8

Properties of Polarization-Dependent Loss and Polarization-Mode Dispersion

Polarization-dependent loss (PDL) and polarization-mode dispersion (PMD) are two linear properties that are encountered when dealing with long spans of single-mode fiber. PDL and PMD are not confined to fiber, however, and may be present in components, simple optical elements such as birefringent crystals and polarizers, and instruments that generate these effects. Yet the study of PDL and PMD and their concatenation properties was motivated by research conducted over the last 30 years on fiber-optic communication links. Indeed, polarization-dependent effects in fiber optics have been studied since the 1970s [37, 55].

Polarization-dependent loss refers to energy loss that is preferential to one polarization state. In the Jones picture, one axis suffers more loss than the other. This differential loss changes the output polarization state and imparts a common loss to an unpolarized light beam. The complement of polarization-dependent loss is polarization-dependent gain (PDG). PDG is an effect that is related to preferential gain between signal and noise in optical amplifiers, which, if left unaddressed, can cause impairments of its own sort. The specifics of PDL detailed in this work are equally applicable to PDG, but impairment analysis of the latter must include the combined effects of noise and signal. This is beyond the scope of the present text and the Reader is referred to Desurvire [7].

Polarization-mode dispersion refers to the polarization effects of concatenated lossless birefringent segments. Each homogeneous segment produces differential-group delay. PMD is generated when two or more differential-group delay segments are placed in cascade. While PMD can have the properties of a single homogeneous birefringent element (when the birefringent axes are aligned along a multi-segment cascade), a single homogeneous birefringent element does not possess all of the properties of PMD. PMD is not defined according to an eigenvector analysis of the transformation matrix as one is accustomed to for a single birefringent element.

The combination of PDL and PMD creates effects which are quite complicated and which can impair a communication system more than either effect

Jay N. Damask: Polarization Optics in Telecommunications, SSOS 101, 297–384 (2005)

alone. For example, PDL is generally wavelength independent, while PMD is generally wavelength dependent. Addition of some PMD to PDL will generally result in wavelength-dependent PDL. The addition of some PDL to PMD can in some cases result in differential-group delay that exceeds what one would expect from the PMD alone.

In terms of partial polarization, PDL and PMD have opposite effects. PDL tends to polarize a partially polarized signal since it is a partial polarizer; PMD, however, tends to depolarize a signal since PMD creates polarization-state dispersion in frequency. Lyot depolarizers exploit polarization-state dispersion to (pseudo-)depolarize a transmission signal (cf. §1.5.3), but since this is tantamount to adding PMD to the system such one- and two-stage depolarizers, studied extensively in the late 1980s, have fallen out of favor.

The Mueller matrix for PDL and PMD highlights some of the properties one can expect from these two effects. PDL, being the effect of a partial polarizer, is represented by a Jones matrix that is Hermitian. When converted to a Mueller matrix $\mathbf{M}_{\mathrm{PDL}}$ in Stokes space, all 16 entries may be occupied:

$$H \mapsto \mathbf{M}_{\mathrm{PDL}} = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{pmatrix}, \quad \text{and} \quad \underline{H} \mapsto \mathbf{M}_{\mathrm{PMD}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix}$$

Given this form, PDL is not reversible unless gain is added. PMD, being a physical quantity that is measurable, is also represented by a Hermitian Jones matrix. Unlike PDL, however, the PMD matrix is identically traceless. One can argue this simply based on the lossless cascade of retarders that generates PMD. Accordingly, its Mueller matrix $\mathbf{M}_{\mathrm{PMD}}$ only contains entries in the lower sub-matrix. The effect of PMD, therefore, can in principle be reversed. Finally, the addition of PDL anywhere along a birefringent cascade scatters the PMD Mueller-matrix entries into all 16 sites. The combination cannot therefore be strictly inverted.

The following sections of this chapter define the PDL and PMD vectors, show their effects on input states of polarization, derive equations of motion for concatenated vectors, and illustrate how these effects impact a waveform. The following chapter details the statistical properties of these effects.

8.1 Polarization-Dependent Loss

There are many mechanisms that generate polarization-dependent loss along a fiber-optic link. Micro-optic components such as those studied in the preceding chapters generate PDL when there is preferential coupling through a focusing lens of one polarization-diversity path over another. Integrated optic filters generate PDL because the passband frequency locations are polarization dependent. Fused fiber couplers generate PDL due to polarization-dependent

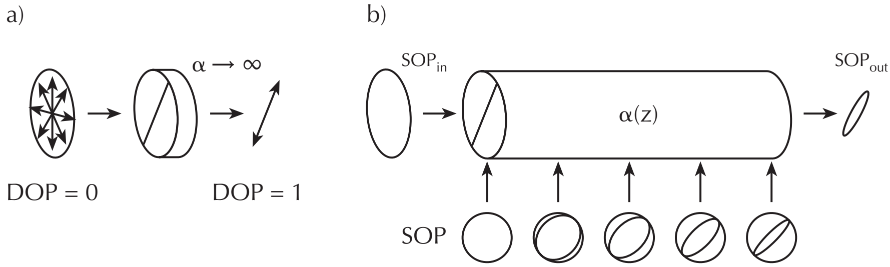

Fig. 8.1. Effects of PDL. a) A completely depolarized signal is completely polarized after transmission through a perfect polarizer. b) A partial PDL element continuously changes the polarization state of input light and diminishes the overall intensity.

coupling ratios. Optical amplifiers generate PDG due to preferential gain of the signal polarization over the orthogonal noise polarization [59]. Finally, micro-bends in fiber will generate PDL. Materials themselves can generate PDL because of dichroism of the molecules, such as in polymer waveguides. Regardless of the origin, however, a single formalism describes the effects.

