigma} |s\rangle. \tag{3.10}$$

Eq. 3.10 shows another important connection between Jones and Stokes vectors: the Jones vector $|s\rangle$ is an eigenvector of $\hat{s}\cdot\vec{\sigma}$ with unit eigenvalue.

Dot Products. An immediate application of the projector rule 3.9 is the establishment of a connection between the dot (scalar) products of two Jones vectors $|p\rangle$ and $|q\rangle$ and of their associated Stokes vectors \hat{p} and \hat{q} . We use 3.9 with $s = p$ and multiply by $|q\rangle$ from the right and $\langle q|$ from the left, obtaining

$$\langle q|p\rangle\langle p|q\rangle = \frac{1}{2}(1 + \hat{p}\cdot\hat{q}). \quad [3.11]$$

A well-known special case of this connection is a pair of orthogonal Jones vectors with $\langle q|p\rangle = 0$. Their Stokes vectors are antiparallel—i.e., $\hat{q} = -\hat{p}$. In the following we will label such pairs of orthogonal Jones vectors, with antiparallel Stokes vectors, as $|p\rangle$ and $|p-\rangle$, for example. Any such pair forms a complete orthogonal set in Jones space, since we have

$$\langle p-|p\rangle = 0 \quad \text{and} \quad |p\rangle\langle p| + |p-\rangle\langle p-| = I, \quad [3.12]$$

where the second (completeness) relation follows from 3.9.

Let us now look at a pair of input Jones vectors $|p_s\rangle, |q_s\rangle$ and the transmitted vectors $|p_t\rangle, |q_t\rangle$ at the fiber output. They are related by Eq. 3.3. Using Eqs. 3.1 and 3.3, it is easy to show that the input and output dot products are preserved—i.e.,

$$\langle p_t|q_t\rangle = \langle p_s|q_s\rangle. \quad [3.13]$$

When 3.13 is inserted in 3.11, we see that the magnitude of Stokes dot products is also preserved on transmission: i.e., transmission through the fiber is represented by a rotation of the Stokes vectors. This rotation property is, of course, a direct consequence of the orthogonality 3.1 of the matrix R .

Vector Superposition. Here we consider a Jones vector $|s\rangle$ represented as a superposition

$$|s\rangle = a|p\rangle + b|p-\rangle \quad [3.14]$$

of any two orthogonal vectors $|p\rangle$ and $|p-\rangle$. How is this superposition mirrored in Stokes space? For the complex constants a and b , using Eqs. 3.12 and 3.14, we have

$$a = \langle p|s\rangle; \quad b = \langle p-|s\rangle; \quad \langle s|s\rangle = 1 = aa^* + bb^*. \quad [3.15]$$

The Stokes vectors of $|s\rangle, |p\rangle$, and $|p-\rangle$, are \hat{s}, \hat{p} , and $-\hat{p}$. Inserting 3.14 into 3.5, one gets

$$\hat{s} = (aa^* - bb^*)\hat{p} + ab^*\langle p-|\vec{\sigma}|p\rangle + a^*b\langle p|\vec{\sigma}|p-\rangle. \quad [3.16]$$

Combining the eigenvector relation 3.10 and the spin vector relation A.3, we find that $\langle p-|\vec{\sigma}|p\rangle = \langle p-|\vec{\sigma}(\hat{p}\cdot\vec{\sigma})|p\rangle = j\hat{p} \times \langle p-|\vec{\sigma}|p\rangle$ can be expressed by two real-valued Stokes vectors \hat{p}_2 and \hat{p}_3 in the form $\langle p-|\vec{\sigma}|p\rangle = \hat{p}_2 + j\hat{p}_3$, where \hat{p}, \hat{p}_2 , and \hat{p}_3 are a right-handed orthogonal set in Stokes space. Thus Eq. 3.16 becomes

$$\hat{s} = (aa^* - bb^*)\hat{p} + (ab^* + a^*b)\hat{p}_2 + j(ab^* - a^*b)\hat{p}_3, \quad [3.17]$$

representing a 3-D superposition in Stokes space. Note the similarity to Eq. 2.2: The two Jones vectors corresponding to polarization along the x and y axes form just such an orthogonal pair. In general, the vector \hat{p} defines an axis on the Poincaré sphere. Jones vectors with equal power split aa^*/bb^* between $|p\rangle$ and $|p-\rangle$ appear on a circle perpendicular to this axis. The phase difference between the superposition coefficients a and b determines the azimuth on that circle.

4. Rotational Matrix Expressions

We have noted above that the change of the polarization of light on transmission through a fiber can be described as the rotation of its Stokes vector. There are matrix forms that highlight these rotational properties—i.e. the rotation axis \hat{r} and the rotation angle φ . These matrices are basic for an understanding of PMD fundamentals, and they have been used for the measurement of PMD vectors in the laboratory (14). We devote this section to a discussion of several expressions for such matrices in both Jones and Stokes space. First we establish a general connection between the matrices U and R of the two spaces, corresponding to the special unitary and special orthogonal Lie groups $SU(2)$ and $SO(3)$ (9).

Connection Between U and R . To derive this connection we use Eqs. 2.3, 2.4, 2.5, and 3.5 to write two expressions for the output Stokes vector \hat{t}

$$\hat{t} = R\hat{s} = \langle s|R\vec{\sigma}|s\rangle; \quad \hat{t} = \langle t|\vec{\sigma}|t\rangle = \langle s|U^\dagger\vec{\sigma}U|s\rangle. \quad [4.1]$$

Since both expressions are valid for any input state $|s\rangle$, we can extract from Eq. 4.1 the equation

$$R\vec{\sigma} = U^\dagger\vec{\sigma}U, \quad [4.2]$$

which is the desired connection between the matrix R of Stokes space and the Jones matrix U . Note that, even though $\det(U) = 1$, Eq. 4.2 does not determine the algebraic sign of U , which causes some difficulty in the determination of U from experimental results. The direct parallel of this sign uncertainty is the double-valued spinor representation of $SO(3)$ in quantum mechanics (9).

Rotational Forms of the Jones Matrix. An input Stokes vector \hat{s} that is collinear with the rotation axis \hat{r} of R will not be rotated upon transmission through the fiber. The Jones vectors corresponding to \hat{r} and $-\hat{r}$ are $|r\rangle$ and $|r-\rangle$, which must therefore be the eigenvectors of the corresponding U . Since $|r\rangle$ and $|r-\rangle$ constitute a complete orthogonal set of Jones vectors, we can express U in the form $U = \lambda_1|r\rangle\langle r| + \lambda_2|r-\rangle\langle r-|$, where λ_1 and λ_2 are its eigenvalues. Since $UU^\dagger = I$ and $\det(U) = 1$, the eigenvalues of U must be of unit magnitude and their product must be unity. Thus, we can write the Jones matrix U in the rotational form

$$U = e^{-j\varphi/2}|r\rangle\langle r| + e^{j\varphi/2}|r-\rangle\langle r-|. \quad [4.3]$$

Using Eq. 3.9, with r replacing s , we can express Eq. 4.3 in the alternate form

$$U = I \cos(\varphi/2) - j\hat{r}\cdot\vec{\sigma} \sin(\varphi/2), \quad [4.4]$$

where φ is the rotation angle in Stokes space, as we shall see below. This U also has the concise form

$$U = e^{-j(\varphi/2)\hat{r}\cdot\vec{\sigma}} \quad [4.5]$$

(see Eq. A.12). To get from 4.5 to 4.4 entails the power series expansion of the exponential in 4.5 followed by the use of $(\hat{r}\cdot\vec{\sigma})^2 = I$ to reduce the result to 4.4. Recall that $|r\rangle$ and $|r-\rangle$ are the eigenstates of $\hat{r}\cdot\vec{\sigma}$, with eigenvalues 1 and -1 . Warning: the reader should not confuse the eigenstates discussed here with the principal states of PMD to be discussed in the next section. Eigenstates of U have the same polarization at input and output. Only for very special cases, such as phase plates, are these eigenstates the same as the principal states.

Rotational Forms in Stokes Space. Here we use the relation 4.2 to convert the rotational form 4.4 of the Jones matrix U into the

isomorphic form for the matrix R in Stokes space. We insert 4.4 into the right-hand side of 4.2 and obtain

$$U^\dagger \vec{\sigma} U = (\cos \varphi) \vec{\sigma} + (1 - \cos \varphi) \hat{r}(\hat{r} \cdot \vec{\sigma}) + (\sin \varphi) \hat{r} \times \vec{\sigma}, \quad [4.6]$$

having made use of the spin vector rules A.3, A.4, A.7, and A.13 to simplify the resulting expressions. Comparison of 4.6 with 4.2 yields the rotational form of R

$$R = (\cos \varphi) I + (1 - \cos \varphi) \hat{r} \hat{r} + (\sin \varphi) \hat{r} \times, \\ = \hat{r} \hat{r} + (\sin \varphi) \hat{r} \times - (\cos \varphi) (\hat{r} \times) (\hat{r} \times), \quad [4.7]$$

where the 3-D dyadic $\hat{r} \hat{r}$ is the projection operator and $\hat{r} \times$ is the crossproduct operator

$$\hat{r} \hat{r} = \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}; \quad \hat{r} \times = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}. \quad [4.8]$$

One can see from 4.7 that for any Stokes vector \hat{s} , $R\hat{s}$ represents a right-handed rotation of \hat{s} through an angle φ about the direction \hat{r} . The Muller Matrix method for measuring PMD vectors uses this form for extracting rotation axis and angle from measured data for R (14).

The elegant compact expression for U given in Eq. 4.5 has an isomorphic counterpart in Stokes space (15) in

$$R = e^{\varphi(\hat{r} \times)}. \quad [4.9]$$

Here the argument of the exponential is a 3×3 matrix operator. To prove the equivalence of this form with 4.7 one uses the power series expansion of the exponential, and the identities

$$(\hat{r} \times) (\hat{r} \times) = -I + \hat{r} \hat{r}; \quad (\hat{r} \times) (\hat{r} \times) (\hat{r} \times) = -\hat{r} \times \quad [4.10]$$

to collect the terms appearing in this expansion.

Elementary Rotations. Elementary rotations (6) in Stokes space are those that rotate the Stokes vectors around the 1, 2, and 3 axes of the Poincaré sphere. The matrices describing those are special cases of 4.4 and 4.7. We call them U_1, U_2, U_3 and R_1, R_2, R_3 . In 4.4, $\hat{r} \cdot \vec{\sigma}$ reduces to σ_i , yielding

$$U_i = I \cos \varphi/2 - j \sigma_i \sin \varphi/2, \quad i = 1, 2, 3, \quad [4.11]$$

while in 4.7 only $r_i = 1$ is different from zero. The elements of U_i and R_i are listed in Table 1.

Note that U_1 and R_1 describe the rotation caused by a birefringent phase plate with the slow principal axis aligned with the x -axis in Jones space. U_2 and R_2 correspond to a phase plate

Table 1. Elementary rotations

Rotation axis	Jones matrix	Stokes space rotation
1	$U_1 = \begin{pmatrix} e^{-j\varphi/2} & 0 \\ 0 & e^{j\varphi/2} \end{pmatrix}$	$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$
2	$U_2 = \begin{pmatrix} \cos \varphi/2 & -j \sin \varphi/2 \\ -j \sin \varphi/2 & \cos \varphi/2 \end{pmatrix}$	$R_2 = \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix}$
3	$U_3 = \begin{pmatrix} \cos \varphi/2 & -\sin \varphi/2 \\ \sin \varphi/2 & \cos \varphi/2 \end{pmatrix}$	$R_3 = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$

set at 45° angle in Jones space. U_3 and R_3 describe a rotation by $\varphi/2$ around the z -axis.

5. PMD Vectors

PMD phenomena in optical fibers typically used for communications occur because of the presence of birefringence in the fiber. This birefringence changes randomly along the fiber length (3, 4). It stems from asymmetries in the fiber stress and geometry, such as elliptical cross sections, microbends, or microtwists. Often such a fiber is visualized or modeled as a sequence of random birefringent sections whose birefringence axes and magnitude change randomly with z (along the fiber). There are different manifestations of PMD depending on the view taken. In the frequency domain view one sees, for a fixed input polarization, a change with frequency ω of the output polarization. In the time domain one observes a mean time delay of a pulse traversing the fiber which is a function of the polarization of the input pulse. The two phenomena are intimately connected.

There exist special orthogonal pairs of polarization at the input and the output of the fiber called the PSPs. Light launched in a PSP does not change polarization at the output to first order in ω . These PSPs have group delays, τ_g , which are the maximum and minimum mean time delays of the time domain view. The difference between these two delays is called the DGD. Typical mean values of the DGD are 1 to 50 ps for a 500-km long fiber, depending on fiber type. The PMD vector $\vec{\tau}$ describes both the PSPs and the DGD in the fiber. It is a Stokes vector pointing in the direction of the slow PSP with a length equal to the DGD.

Some insight into the PMD problem can be had simply by contemplating a piece of polarization-maintaining fiber. Its PSPs are the polarizations along the principal axes of birefringence of the fiber. In this case the two axes can be treated separately, and in general have different phase shifts ϕ and different group delays $d\phi/d\omega$. One can see that the different values of $\phi(\omega)$ will also produce changes in the output state of polarization as a function of frequency unless the input is launched in one of the PSPs. It may seem surprising that PSPs occur in a fiber exhibiting random birefringence as a function of z . However, they do, as discovered by Poole and Wagner (2). The DGD grows roughly as the square root of the length of fiber, as is characteristic of a random walk problem. This section is devoted to the discussion of the PMD vectors.

We will discuss four mathematical expressions for the PMD vector, one based on Jones matrix eigenanalysis, one on σ -expansion, one on time domain moments, and one on Müller matrices. These expressions correspond to four different yet closely related views of PMD and form the basis for different experimental approaches, such as time-domain and frequency-domain PMD measurements.

Jones Matrix Eigenvector Analysis. The analysis of Poole and Wagner (2) is based on the Jones matrix U and the frequency-domain view, identifying the group delay of a narrow-band pulse propagating through the fiber with the frequency derivative of the phase of the field. As is the case for polarization-maintaining fiber, they found two output polarization states \hat{t} that do not change to first order in frequency. These states have different group delays, and also exhibit the maximum and minimum mean time delays for narrow-band pulses of arbitrary input polarization. Note that our PMD vector $\vec{\tau}$ is defined for right-circular Stokes space, while the widely used PMD vector $\vec{\Omega}$ of Poole *et al.* (2, 3) is defined for left-circular Stokes space. The connection between $\vec{\tau}$ and $\vec{\Omega}$ is detailed in *Appendix B* in the supplemental data at www.pnas.org. We have chosen the symbol $\vec{\tau}$ as appropriate for a quantity with the dimension of time.