8.1.1 Definitions

Table 8.1 gives a symbol list for the terms used in the following analysis. There is some inconsistency of notation in the literature, but the notation adopted here is intended to include the broadest overlap.

Figure 8.1 illustrates two effects of PDL. When PDL is complete, the element acts as a perfect polarizer (Fig. 8.1(a)). A perfect polarizer transmits only states that have a finite projection along the polarizer axis. After polarization, a completely depolarized signal becomes completely polarized. The PDL studied below is for partial polarization, not complete. Moreover, polarization occurs continuously through a differentially lossy medium. Figure 8.1(b) illustrates the evolution of light that is initially circularly polarized along a homogeneous PDL element. As the light travels the intensity along the loss axis diminishes while that along the neutral axis is unaltered. One can see how PDL transforms the polarization state along the element. Light launched along the PDL axis is unaltered and undiminished, and light launched along the orthogonal axis is unaltered in state by suffers loss.

Polarization-dependent loss ρ_{dB} is defined by international standards bodies such as the TIA and IEC as [34, 35, 60]

$$\rho_{\mathrm{dB}} \equiv 10 \log_{10} \left(\frac{T_{\max}}{T_{\min}} \right) \tag{8.1.1}$$

where $T_{\max}$ and $T_{\min}$ are the maximum and minimum transmission intensities through the system. PDL is defined in decibels and is positive. Maximum transmission intensity occurs when the polarization state of a completely polarized probe beam is aligned to the maximum transmission axis of the PDL element. That axis may be anywhere on the Poincaré sphere. Minimum transmission occurs for the orthogonal polarization (even in the presence of PMD).

Consider an intuitive example. The Jones matrix of a PDL element aligned to the horizontal (S_1 in Stokes space) may be written as

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & \\ & \mathrm{e}^{-\alpha} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{8.1.2}$$

where α is the loss coefficient. When the input polarization is $(1,0)^T$ the output intensity is $|v_1|^2 = 1$. Similarly, for $(0,1)^T$ the output intensity is $|v_2|^2 = \exp(-2\alpha)$. Therefore the PDL is

$$\rho_{\mathrm{dB}} = \alpha(20 \log_{10} \mathrm{e}) \tag{8.1.3}$$

This relation holds true for any orientation of the PDL vector. Note that $20 \log_{10} \mathrm{e} \simeq 8.686$. Moreover, for $\rho_{\mathrm{dB}} = 3$ dB, $\alpha \simeq 0.345$.

In a fiber-optic link a PDL element is generally located somewhere between the terminations. As single-mode fiber does not preserve polarization in a practical link, the apparent orientation of the PDL vector at the fiber termination generally will not give a purely diagonal Jones matrix. Instead, the PDL vector can point in any direction. In particular the output matrix is $P' = UPV^{\dagger}$, where U, V are unitary operators.

The generalization of the PDL operator P, whose simple case is that in (8.1.2), comes in the matrix exponential form:

$$P = \mathrm{e}^{-\alpha/2} \exp\left(\frac{\vec{\alpha} \cdot \vec{\sigma}}{2}\right) \tag{8.1.4}$$

where local PDL vector $\vec{\alpha} = \alpha\hat{\alpha}$ and $\hat{\alpha}$ is a unit vector in Stokes space that points in the direction of maximum transmission. This matrix exponential operator is expanded using (2.5.77) on page 62, yielding

$$P = \mathrm{e}^{-\alpha/2} \left(I \cosh(\alpha/2) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha/2)\right) \tag{8.1.5}$$

Indeed (8.1.2) is recovered with $\hat{\alpha} = \hat{s}_1$.

An input state of polarization is altered by the PDL element according to

$$|t\rangle = P\,|s\rangle$$

Assuming that $\langle s\,|s\rangle = 1$, the transmission is

$$\langle t\,|t\rangle = \langle s\,|\,P^{\dagger}P\,|s\rangle$$

As the exponent of P is real, one has $P^{\dagger}P = P^2$. Therefore,

Table 8.1. Symbol Definitions for PDL

Symbol	Range	Definition
ρ_{dB} :	$0 \leq \rho_{\mathrm{dB}} < \infty$	Polarization-dependent loss in decibels (dB) as defined by the TIA and IEC
$\vec{\alpha}, \hat{\alpha}$:		PDL vector and unit vector in Stokes space. $\vec{\alpha}$ points in the direction of maximum transmission
α :	$0 \leq \alpha < \infty$	Loss coefficient $\lvert\vec{\alpha}\rvert$, related to ρ_{dB} via: $\rho_{\mathrm{dB}} = \alpha(20 \log_{10} \mathrm{e})$
Γ :	$0 \leq \Gamma \leq 1$	Unit loss; normalization of α via: $\Gamma = \tanh \alpha$
$\vec{\Gamma}$:		Vector of cumulative PDL over a concatenation. $\vec{\Gamma}$ points in the direction of maximum transmission
$\vec{\alpha}(z)$:		PDL vector per unit length
$\vec{\Gamma}(z)$:		Cumulative PDL vector per unit length

$$
\begin{aligned}
\langle t \,|t\rangle &= \mathrm{e}^{-\alpha} \langle s \,|\, (I \cosh(\alpha) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha)) \,|s\rangle \\
&= \mathrm{e}^{-\alpha} \left(\cosh \alpha + (\hat{\alpha} \cdot \hat{s}) \sinh \alpha\right)
\end{aligned}
$$