According to Eqs. 2.3 and 3.3 we have a transmission equation

$$|t\rangle = e^{-j\phi_0} U |s\rangle, \quad [5.1]$$

relating the input and output Jones vectors $|s\rangle$ and $|t\rangle$. As pulses are described by wave packets with a finite frequency band, we need to consider the frequency dependence of $|t\rangle$. We assume fixed input polarization and phase—i.e., $|s\rangle_\omega = 0$ (hence $\delta_\omega = 0$), as is appropriate for a simple pulse entering the fiber at time zero. By differentiating 5.1 and eliminating $|s\rangle$, we obtain for the change of the output Jones vector

$$|t\rangle_\omega = -j(d\phi_0/d\omega + jU_\omega U^\dagger)|t\rangle. \quad [5.2]$$

Eq. 5.2 tells us that for most input polarizations, the output polarization will change with frequency in first order. The frequency derivative of the common phase ϕ_0 identifies a mean group delay τ_0 common to all polarizations

$$\tau_0 = d\phi_0/d\omega. \quad [5.3]$$

Poole *et al.* (2, 3), in effect, noted that if $|t\rangle$ is either of the two orthogonal eigenstates of the operator $jU_\omega U^\dagger$, then $\hat{t}_\omega = 0$, and so $|t\rangle_\omega$ should be expressible in the form

$$|t\rangle_\omega = -j(d\phi/d\omega)|t\rangle, \quad [5.4]$$

where ϕ is the phase of $|t\rangle$. We show below that the operator $jU_\omega U^\dagger$ is Hermitian and that its trace is zero. Thus its eigenvalues are real and add to zero. We designate them $\tau/2$ and $-\tau/2$. Comparing 5.2 with 5.4, identifying $d\phi/d\omega$ in 5.4 with a group delay τ_g , and using 5.3, we get the two values of group delay

$$\tau_g = \tau_0 \pm \tau/2. \quad [5.5]$$

The eigenstate of $jU_\omega U^\dagger$ associated with the larger value of group delay we designate $|p\rangle$, so that the orthogonal eigenstate having the smaller value of group delay becomes $|p-\rangle$. The slow PSP, $|p\rangle$, thus satisfies the Jones matrix eigenvector equation

$$\frac{1}{2}\tau|p\rangle = jU_\omega U^\dagger|p\rangle. \quad [5.6]$$

According to 5.5 the DGD is τ . Since the determinant of a Hermitian matrix is the product of its eigenvalues, we have $\det(jU_\omega U^\dagger) = -\tau^2/4$. From this, noting that $\det(U) = 1$, we can extract an expression for τ , namely

$$\tau = 2\sqrt{|\det U_\omega|}. \quad [5.7]$$

The PMD vector $\vec{\tau}$ is the Stokes vector of $|p\rangle$ multiplied with the DGD

$$\vec{\tau} = \tau\hat{p}. \quad [5.8]$$

Pauli Spin Matrix Expansion. We now pause to examine the operator $jU_\omega U^\dagger$. Differentiation of the unitary condition 3.1 yields the relation

$$jU_\omega U^\dagger = -jUU^\dagger_\omega. \quad [5.9]$$

Since the Hermitian conjugate of the product of two operators is the product of the Hermitian conjugates of the individual operators taken in the opposite order, we see that the matrix product $jU_\omega U^\dagger$ is Hermitian. We also have $\det(U) = 1$. Consider the expansion

$$U(\omega + d\omega) = U + d\omega U_\omega = (I + d\omega U_\omega U^\dagger)U. \quad [5.10]$$

Since the determinant of the product of two matrices is the product of their determinants, and $\det(U) = 1$, one can see that $\det(U(\omega + d\omega)) = 1$ also only if the trace of $U_\omega U^\dagger$ is zero. This is the argument necessary to arrive at 5.5. Given these properties of $jU_\omega U^\dagger$, and its eigenvalues $\pm\tau/2$, it follows that it can be written in the form

$$jU_\omega U^\dagger = \frac{1}{2} \begin{pmatrix} \tau_1 & \tau_2 - j\tau_3 \\ \tau_2 + j\tau_3 & -\tau_1 \end{pmatrix} = \frac{1}{2} \vec{\tau} \cdot \vec{\sigma}, \quad [5.11]$$

where the last expression is in the spin vector form. To show that 5.11 and 5.8 identify the same vector $\vec{\tau}$, we substitute 5.11 into the eigenvalue equation 5.6, yielding

$$(\vec{\tau} \cdot \vec{\sigma})|p\rangle = \tau|p\rangle. \quad [5.12]$$

In view of 3.10, we observe that $\vec{\tau} = \tau\hat{p}$, as in 5.8, is the solution of 5.12. The σ -expansion of the Jones matrix product of 5.11 is, thus, a very practical tool, because it yields the PMD Stokes vector components as the coefficients of the expansion.

In Section 4 we derived the spin vector expansion of a Jones matrix U in terms of the rotational variables \hat{r} and φ . By substituting 4.4 into 5.11 we can gain an expression for $\vec{\tau}$ in terms of these variables and their frequency derivatives. The result, using the spin vector rules A.5 and A.6 to reduce the intermediate factors, and the identity $\hat{r}_\omega \cdot \hat{r} = 0$, is

$$\vec{\tau} = \varphi_\omega \hat{r} + \sin \varphi \hat{r}_\omega - (1 - \cos \varphi) \hat{r}_\omega \times \hat{r}. \quad [5.13]$$

When U is given in the Caley/Klein form

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \quad [5.14]$$

we obtain the DGD from 5.7 as (2)

$$\tau = 2\sqrt{\alpha_\omega \alpha_\omega^* + \beta_\omega \beta_\omega^*}, \quad [5.15]$$

and for the components τ_i

$$\begin{aligned} \tau_1 &= 2j(\alpha_\omega \alpha_\omega^* + \beta_\omega \beta_\omega^*) \\ \tau_2 &= 2 \operatorname{Im}(\alpha_\omega \beta - \beta_\omega \alpha) \\ \tau_3 &= 2 \operatorname{Re}(\alpha_\omega \beta - \beta_\omega \alpha). \end{aligned} \quad [5.16]$$

Input and Output PMD Vectors. The above discussion has focused on expressions for the PMD vector at the fiber output. There are practical cases where the corresponding PMD vector at the input is needed. This can be obtained by a simple transformation. We distinguish input and output quantities by the subscripts s and t (see Fig. 1). The relation between the PSPs $|p_s\rangle$ and $|p_t\rangle$ is governed by 3.3, which transforms the Jones vectors of the PSPs in the form

$$|p_t\rangle = T|p_s\rangle. \quad [5.17]$$

Matrices M_s operating on vectors at the input transform to output matrix operators $M_t = TM_s T^\dagger$. Applying this transformation to the matrix $\vec{\tau} \cdot \vec{\sigma}$ gives

$$\vec{\tau}_t \cdot \vec{\sigma} = U^\dagger \vec{\tau}_s \cdot \vec{\sigma} U = 2jU^\dagger U_\omega, \quad [5.18]$$

where we have used 5.11 and $T = e^{-j\phi_0 U}$. Thus, the components of the input PMD vector $\vec{\tau}_s$ are the coefficients of the σ -expansion of the matrix product $U^\dagger U_\omega$. As the PMD vectors are Stokes vectors they are transformed by the Müller rotation matrix R isomorphic to U (see Eq. 4.2) as

$$\vec{\tau}_t = R\vec{\tau}_s. \quad [5.19]$$

Moments and Mean Signal Delay. Moments are widely used to characterize pulse transmission, and they have also been applied to PMD phenomena (15–18). Here we describe their use for determining and defining PMD vectors. One should note that moments are very appropriate for the description of signal delays in systems involving two (or more) modes of propagation such as

those involving PMD effects, whereas concepts such as group velocity and group delay have a strict definition for an individual mode only. The PMD phenomenon can split an input pulse into two or more pulses at the fiber output, leading to polarization-dependent pulse shapes. To maintain our awareness of this complexity we use the word “signal” for the output rather than “pulse.”

To prepare for the discussion of moments we represent a signal at fiber position z by the complex field vector $\mathbf{E}(z, t)$ in the time domain and by its Fourier transform $\tilde{\mathbf{E}}(z, \omega)$ in the frequency domain. The tilde distinguishes between them.

$$\mathbf{E} = \frac{1}{\sqrt{2\pi}} \int d\omega \tilde{\mathbf{E}} e^{j\omega t}, \quad \tilde{\mathbf{E}} = \frac{1}{\sqrt{2\pi}} \int dt \mathbf{E} e^{-j\omega t}, \quad [5.20]$$

where all integrals extend from $-\infty$ to $+\infty$. Let the fields be normalized so that the energy in the pulse is

$$W = \int dt \mathbf{E}^\dagger \mathbf{E} = \int d\omega \tilde{\mathbf{E}}^\dagger \tilde{\mathbf{E}}, \quad [5.21]$$

where the right equation is Parseval's theorem.

The mean time $t_g(z)$ at which a signal passes location z is defined by its center of gravity or first moment W_1

$$t_g = W_1(z)/W. \quad [5.22]$$

Using the convolution theorem, the first moment can be written in either the time domain or the frequency domain

$$W_1 = \int dt t \cdot \mathbf{E}^\dagger \mathbf{E} = j \int d\omega \tilde{\mathbf{E}}^\dagger \tilde{\mathbf{E}} \omega. \quad [5.23]$$

We now use the transmission and Jones matrices to express the output field $\tilde{\mathbf{E}}_t$ in terms of the input $\tilde{\mathbf{E}}_s$, $\tilde{\mathbf{E}}_t = T \tilde{\mathbf{E}}_s$, and use the derivatives,

$$\tilde{\mathbf{E}}_{t\omega} = T_\omega \tilde{\mathbf{E}}_s + T \tilde{\mathbf{E}}_{s\omega}, \quad [5.24]$$

$$T_\omega = (-j\tau_0 U + U_\omega) e^{-j\phi_0} = -jU(\tau_0 I + \frac{1}{2} \tilde{\tau}_s \cdot \vec{\sigma}) e^{-j\phi_0}, \quad [5.25]$$

to substitute in 5.22 and 5.23. Defining the mean signal delay τ_g as the difference between the mean arrival times at output (t) and input (s), one gets

$$\tau_g = t_{gt} - t_{gs} = [W_{1t} - W_{1s}]/W = \frac{1}{W} \int d\omega \tilde{\mathbf{E}}_s^\dagger (\tau_0 + jU^\dagger U_\omega) \tilde{\mathbf{E}}_s, \quad [5.26]$$

where $\tau_0 = d\phi_0/d\omega$ as before. Substituting the σ -expansion 5.18 for $jU^\dagger U_\omega = \frac{1}{2} \tilde{\tau}_s \cdot \vec{\sigma}$, the mean signal delay can be written in the form

$$\tau_g = \frac{1}{W} \int d\omega \tilde{\mathbf{E}}_s^\dagger (\tau_0 + \frac{1}{2} \tilde{\tau}_s \cdot \vec{\sigma}) \tilde{\mathbf{E}}_s. \quad [5.27]$$

We proceed by representing the input field $\tilde{\mathbf{E}}_s$ by its complex amplitude e , its unit Jones vector $|s\rangle$, and the corresponding Stokes vector \hat{s} (see Eq. 3.5),

$$\tilde{\mathbf{E}}_s = e(\omega)|s\rangle; \quad \hat{s}(\omega) = \langle s|\vec{\sigma}|s\rangle, \quad [5.28]$$

where the signal energy becomes $W = \int d\omega e^* e$. Inserting this above, we obtain (15, 18) for the mean signal delay

$$\tau_g = \frac{1}{W} \int d\omega e(\omega) e^*(\omega) (\tau_0 + \frac{1}{2} \tilde{\tau}_s \cdot \hat{s}) = \langle \tau_0(\omega) \rangle + \frac{1}{2} \langle \tilde{\tau}_s(\omega) \cdot \hat{s}(\omega) \rangle. \quad [5.29]$$

where $\tilde{\tau}_s$ is the input PMD vector as before. This expresses τ_g as the spectral mean (denoted by the $\langle \rangle$ bracket) of the spectral density $e^* e$ weighing $\tau_0(\omega)$ and the dot product between input PMD vector $\tilde{\tau}_s(\omega)$ and the input Stokes vector. This expression for the mean signal delay is valid for any input pulse shape and spectral variation of its polarization.

Specializing to frequency-independent input polarization \hat{s} and a narrow pulse spectrum such that τ_0 and $\tilde{\tau}_s$ can be regarded constant over the signal band, 5.29 simplifies to (16)

$$\tau_g = \tau_0 + \frac{1}{2} \tilde{\tau}_s \cdot \hat{s}. \quad [5.30]$$

This agrees with 5.5, which was derived for the special case of input polarizations aligned with the PSPs. Eq. 5.30 gives the mean signal delay for any input polarization. Inputs at the PSPs are seen to lead to the maximum and minimum delays τ_g . In fact, the mean signal delay τ_g can be simply interpreted as the power-weighted average of the two PSP delays. To see this, consider the representation of the input polarization as the superposition 3.14, $|s\rangle = a|p\rangle + b|p-\rangle$, of the PSPs. These delays are $(\tau_0 \pm \tau/2)$ and the power-weighted average delay is

$$\tau_g = aa^*(\tau_0 + \tau/2) + bb^*(\tau_0 - \tau/2). \quad [5.31]$$

We can now use the dot-product rule 3.11 to show the equivalence of 5.31 and 5.30. The simple form of 5.30 allows a determination of the PMD vector $\tilde{\tau}_s$ in the time domain (19).