Denoting $T_p = \langle t \,|t\rangle$, the transmission intensity is

$$
T_p = \frac{1}{1 + \tanh \alpha} \left(1 + \tanh \alpha (\hat{\alpha} \cdot \hat{s})\right) \tag{8.1.6}
$$

The transmission depends not only on the loss coefficient α but the relative orientation of the PDL $\hat{\alpha}$ to the incoming state of polarization $\hat{s}$. Note that $\hat{s}$ is the state incident on the PDL element, not necessarily the state launched into a fiber far away from the element. The extrema of transmission are

$$
T_p = \begin{cases} 1 & \hat{\alpha} \cdot \hat{s} = 1 \\ \mathrm{e}^{-2\alpha} & \hat{\alpha} \cdot \hat{s} = -1 \end{cases}
$$

Calculation of the PDL in dB recovers the definition (8.1.3). The relation between loss coefficient α PDL in decibels clearly holds in general.

One point to note is that transmission coefficients T_p do not multiply. That is, $T_{p21} \neq T_{p2} T_{p1}$. This is of course the result of change in the output polarization state from each PDL element, and a cascade exhibits multiple polarization alternations along its length. The correct calculation is $T_{p12} = \langle s \,|\, P_1^\dagger P_2^\dagger P_2 P_1 \,|s\rangle$.

The scalar transmission T_p can be mapped in Stokes space to show all possible polarization inputs. Figure 8.2 illustrates four such examples. The surfaces are the product of the transmission coefficient with input states that lie on the unit sphere. Figures 8.2(a,b) illustrate single PDL elements with 3 dB and 30 dB PDL, respectively. Note that the latter surface goes concave. The transmission contour along the equator is projected to the bottom surface, where the orthogonal directions of no-loss and high-loss are apparent. Figures 8.2(c,d) illustrate the transmission through two PDL elements in cascade, both having 3 dB PDL. In the first case the PDL elements are aligned, resulting in a 6 dB overall PDL. However, in the second case the PDL elements are crossed. This results in a sphere with radius 0.5. There is no PDL at the output, but there is a uniform insertion loss of 3 dB. This shows a simple example of the more general fact that PDL can increase or decrease depending on the relative orientation of multiple elements, but the insertion loss always accumulates.

Thus far the input has been considered completely polarized. When the input is completely depolarized, the transmission is averaged over all polarization states. Since it is clear that $\langle \hat{\alpha} \cdot \hat{s} \rangle = 0$, the average transmission for unpolarized light is

$$\langle T_p \rangle = T_{\text{depol}}$$

where T_{depol} is the transmission coefficient for depolarized light. To connect with the literature, (8.1.6) is recast into

$$T_p = T_{\text{depol}}\,(1 + \Gamma(\hat{\alpha} \cdot \hat{s})) \tag{8.1.7}$$

where the loss coefficient is functionally normalized as

$$\Gamma \equiv \tanh \alpha \tag{8.1.8}$$

and the transmission for depolarized light is

$$T_{\text{depol}} = \frac{1}{1 + \Gamma} \tag{8.1.9}$$

The minimum and maximum transmissions are therefore

$$\begin{aligned} T_{\max} &= T_{\text{depol}}\,(1 + \Gamma) \\ T_{\min} &= T_{\text{depol}}\,(1 - \Gamma) \end{aligned}$$

The relationship between the normalized loss coefficient and the transmission extrema is

$$\Gamma = \frac{T_{\max} - T_{\min}}{T_{\max} + T_{\min}}, \quad \text{and} \quad \frac{T_{\max}}{T_{\min}} = \frac{1 + \Gamma}{1 - \Gamma} \tag{8.1.11}$$

The decibel expression for PDL can be written in these terms:

$$\rho_{\text{dB}} = 10 \log_{10} \left(\frac{1 + \Gamma}{1 - \Gamma} \right) \tag{8.1.12}$$

Moreover, these definitions are used to write the Jones and Mueller matrices for a PDL element oriented along the horizontal:

$$\mathbf{J}_{\hat{s}_1} = T_{\text{depol}}^{1/2} \begin{pmatrix} \sqrt{1 + \Gamma} & \\ & \sqrt{1 - \Gamma} \end{pmatrix} \tag{8.1.13}$$

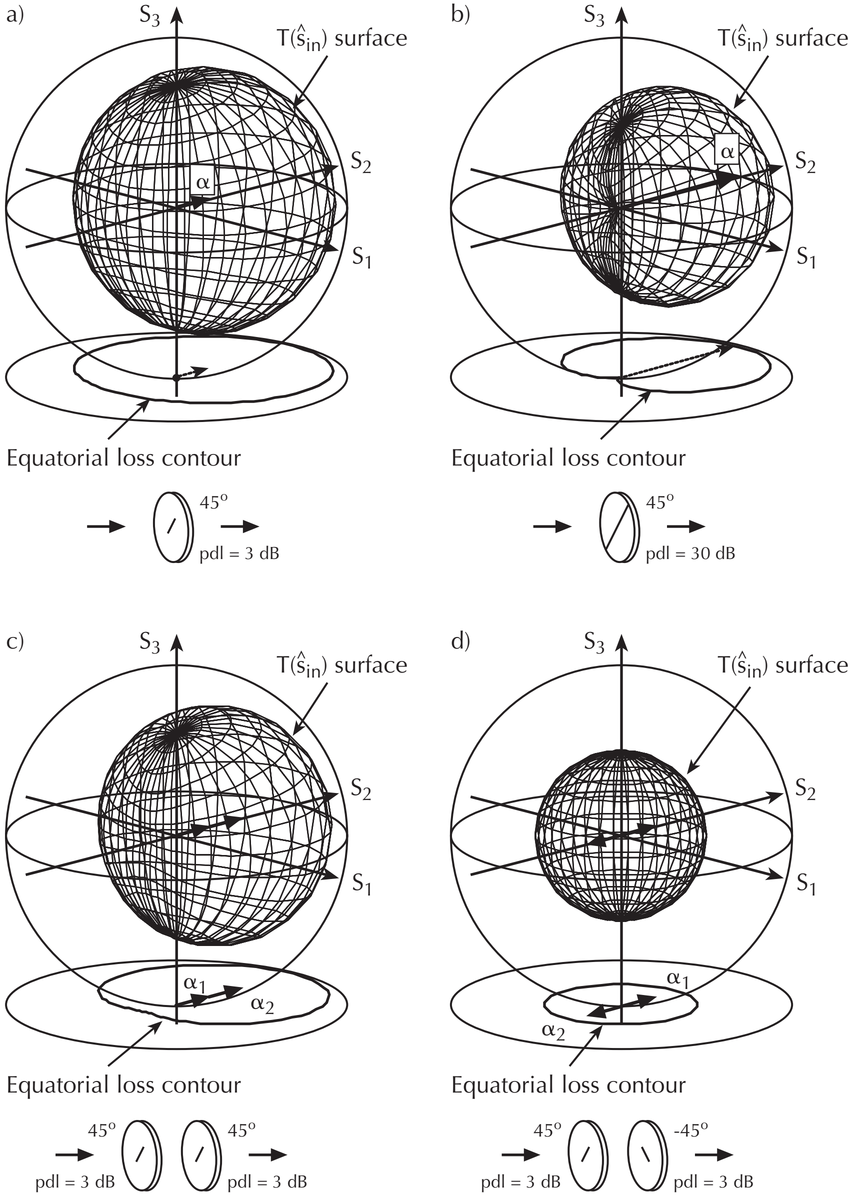

Fig. 8.2. These surfaces plot the transmission T_p as a function of input polarization state $\hat{s}_{\rm in}$. The contours at the bottom are projections of T_p along the equator. a) Transmission surface for single 3 dB PDL vector at $+45°$. b) Transmission surface for same PDL element but with 30 dB PDL. Note this surface is concave. c) Transmission surface after two aligned PDL elements, 3 dB PDL each. d) Transmission surface after two orthogonally aligned PDL elements, 3 dB PDL each. Note surface is spherical (no PDL) but has a smaller radius reflecting the insertion loss.

and

$$\mathbf{M}_{\hat{s}_1} = T_{\text{depol}} \begin{pmatrix} 1 & \Gamma & & \\ \Gamma & 1 & & \\ & & \sqrt{1-\Gamma^2} & \\ & & & \sqrt{1-\Gamma^2} \end{pmatrix} \tag{8.1.14}$$

where the Mueller matrix is related to the Jones matrix through the spin-vector expression (1.4.22) on page 18. Note that the Jones matrix can also be written as $\mathbf{J} = \text{diag}(\sqrt{T_{\max}}, \sqrt{T_{\min}})$.

8.1.2 Change of Polarization State

The state of polarization at the output from a PDL element generally differs from the input state. This is easily imaged: a right-hand circular state that is transmitted through a PDL element has one axis shortened. Accordingly, the state is altered from circular to elliptical.

The output polarization state is determined from the PDL operator P in the following way. The output unit Stokes vector is

$$\hat{t} = \frac{\langle t\,|\vec{\sigma}|\,t\rangle}{\langle t\,|t\rangle} \tag{8.1.15}$$

Substitution of the P expansion (8.1.5) into the numerator yields

$$\langle t\,|\vec{\sigma}|\,t\rangle = \mathrm{e}^{-\alpha}\langle s\,|\Big(\,\big(Ic_{\alpha/2} + s_{\alpha/2}(\hat{\alpha}\cdot\vec{\sigma})\big)\,\vec{\sigma}\,\big(Ic_{\alpha/2} + s_{\alpha/2}(\hat{\alpha}\cdot\vec{\sigma})\big)\Big)\,|s\rangle$$

where $s_{\alpha/2} = \sinh(\alpha/2)$ and $c_{\alpha/2} = \cosh(\alpha/2)$. Application of the identities

$$\begin{aligned} \langle s\,|\vec{\sigma}|\,s\rangle &= \hat{s} \\ \langle s\,|\vec{\sigma}(\hat{\alpha}\cdot\vec{\sigma}) + (\hat{\alpha}\cdot\vec{\sigma})\vec{\sigma}|\,s\rangle &= 2\hat{\alpha} \\ \langle s\,|(\hat{\alpha}\cdot\vec{\sigma})\vec{\sigma}(\hat{\alpha}\cdot\vec{\sigma})|\,s\rangle &= 2\langle s\,|(\hat{\alpha}\hat{\alpha}\cdot)\vec{\sigma}|\,s\rangle - \langle s\,|\vec{\sigma}|\,s\rangle \\ &= 2\hat{\alpha}(\hat{\alpha}\cdot\hat{s}) - \hat{s} \end{aligned}$$

produces

$$\langle t\,|\vec{\sigma}|\,t\rangle = \mathrm{e}^{-\alpha}\left[\,\hat{s} + s_\alpha\left(1 + t_{\alpha/2}(\hat{\alpha}\cdot\hat{s})\right)\hat{\alpha}\right]$$