The higher moments $W_n = \int dt t^n \mathbf{E}^\dagger \mathbf{E}$ can be used to determine further detail on the properties of transmitted pulses. The second moment

$$W_2(z) = \int dt t^2 \mathbf{E}^\dagger \mathbf{E}(z) = \int d\omega \tilde{\mathbf{E}}_\omega^\dagger \tilde{\mathbf{E}}_\omega \omega^2 \quad [5.32]$$

provides information about rms pulse spreading. Proceeding along similar lines as above and the substitution of 5.24, we obtain

$$W_{2t} - W_{2s} = \int d\omega \{ \tilde{\mathbf{E}}_s^\dagger T_\omega^\dagger T_\omega \tilde{\mathbf{E}}_s + (\tilde{\mathbf{E}}_s^\dagger T_\omega^\dagger T \tilde{\mathbf{E}}_{s\omega} + cc) \}. \quad [5.33]$$

From 5.25 it follows that

$$T_\omega^\dagger T_\omega = (\tau_0 I + \frac{1}{2} \tilde{\tau}_s \cdot \vec{\sigma})^2 = \tau_0^2 I + \tau_0 \tilde{\tau}_s \cdot \vec{\sigma} + \frac{1}{4} \tau_s^2 I, \quad [5.34]$$

where we have used spin-vector rule A.8. To simplify 5.33 further we use 5.28, assume frequency-independent input polarization $|s\rangle$, and characterize the complex input amplitude $e(\omega)$ by its real amplitude $a(\omega)$ and phase $\delta(\omega)$

$$e = a e^{-j\delta}. \quad [5.35]$$

The result is

$$W_{2t} - W_{2s} = \int d\omega e^* e \{ (\tau_0 \hat{s} + \frac{1}{2} \tilde{\tau}_s)^2 + 2\delta_\omega (\tau_0 + \frac{1}{2} \tilde{\tau}_s \cdot \hat{s}) \}. \quad [5.36]$$

For the special case of $\delta_\omega = 0$, this form agrees with the result of Karlsson (15). This special case assumes freedom from frequency chirp in the optical pulse and a symmetric pulse shape, or, more precisely, that $\mathbf{E}^*(-t) = \mathbf{E}(t)$.

Müller Matrix Expression for $\vec{\tau}$. A fourth way of determining the PMD vector uses Stokes space and the Müller (rotation) matrix R isomorphic to U . The background for this will be given in Section 6. For completeness we show the result here. It is

$$\vec{\tau} \times = R_\omega R^\dagger, \quad [5.37]$$

where \times denotes the cross-product operator as in 4.8.

This method relies on the determination of the rotation matrix R and its derivative in the frequency domain.

6. Laws of Infinitesimal Rotation

The laws of rotation assume a particularly simple form when the rotations are infinitesimally small—e.g., for a small change in fiber length or frequency. These are well known in mechanics (6) and are widely used in the related phenomena of PMD, coupled modes of propagation, and two-waveguide systems (3, 20–24). Here these laws allow simple geometrical interpretations in Stokes space (on the Poincaré sphere). We shall discuss infinitesimal laws for the birefringence vector, the PMD vector, and the dynamical PMD equation.

Birefringence Vector. Consider the change of the polarization $|s\rangle(z)$ of light at fiber location z due to a small length addition dz of fiber. This change is influenced by the fiber's local birefringence characterized by its effective relative dielectric tensor $\varepsilon(z)$ —i.e., a cross-sectional average of the fiber characteristics for the fiber mode of interest. The change is governed by the wave equation for a spectral component of the effective transverse field vector $\vec{E}(z)$ of the mode,

$$\frac{d^2 \vec{E}}{dz^2} + \varepsilon k_0^2 \vec{E} = 0, \quad [6.1]$$

where $k_0 = 2\pi/\lambda_0$ is the propagation constant of free space, and λ_0 is the free-space wavelength. To proceed, we use a σ -expansion of the ε -tensor of the form

$$\varepsilon k_0^2 = \beta_0^2 I + \beta_0 \vec{\beta} \cdot \vec{\sigma} = \beta_0^2 I + \beta_0 \begin{pmatrix} \beta_1 & \beta_2 - j\beta_3 \\ \beta_2 + j\beta_3 & -\beta_1 \end{pmatrix}, \quad [6.2]$$

where β_0 is the common propagation constant. The coefficients β_i of the expansion are the components of the local birefringence vector $\vec{\beta}(z)$ in Stokes space. This vector has the character of a propagation constant and has been a useful tool in describing birefringence in PMD phenomena (3, 22, 25) (see Appendix B in supplemental data at www.pnas.org for the connection of $\vec{\beta}$ to the vector \vec{W} used in part of the literature). In addition to this expansion we use an adiabatic approximation assuming that the polarization, $|s\rangle(z)$, and $\varepsilon(z)$ all vary slowly with z , and by setting

$$\vec{E} = e^{-j\beta_0 z} |s\rangle, \quad [6.3]$$

where $|s\rangle$ includes a slowly varying phase. We continue by inserting 6.2 and 6.3 into 6.1. We drop the $d^2|s\rangle/dz^2$ term in accordance with the adiabatic assumption and find the adiabatic wave equation for the Jones vector $|s\rangle(z)$,

$$\frac{d|s\rangle}{dz} + \frac{1}{2} j \vec{\beta} \cdot \vec{\sigma} |s\rangle = 0. \quad [6.4]$$

The right-hand side of 6.2 shows the form of the fiber's dielectric tensor as expressed by the components β_i of the birefringence vector $\vec{\beta}$. It is apparent that optical activity (circular birefringence) is included in this description. For the special case of linear birefringence aligned with the fiber's x -axis we have $\beta_2 = \beta_3 = 0$ and $\beta_1 = \Delta n k_0$, where Δn is the differential refractive index. The corresponding ordinary and extraordinary

indices are

$$n = n_0 \pm \Delta n / 2. \quad [6.5]$$

There is a close relationship between the birefringence vector $\vec{\beta}$ and the concept of a local normal mode used in the theory of optical waveguides for the analysis of guides whose characteristics change with length z , such as our $\varepsilon(z)$. At a given location z_0 , the local normal modes are defined as the modes of a uniform guide with the characteristics of location z_0 —i.e., a guide with uniform $\varepsilon(z) = \varepsilon(z_0)$ in our case. Inserting this assumption into 6.1, we obtain the field of the two local normal modes $\vec{E}(z) = e^{-j\beta_M z} |s_M\rangle$ whose polarization $|s_M\rangle$ does not change along the length. Because of 3.10 we know that $|s_M\rangle$ must be an eigenvector of $\vec{\beta} \cdot \vec{\sigma}$. The polarization of the local normal modes is, thus, described by the Stokes vectors $\pm \hat{\beta}$. The propagation constant β_M of the modes depends on their polarization in the form

$$\beta_M = \beta_0 + \frac{1}{2} \vec{\beta} \cdot \hat{s}_M = \beta_0 \pm \beta, \quad [6.6]$$

assuming that $\beta_0 \gg \beta$.

We now return to the case of interest where $\vec{\beta}(z)$ changes along the fiber length z . To describe the change of polarization with z we differentiate the Stokes vector $\hat{s} = \langle s | \vec{\sigma} | s \rangle$ and obtain

$$\frac{d\hat{s}}{dz} = \frac{d\langle s | \vec{\sigma} | s \rangle}{dz} + \langle s | \vec{\sigma} \frac{d|s\rangle}{dz}. \quad [6.7]$$

Finally, we combine 6.4 and 6.7, use the spin vector rules A.6, A.7, and A.13, and obtain the law of infinitesimal rotation for birefringence

$$\frac{d\hat{s}}{dz} = \vec{\beta} \times \hat{s}. \quad [6.8]$$

Integration of 6.8 yields a formal expression for the Müller matrix R of the fiber as in 2.5.

PMD Vector $\vec{\tau}$. Here we consider the change of polarization at the fiber output due to a small change in frequency ω . The polarization at the input is held constant. We start with Eq. 5.2 written as

$$|s\rangle_\omega = -j(\tau_0 + jU_\omega U^\dagger) |s\rangle \quad [6.9]$$

for the change of the output polarization $|s\rangle_\omega$, where the subscript ω indicates differentiation. The differential of the corresponding Stokes vector is

$$\hat{s}_\omega = \langle s | \vec{\sigma} | s \rangle + \langle s | \vec{\sigma} (|s\rangle)_\omega. \quad [6.10]$$

Finally, we combine 6.9 and 6.10, using the σ -expansion 5.11 for the product $U_\omega U^\dagger$, apply spin vector rule A.6, A.7, and obtain the law of infinitesimal rotation

$$\hat{s}_\omega = \vec{\tau} \times \hat{s}. \quad [6.11]$$

The geometrical interpretation of this simple law is a rotation of the output Stokes vector on the Poincaré sphere as ω changes. The rotation axis is the PSP \hat{p} and the rotation rate is the DGD τ .