Using the previously determined expression for $\langle t\,|t\rangle$, the unit Stokes vector $\hat{t}$ is governed by the relation [17]

$$\hat{t} = \frac{\sqrt{1-\Gamma^2}}{1 + (\Gamma\hat{\alpha}\cdot\hat{s})}\,\hat{s} + \frac{1 + \Gamma^{-2}\left(1 - \sqrt{1-\Gamma^2}\right)(\Gamma\hat{\alpha}\cdot\hat{s})}{1 + (\Gamma\hat{\alpha}\cdot\hat{s})}\,\Gamma\hat{\alpha} \tag{8.1.16}$$

where the following identification is made

$$\frac{\tanh(\alpha/2)}{\tanh\alpha} = \frac{1 - \sqrt{1-\Gamma^2}}{\Gamma^2}$$

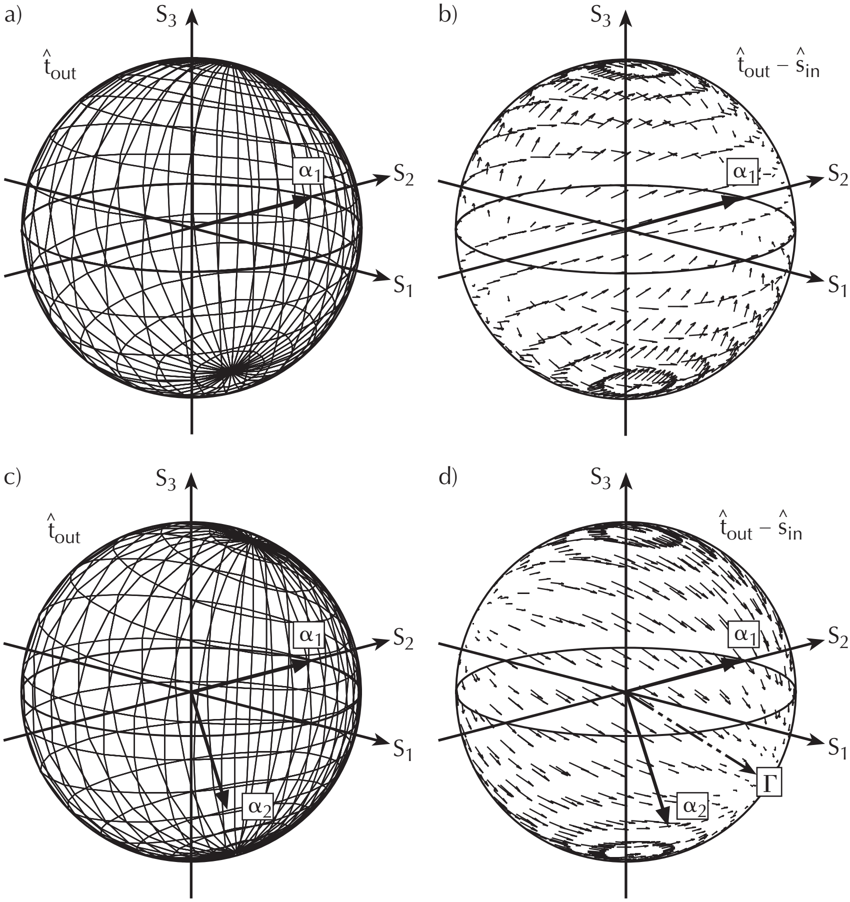

Fig. 8.3. Two examples of $\hat{t}_{\text{out}}$. Left figures show mesh of normalized $\hat{t}_{\text{out}}$ for one and two PDL elements. Right figures show vector difference plot to indicate the change $\hat{t}_{\text{out}} - \hat{s}_{\text{in}}$. a) and b) Single 3 dB PDL element aligned along S_2. All states but $\pm S_2$ are pulled toward S_2. c–d) Two 3 dB PDL elements cascaded, second element having elliptical PDL vector (or intermediate unitary rotation and linear PDL vector). Note in d) that cumulative PDL vector $\vec{\Gamma}$ points in a direction between $\vec{\alpha}_1$ and $\vec{\alpha}_2$. It is along $\vec{\Gamma}$ that the output states are pulled.

The transformation expression (8.1.16) has a characteristically different form than encountered previously in this text (Fig. 8.3). One is accustomed to $\hat{t} = R\hat{s}$, where R is a rotation operator in Stokes space. Instead, polarization-transformation through a PDL element has the form $\hat{t} = u\hat{s} + v\hat{\alpha}$, where u, v are positive coefficients that are themselves a function of the relative orientation of $\hat{s}$ and $\hat{\alpha}$. Physically, the input polarization state is "pulled" toward the PDL vector direction, where the pull strength is dictated by Γ and the relative orientation. This pulling effect is shown in Fig. 8.3(b,d).