The infinitesimal rotation law 6.11 allows us to express the PMD vector in terms of the rotation (Müller) matrix R that relates the input and output Stokes vectors $\hat{s}(0)$ and $\hat{s}(z)$ by $\hat{s}(z) = R\hat{s}(0)$. We differentiate this relation [while keeping the input $\hat{s}(0)$ fixed] and obtain

$$\hat{s}_\omega(z) = R_\omega \hat{s}(0) = R_\omega R^\dagger \hat{s}(z). \quad [6.12]$$

Comparison of 6.11 and 6.12 yields the operator relationship

$$\vec{\tau} \times = R_\omega R^\dagger, \quad [6.13]$$

which we have already listed in 5.37.

The Dynamical PMD Equation. The dynamical PMD equation results from a combination of the infinitesimal rotation laws for PMD and birefringence (22). It is the basis for the statistical theory of PMD (23). Poole's derivation starts by differentiation of 6.8 with respect to ω and of 6.11 with respect to z . The results are combined by eliminating $\partial^2 \hat{s} / \partial z \partial \omega$, and simplified by the relation $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$, yielding

$$\frac{\partial \vec{\tau}}{\partial z} \times \hat{s} = \vec{\beta}_\omega \times \hat{s} + (\vec{\beta} \times \vec{\tau}) \times \hat{s}. \quad [6.14]$$

As this is valid for any \hat{s} , we can extract the included equation

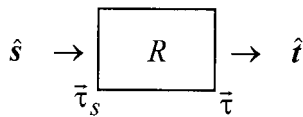
$$\frac{\partial \vec{\tau}}{\partial z} = \frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\tau}. \quad [6.15]$$

This is the dynamical PMD equation describing the evolution of the PMD vector with distance. In the next section we shall see that it is intimately connected to the powerful PMD concatenation rules.

7. PMD Vector Concatenation Rules

The PMD vector concatenation rules (13, 16, 22, 23, 26) are a powerful set of simple tools that allow the determination of the PMD vector of an assembly of concatenated fiber sections when the PMD vectors of the individual sections are known. Among their uses is the analysis of the evolution of the PMD vector with fiber length (26), statistical PMD modeling (23), PMD simulation, and the design of multisection PMD compensators. They appear in a variety of related forms, including the sum, differential, and integral formulations for both the first-order and second-order PMD vectors. As a preparation for the discussion of these formulations, we shall first review the relations between the output and input PMD vectors of a single fiber section.

Transformation of PMD Vectors. Consider a fiber section with rotation matrix R and input and output Stokes vectors \hat{s} and \hat{t} as shown in Scheme 2. The corresponding PMD vectors at input



Scheme 2.

and output are $\vec{\tau}_s$ and $\vec{\tau}$. They are related by

$$\vec{\tau} = R \vec{\tau}_s, \quad [7.1]$$

as discussed earlier (5.19). The expression for the output vector $\vec{\tau}$ in terms of R was already given in 6.13; it is

$$\vec{\tau} \times = R_\omega R^\dagger. \quad [7.2]$$

As cross-product operators transform like matrices from output to input we obtain the expression for the input PMD vector as

$$\vec{\tau}_s \times = R^\dagger (\vec{\tau} \times) R = R^\dagger R_\omega. \quad [7.3]$$

When higher-order PMD effects are considered, $\vec{\tau}$ is usually called the “first-order” PMD vector, its frequency derivative $\vec{\tau}_\omega$ is the “second-order” PMD vector, the second derivative $\vec{\tau}_{\omega\omega}$ is the “third-order” PMD vector, etc. The input/output transformation for the second-order PMD vectors $\vec{\tau}_{s\omega}$ and $\vec{\tau}_\omega$ is derived by differentiating 7.1 and substituting 7.2, giving

$$\vec{\tau}_\omega = R_\omega \vec{\tau}_s + R \vec{\tau}_{s\omega} = R_\omega R^\dagger \vec{\tau} + R \vec{\tau}_{s\omega} = \vec{\tau} \times \vec{\tau} + R \vec{\tau}_{s\omega},$$

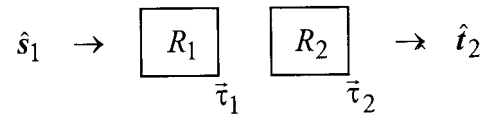
with the result

$$\vec{\tau}_\omega = R \vec{\tau}_{s\omega}. \quad [7.4]$$

This shows that the input and output second-order PMD vectors transform the same way as the first-order vectors of the section. For the third-order PMD vectors one gets the somewhat more complicated relationship

$$\vec{\tau}_{\omega\omega} = R \vec{\tau}_{s\omega\omega} + \vec{\tau} \times \vec{\tau}_\omega. \quad [7.5]$$

PMD of Two Concatenated Sections. Here, we consider two concatenated fiber sections with rotation matrices R_1 and R_2 as shown in Scheme 3. The output PMD vectors of each of the individual sections are $\vec{\tau}_1$ and $\vec{\tau}_2$.



Scheme 3.

Our goal is to determine the output PMD vector of the combined assembly. The rotation matrix R of the two-section assembly is the matrix product

$$R = R_2 R_1. \quad [7.6]$$

Combining this with 7.2, we find for the PMD vector $\vec{\tau}$ of the assembly that

$$\vec{\tau} \times = R_\omega R^\dagger = R_{2\omega} R_1 R_1^\dagger R_2^\dagger + R_2 R_{1\omega} R_1^\dagger R_2^\dagger. \quad [7.7]$$

Now we recall that $R_1 R_1^\dagger = I$, apply 7.2 for the individual sections as well as a transformation of matrix operators as in 5.18

$$(R_2 \vec{\tau}_1) \times = R_2 (\vec{\tau}_1 \times) R_2^\dagger \quad [7.8]$$

to simplify 7.7 into the form

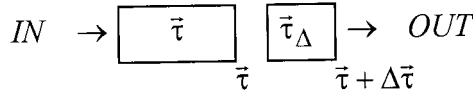
$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1. \quad [7.9]$$

This is the basic concatenation rule. It can be generalized to multiple sections as well as differentially small sections. It can also be transformed to the input of the fiber, or, for that matter to any desired fiber cross section. The rule is very similar to that for impedances of a transmission line: to get the PMD vector of an assembly, transform the PMD vectors of each individual section to a common reference cross section and take the sum of all those vectors.

The concatenation rule for the second-order PMD vector $\vec{\tau}_\omega$ follows from 7.9 by differentiation and substitutions similar to those above. One obtains

$$\vec{\tau}_\omega = \vec{\tau}_{2\omega} + R_2 \vec{\tau}_{1\omega} + \vec{\tau}_2 \times \vec{\tau}. \quad [7.10]$$

The Differential Concatenation Rule. Here, we apply the two-section concatenation rule to a long piece of fiber with output PMD vector $\vec{\tau}$ and a differentially small fiber addition of length Δz as shown in Scheme 4.



Scheme 4.

We want to answer the question: How does $\vec{\tau}(z)$ change due to the differential addition of length Δz ? The answer to this question is already provided by the dynamical PMD equation 6.15. Our discussion will show that the latter is identical to the differential form of the concatenation rule. It will establish a link between this rule and the laws of infinitesimal rotation treated in Section 6. For this purpose we need to determine the rotation matrix R_Δ and the PMD vector $\vec{\tau}_\Delta$ of the differential addition Δz . We use the local birefringence vector $\vec{\beta}$ of the addition and the law of infinitesimal rotation (6.8) and get for the change $\Delta \hat{s}$ of the Stokes vector \hat{s} entering the element

$$\Delta \hat{s} = \Delta z \vec{\beta} \times \hat{s}. \quad [7.11]$$

In terms of R_Δ , the output Stokes vector $\hat{s} + \Delta \hat{s}$ can be expressed as

$$\hat{s} + \Delta \hat{s} = R_\Delta \hat{s}. \quad [7.12]$$

Comparison of 7.11 and 7.12 yields the desired expression for the rotation matrix of the differential addition

$$R_\Delta = I + \Delta z \vec{\beta} \times. \quad [7.13]$$

To obtain its PMD vector $\vec{\tau}_\Delta$ we differentiate 7.13, apply 7.2, and drop terms of second order in Δz with the result

$$\vec{\tau}_\Delta = \Delta z \vec{\beta}_\omega. \quad [7.14]$$

The PMD concatenation rule 7.9 for the addition of the differential fiber element is

$$\vec{\tau} + \Delta \vec{\tau} = \vec{\tau}_\Delta + R_\Delta \vec{\tau}. \quad [7.15]$$

Now we can insert 7.13 and 7.14, divide by Δz , and get

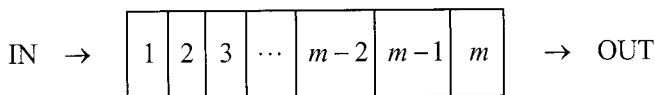
$$\Delta \vec{\tau} / \Delta z = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau}, \quad [7.16]$$

which is the dynamical PMD equation 6.15 on transition to infinitesimal Δz .

The differential concatenation rule for the second-order PMD vector is obtained along similar lines or by differentiation of 7.16. It is

$$d\vec{\tau}_\omega / dz = \vec{\beta}_{\omega\omega} + \vec{\beta} \times \vec{\tau}_\omega + \vec{\beta}_\omega \times \vec{\tau}. \quad [7.17]$$

Concatenation Rules for Many Sections. Consider now a fiber consisting of m individual sections, which is sketched in Scheme 5.



Scheme 5.

The rotation matrices R_n and the output PMD vectors $\vec{\tau}_n$ of each section are considered known. The PMD vectors $\vec{\tau}$ and $\vec{\tau}_\omega$ of the combined assembly can be determined by repeated application of the two-section rules 7.9 and 7.10. The resulting expressions simplify if we define the rotation matrix of the last $m - n + 1$ sections as

$$R(m, n) = R_m R_{m-1} \cdots R_n, \quad [7.18]$$

where $R(m, m) = R_m$ and $R(m, m + 1) = I$, and the output PMD vector of the first n sections as $\vec{\tau}(n)$.

Using these definitions, the sum rules for the first- and second-order PMD vectors of the assembly are

$$\vec{\tau} = \sum_{n=1}^m R(m, n + 1) \vec{\tau}_n, \quad [7.19]$$

and

$$\vec{\tau}_\omega = \sum_{n=1}^m R(m, n + 1) \{ \vec{\tau}_{n\omega} + \vec{\tau}_n \times \vec{\tau}(n) \}. \quad [7.20]$$

Integral Form of the Concatenation Rule. The above sums turn into integrals when one makes the transition to infinitesimal sections of length dz and uses the birefringence vector $\vec{\beta}(z)$ to express the PMD vector $\vec{\tau}(z)$ of an infinitesimal section

$$\vec{\tau}(z) = \Delta z \vec{\beta}_\omega(z) \quad [7.21]$$

as in 7.14. For a fiber of length L , we express the rotation from z to L as $R(L, z)$ corresponding to 7.18 and get the integral formulations of the output PMD vectors at L as

$$\vec{\tau}(L) = \int_0^L dz R(L, z) \vec{\beta}_\omega(z) \quad [7.22]$$

and

$$\vec{\tau}_\omega(L) = \int_0^L dz R(L, z) \{ \vec{\beta}_{\omega\omega} + \vec{\beta}_\omega \times \vec{\tau}(z, 0) \}, \quad [7.23]$$

where $\vec{\tau}(z, 0)$ is the PMD vector for the piece of the fiber extending from 0 to z .

Appendix A: Pauli Spin Matrices and Spin Vectors

In this section we give a brief review of the properties of the Pauli spin matrices defined in Section 2 and of some key formulas for the spin vectors $\vec{\sigma}$ which will be used in the main body of the paper (9). The spin vector notation provides a convenient way of dealing with the complex 2×2 matrices that occur in the Jones matrix analysis of lightwave transmission. In addition it simplifies the connection between the complex 2-D Jones vectors and their corresponding real 3-D Stokes vectors. The spin matrices are Hermitian and unitary—i.e.,

$$\sigma_i = \sigma_i^\dagger \quad \text{and} \quad \sigma_i^\dagger = \sigma_i^{-1} \quad [A.1]$$

and have zero trace, $\text{Tr} \sigma_i = 0$. They obey the well-known multiplication rules

$$\sigma_i^2 = I; \quad \sigma_i \sigma_j = -\sigma_j \sigma_i = j \sigma_k, \quad [A.2]$$

where the indices (i, j, k) can be any cyclic permutation of $(1, 2, 3)$. The spin vector notation consists of treating the three spin matrices of Jones space as the three components of a vector in

Stokes space. For example, the dot product $\vec{a} \cdot \vec{\sigma} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$ is a 2×2 matrix in Jones space. Thus, the spin vector $\vec{\sigma}$ contains elements of both spaces, and is useful in examining their connection. By using rule **A.2**, any function of the σ matrices that can be expanded in a power series can be reduced to an expression linear in the σ matrices and the unit matrix I . Useful examples of such reductions, expressed in spin vector notation, are

$$\vec{\sigma}(\vec{a} \cdot \vec{\sigma}) = \vec{a}I + j\vec{a} \times \vec{\sigma}, \quad [\text{A.3}]$$

$$(\vec{a} \cdot \vec{\sigma})\vec{\sigma} = \vec{a}I - j\vec{a} \times \vec{\sigma}, \quad [\text{A.4}]$$

$$(\vec{a} \cdot \vec{\sigma})(\vec{a} \cdot \vec{\sigma}) = a^2I, \quad [\text{A.5}]$$

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b})I + j(\vec{a} \times \vec{b}) \cdot \vec{\sigma}, \quad [\text{A.6}]$$

$$(\vec{a} \cdot \vec{\sigma})\vec{\sigma}(\vec{a} \cdot \vec{\sigma}) = 2\vec{a}(\vec{a} \cdot \vec{\sigma}) - a^2\vec{\sigma}, \quad [\text{A.7}]$$

where I is the 2×2 unit matrix and \vec{a} and \vec{b} are any vectors in Stokes space, and a is the length of the vector \vec{a} . Although limited space does not permit us to write out the detailed derivations for the above, note that **A.4** is the complex conjugate of **A.3** and that **A.5** is a special case of **A.6**. Note also that $\vec{a} \cdot \vec{\sigma}$ does not commute with $\vec{b} \cdot \vec{\sigma}$ unless $\vec{a} \times \vec{b} = 0$.