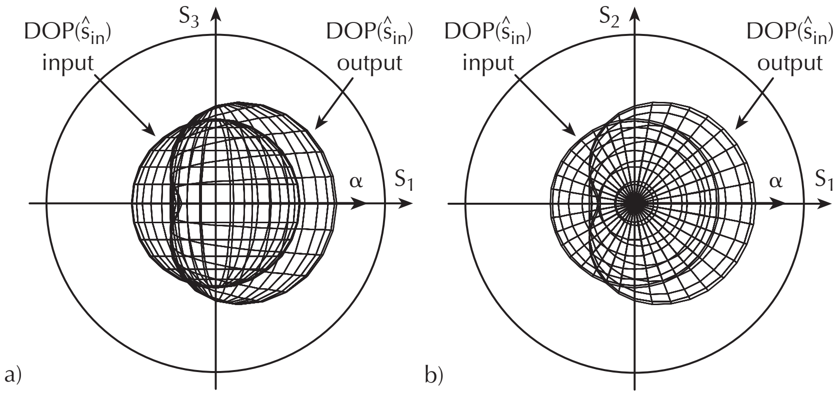

Fig. 8.4. Degree-of-polarization surfaces, before and after single PDL element with 3 dB loss and aligned along S_1, illustrate repolarization. a) View of the (S_1, S_3) plane. b) View of the (S_1, S_2) plane. In both cases, spherical surface has $\mathcal{D}_{\text{in}} = 0.5$. The $\mathcal{D}_{\text{out}}$ surface is a map of $\mathcal{D}_{\text{out}}(\hat{s}_{\text{in}})$, or the DOP as a function of input polarization. The right hemisphere of both plots shows repolarization of the signal ($\mathcal{D}_{\text{out}} \geq \mathcal{D}_{\text{in}}$). The left hemisphere shows increased depolarization.

8.1.3 Repolarization

Intuitively one expects that depolarized light which passes through an ideal polarizer attains a unity degree of polarization. Repolarization [45] is the effect of generating partially polarized light from depolarized light by transiting through one or more PDL elements. However, transit of a PDL element with partially polarized light (cf. §1.5) can either increase or decrease the degree of polarization, depending on orientation [17].

The Mueller matrix for a single PDL element (8.1.14) governs re- and depolarization. Without loss of generality the matrix elements will remain fixed in the analysis below while the input polarization state varies. To summarize the findings:

$$
\begin{aligned}
\mathcal{D}_{\text{in}} = 1 \quad &\longrightarrow \quad \mathcal{D}_{\text{out}} = 1 \\
\mathcal{D}_{\text{in}} = 0 \quad &\longrightarrow \quad \mathcal{D}_{\text{out}} = \Gamma \\
\mathcal{D}_{\text{in}} = d \quad &\longrightarrow \quad \mathcal{D}_{\text{out}} = f(d, \Gamma, \hat{\alpha} \cdot \hat{s})
\end{aligned}
$$

where f is the function (8.1.18). These cases are derived as follows. The output Stokes vector for an arbitrary input having $\mathcal{D}_{\text{in}} = d$ is

$$
S_{\text{out}} = T_{\text{depol}} \begin{pmatrix} 1 & \Gamma & & \\ \Gamma & 1 & & \\ & & \sqrt{1-\Gamma^2} & \\ & & & \sqrt{1-\Gamma^2} \end{pmatrix} \begin{pmatrix} 1 \\ d\cos\phi\sin\theta \\ d\sin\phi\sin\theta \\ d\cos\theta \end{pmatrix} \qquad (8.1.17)
$$

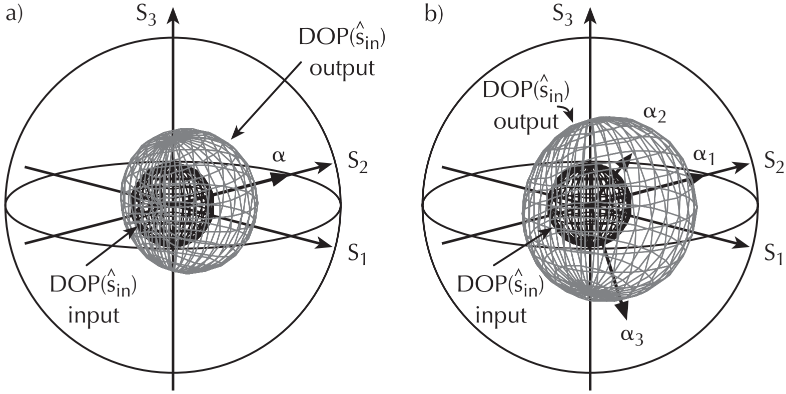

Fig. 8.5. Repolarization surfaces for $\mathcal{D}_{\text{in}} = 0.25$ after one (a) and three (b) PDL elements. All elements have 3 dB PDL. Note that for these large PDL values the repolarization is quickly established.

For $d = 1$ the output partial polarization in all cases is $\mathcal{D}_{\text{out}} = 1$. Conversely, for $d = 0$, only the S_0 and S_1 terms survive so $\mathcal{D}_{\text{out}} = \Gamma$. For an arbitrary d such that $0 \leq d \leq 1$, the output DOP is

$$\mathcal{D}_{\text{out}} = \frac{\sqrt{(1 + d\Gamma c_\phi s_\theta)^2 - (1 - d^2)(1 - \Gamma^2)}}{1 + d\Gamma c_\phi s_\theta} \tag{8.1.18}$$

Two degree-of-polarization surfaces as plotted in Fig. 8.4. The partial polarization at the output increases when the polarized portion of the input light is aligned to the direction of maximum transmission and decreases when aligned to the direction of maximum extinction. This behavior follows what one would expect.

While the above examples used a single PDL vector aligned along S_1, a Mueller matrix can be calculated for any PDL operator P or concatenation of operators. Such a generalized Mueller matrix is used to calculate the repolarization surfaces illustrated in Fig. 8.5. In both cases $\mathcal{D}_{\text{in}} = 0.25$ and PDL for each segment is 3 dB. While the single-segment case exhibits both re- and de-polarization, the three-segment case exhibits repolarization for all input states.

The effect of repolarization is statistical when many PDL elements are distributed along a birefringent cascade. This is the model of a transmission link. Repolarization can be an impairment for long-haul transmission systems that launch polarized light to mitigate polarization-dependent gain effects from the erbium amplifiers. As the light becomes repolarized, PDG impairments begin to accumulate. The statistics of repolarization are derived in [45].

8.1.4 PDL Evolution Equations

Polarization-dependent loss in a communications link comes from optical components or imperfections along the optical fiber. In either case a local PDL vector $\vec{\alpha}$ each source. Polarization-dependent loss through a chain of PDL element accumulates and is denoted by $\vec{\Gamma}$. The cumulative PDL vector will always try to track to the local PDL element, but if the elements have low PDL and are randomly distributed, the cumulative vector eventually decorrelates. In this section the evolution of $\vec{\Gamma}$ is derived; its driving term is the local PDL $\vec{\alpha}$.

Two slightly different derivations for the evolution of $\vec{\Gamma}$ through a chain of arbitrarily oriented PDL elements are given by Gisin and Huttner, and Mecozzi and Shtaif [44]. The Gisin and Huttner derivation [18, 31] tracks the evolution of the PDL vector as defined at the output. Mecozzi and Shtaif do the opposite, tracking the PDL vector as defined at the input. Consider the output polarization produced by a chain of PDL elements acting on an input state: $|t\rangle = AT\,|s\rangle$ where, as in (8.1.4), A is the common loss and T is an Hermitian operator. The output intensity is then

$$\langle t\,|t\rangle = A^2\langle s\left|T^\dagger T\right| s\rangle \tag{8.1.19}$$

Since T is Hermitian, the operator $T^\dagger T$ may be written in spin-vector form: $T^\dagger T = p_o I + \vec{p}\cdot\vec{\sigma}$. Taking $\langle s\,|s\rangle = 1$, one finds

$$\langle t\,|t\rangle = A^2\left(p_o + \vec{p}\cdot\hat{s}\right) \tag{8.1.20}$$

The maximum and minimum output intensities are for alignment and anti-alignment of the input state $\hat{s}$ with $\vec{p}$. Therefore $\vec{p}$ represents the cumulative PDL vector referenced to the input.

Reference to the output state comes from the complement of (8.1.19):

$$\langle s\,|s\rangle = A^{-2}\left(\langle t\,|\left(TT^\dagger\right)^{-1}|t\rangle\right)$$

Denoting $QQ^\dagger$ as the inverse of $TT^\dagger$, the spin-vector form is $QQ^\dagger = q_o I + \vec{q}\cdot\vec{\sigma}$. Still assuming the input intensity is normalized, the output intensity can be expressed as a function of the output polarization state:

$$\langle t\,|t\rangle = \frac{A^{-2}}{q_o + \vec{q}\cdot\hat{t}}$$

Now the minimum and maximum output intensities are for alignment and anti-alignment of the *output* state $\hat{t}$ with $\vec{q}$.

The derivation below follows the Mecozzi and Shtaif choice of reference frame: equations of motion are referred to the input.

Referring back to (8.1.20), the extrema in transmission are clearly

$$T_{\max} = A_N^2\left(p_o + p\right), \quad \text{and} \quad T_{\min} = A_N^2\left(p_o - p\right)$$

where $p = |\vec{p}|$. The PDL is therefore

$$\rho_{\mathrm{dB}} = 10\log_{10}\left(\frac{1+p/p_o}{1-p/p_o}\right) \tag{8.1.21}$$

In light of the analogous expression (8.1.12), one can expect that the cumulative PDL magnitude Γ is related to the spin-vector as $\Gamma = p/p_o$. The vector form follows this relation and will be used in a moment.

The evolution of p_o and $\vec{p}$ is determined by the evolution of $T^\dagger T$. In the continuum limit, the cumulative loss A and PDL operator T are

$$A = \exp\left(-\frac{1}{2}\int_0^z \alpha(z)dz\right), \quad \text{and} \quad T = \exp\left(\frac{1}{2}\int_0^z \vec{\alpha}(z)\cdot\vec{\sigma}\, dz\right) \tag{8.1.22}$$

where $\vec{\alpha}(z)$ represents the derivative in z of the local PDL vector. (Gisin and Huttner use a discretized version to emphasize the form of the derivatives to follow.) The output intensity is

$$\langle t\,|t\rangle = A^2\langle s\left|T^\dagger T\right| s\rangle = A^2\langle s\,|p_o I + \vec{p}\cdot\vec{\sigma}|\, s\rangle \tag{8.1.23}$$