In general, any 2×2 matrix M may be expanded in the form

$$\begin{aligned} M &= \begin{pmatrix} a_0 + a_1 & a_2 - ja_3 \\ a_2 + ja_3 & a_0 - a_1 \end{pmatrix} \\ &= a_0I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \\ &= a_0I + \vec{a} \cdot \vec{\sigma} \end{aligned} \quad [\text{A.8}]$$

with coefficients a_i given by

$$a_0 = \frac{1}{2}\text{Tr}(M); \quad a_i = \frac{1}{2}\text{Tr}(\sigma_i M), \quad [\text{A.9}]$$

where to show **A.9** we use **A.2** along with $\text{Tr}(I) = 2$ and $\text{Tr}(\sigma_i) = 0$. In **A.8** we have used the spin vector $\vec{\sigma}$ to write the expansion in compact notation. An illustrative example of **A.8** is

$$\begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} = \frac{\alpha}{2}(\sigma_2 + j\sigma_3), \quad [\text{A.10}]$$

which is intended to demonstrate its universality. Note that the right-hand sides of **A.3** through **A.7** are matrices of the form of **A.8**.

If M is Hermitian ($M^\dagger = M$), then the a coefficients in the expansion **A.8** are real. If M is unitary ($M^\dagger = M^{-1}$), then it can be expressed in the form

$$M = \exp(-jH) = I + (-jH) + (1/2)(-jH)^2 + \cdots \quad [\text{A.11}]$$

where H is a 2×2 Hermitian matrix, and the right-hand side is the usual power series expansion of the exponent function (12). If we express the matrix H of **A.11** in the form **A.8** with real coefficients, we get for the unitary matrix

$$\begin{aligned} M &= \exp(-ja_0)\exp(-j\vec{a} \cdot \vec{\sigma}) \\ &= \exp(-ja_0)[I \cos(a) - j\hat{a} \cdot \vec{\sigma} \sin(a)], \end{aligned} \quad [\text{A.12}]$$

where $a\hat{a} = \vec{a}$. In **A.12** the first form results because the unit matrix I commutes with the others, and the last form results from expanding $\exp(-j\vec{a} \cdot \vec{\sigma})$ in a power series and using **A.5** to reduce the result to a form linear in I and the Pauli matrices. It is easy to show directly that if M has the form **A.12**, then $M^\dagger M = I$. Note that if the trace of any matrix M is zero, then the coefficient of I must be zero. Also, it is not difficult to show from **A.12** that if a unitary matrix M , such as U in the text, satisfies $\det(M) = 1$, then the coefficient $\exp(-ja_0)$ is restricted to the values 1 or -1 .

A great practical value of the spin vector notation is that $\vec{\sigma}$ can be converted into any Stokes vector (\hat{s}) by forming a quadratic form as in **3.5**. This simplicity carries to more complex expressions such as in

$$\langle s | \vec{a} \cdot \vec{\sigma} | s \rangle = \vec{a} \cdot \langle s | \vec{\sigma} | s \rangle = \vec{a} \cdot \hat{s} \quad [\text{A.13}]$$

$$\langle s | \vec{a} \times \vec{\sigma} | s \rangle = \vec{a} \times \langle s | \vec{\sigma} | s \rangle = \vec{a} \times \hat{s}, \text{ or}$$

$$\langle s | R \vec{\sigma} | s \rangle = R \langle s | \vec{\sigma} | s \rangle = R \hat{s},$$

where R is a 3×3 matrix. In the sense of **A.13** it is often appropriate to regard $\vec{\sigma}$ as an arbitrary vector in Stokes space. For example, in the text we often use the fact that if $\vec{a} \cdot \vec{\sigma} = \vec{b} \cdot \vec{\sigma}$, it follows that $\vec{a} = \vec{b}$. This result also follows from **A.8** and **A.9**, since the vector components a_i can be recovered from the dot product $\vec{a} \cdot \vec{\sigma}$.

We are greatly indebted to our research colleagues for innumerable and extensive discussions on all aspects of this paper and for their many comments leading to valuable improvements of the manuscript. In particular, we owe special thanks to Lynn Nelson, Rene Essiambre, Gerry Foschini, Basil Hakki, Fred Heismann, Bob Jopson, Bill Shieh, and Mark Shtaif.

- Rashleigh, S. C. & Ulrich, M. (1978) *Opt. Lett.* **3**, 60–62.
- Poole, C. D. & Wagner, R. E. (1986) *Elect. Lett.* **22**, 1029–1030.
- Poole, C. D. & Nagel, J. A. (1997) in *Optical Fiber Telecommunications*, eds. Kaminow, I. P. & Koch, T. L. (Academic, San Diego), Vol. IIIA, pp. 114–161.
- Heismann, F. (1998) *ECOC'98 Digest 2* (Sept.), 51–79.
- Menyuk, C. R. (1999) *J. Eng. Math.* **36**, 113–136.
- Goldstein, H. (1959) *Classical Mechanics* (Addison-Wesley, Reading, MA).
- Messiah, A. (1961) *Quantum Mechanics* (North Holland, Amsterdam).
- Weinberg, S. (1995) *The Quantum Theory of Fields* (Cambridge Univ. Press, Cambridge, U.K.).
- Sattinger, D. H. & Weaver, O. L. (1985) *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics* (Springer, Berlin).
- Born, M. & Wolf, E. (1987) *Principles of Optics* (Pergamon, New York), 6th Ed.
- Huard, S. (1997) *Polarization of Light* (Wiley, New York).
- Jones, R. C. (1948) *J. Opt. Soc. Am.* **38**, 671–685.
- Gisin, N. & Pelloux, J. P. (1992) *Opt. Commun.* **89**, 316–323.
- Jopson, R. M., Nelson, L. E. & Kogelnik, H. (1999) *IEEE Photon. Technol. Lett.* **11**, 1153–1155.
- Karlsson, M. (1998) *Opt. Lett.* **23**, 688–690.
- Mollenauer, L. F. & Gordon, J. P. (1994) *Opt. Lett.* **19**, 375–377.
- Galtarossa, A. & Palmieri, L. (1998) *Elect. Lett.* **34**, 492–493.
- Shieh, W. (1999) *IEEE Photon. Technol. Lett.* **11**, 677.
- Nelson, L. E., Jopson, R. M., Kogelnik, H. & Gordon, J. P. (2000) *Optics Express* **6**, 158–167.
- Ulrich, R. (1977) *Opt. Lett.* **1**, 109–111.
- Ulrich, R. & Simon, A. (1979) *Appl. Opt.* **18**, 2241–2251.
- Poole, C. D., Winters, J. H. & Nagel, J. A. (1991) *Opt. Lett.* **16**, 372–374.
- Foschini, G. J. & Poole, C. D. (1991) *J. Lightwave Technol.* **9**, 1439–1456.
- Frigo, N. J. (1986) *IEEE J. Quantum Electron.* **22**, 2131–2140.
- Eickhoff, W., Yen, Y. & Ulrich, R. (1981) *Appl. Opt.* **20**, 3428–3435.
- Curti, F., Daino, B., De Marchis, G. & Matera, F. (1990) *J. Lightwave Technol.* **8**, 1162–1166.