Now, the evolution is determined by the differential change of $T^\dagger T$. Taking the derivative along z (from input to output) of both sides of (8.1.23) gives

$$\frac{dT^\dagger}{dz}T + T^\dagger\frac{dT}{dz} = \frac{d}{dz}\left(p_o I + \vec{p}\cdot\vec{\sigma}\right)$$

with the initial conditions of $p_o = 1$ and $\vec{p} = 0$ at $z = 0$. The spatial evolution of T and its adjoint comes from (8.1.22) and (2.5.82) on page 63

$$\frac{dT^\dagger}{dz}T + T^\dagger\frac{dT}{dz} = \frac{1}{2}\,\left(\left(\vec{\alpha}(z)\cdot\vec{\sigma}\right)T^\dagger T + T^\dagger T\left(\vec{\alpha}(z)\cdot\vec{\sigma}\right)\right)$$

Plugging in the spin-vector form of $T^\dagger T$ and using the anti-commutator relation $\{(a\cdot\vec{\sigma}), (b\cdot\vec{\sigma})\} = 2(a\cdot b)$ gives

$$\frac{d}{dz}\left(p_o I + \vec{p}\cdot\vec{\sigma}\right) = p_o\left(\vec{\alpha}(z)\cdot\vec{\sigma}\right) + \left(\vec{\alpha}(z)\cdot\vec{p}\right)$$

Finally, separation of spin-vector from scalar terms gives the coupled evolution equations

$$\frac{dp_o}{dz} = \vec{\alpha}(z)\cdot\vec{p}, \quad \text{and} \quad \frac{d\vec{p}}{dz} = p_o\,\vec{\alpha}(z) \tag{8.1.24}$$

These equations are converted to equations for the cumulative PDL vector and depolarization transmission. Define the cumulative PDL vector as

$$\vec{\Gamma}(z) \equiv \frac{\vec{p}\,(z)}{p_o(z)} \tag{8.1.25}$$

and the depolarization transmission T_{depol} as

$$T_{\text{depol}} = A_N^2 p_o \tag{8.1.26}$$

Differentiation with respect to length and substitution of (8.1.24) yields the equations of motion for $\vec{\Gamma}$ and T_{depol}:

$$\frac{d\vec{\Gamma}}{dz} = \vec{\alpha} - \left(\vec{\alpha}\cdot\vec{\Gamma}\right)\vec{\Gamma} \tag{8.1.27a}$$

$$\frac{d\,T_{\text{depol}}}{dz} = \left(\vec{\alpha}\cdot\vec{\Gamma} - \alpha\right)T_{\text{depol}} \tag{8.1.27b}$$

These elegant equations predict the direction and length of the PDL vector is Stokes space as well as the depolarized transmission coefficient. Note that $0 \le |\vec{\Gamma}| \le 1$, so the PDL vector is bound by the unit sphere.

As for discussion, note how the local PDL element $\vec{\alpha}$ pulls the cumulative PDL vector. If the cumulative and local vectors are initially orthogonal at the start of the n^{th} segment, then $\vec{\alpha}\cdot\vec{\Gamma} = 0$. The differential equation (8.1.27a) then pulls $\vec{\Gamma}$ strongly in the direction of $\vec{\alpha}$. Moreover, since $\Gamma \le \alpha$ always, $\vec{\Gamma}$ asymptotically approaches $\vec{\alpha}$, but never aligns perfectly unless they are aligned initially. This said, the magnitude of $\vec{\Gamma}$ may increase or decrease depending on the cascade.

In contrast, T_{depol} monotonically decreases in all cases. At best T_{depol} is stationary. Lost power is never recovered when propagating through PDL media. Since $0 \le \Gamma \le 1$, $\vec{\alpha}\cdot\vec{\Gamma} \le \alpha$ always.

Figure 8.6 illustrates four examples of cumulative PDL evolution. The details are given in the caption. One point to note, however, is the behavior of the insertion loss in the case of many small-valued PDL elements. Figures 8.6(c,d) show a trend of linear decrease (on a log scale) of T_{depol}. The origin of this behavior is as follows. The general solution to (8.1.27b) is

$$T_{\text{depol}} = \exp\left(-\int_0^z \left(\vec{\alpha}(z) - \vec{\alpha}(z)\cdot\vec{\Gamma}(z)\right)dz\right)$$

In a long chain of PDL elements, the cumulative vector eventually losses track of the local PDL element vector. That is, $\vec{\alpha}$ and $\vec{\Gamma}$ decorrelate. Once the two vectors decorrelate, one can write that the statistical average of the inner product vanishes: $\left\langle \vec{\alpha}_j\cdot\vec{\Gamma}_j\right\rangle = 0$. Beyond this point, the mean value of the insertion loss goes as

$$T_{\text{depol}} \simeq \exp\left(-\int_0^z \vec{\alpha}(z)dz\right)$$

The insertion loss plots in Figures 8.6(c,d) illustrate this trend. Moreover, the variance of T_{depol} is related to the variance $\operatorname{var}\left(\vec{\alpha}\cdot\vec{\Gamma}\right)$. The variance of the cosine function uniformly distributed in angle is $1/2$. When all PDL elements have the same value, the variance of T_{depol} is approximately

$$\operatorname{var}\left(T_{\text{depol}}\right) \simeq \exp\left(-\vec{\alpha}(z)\left\langle\Gamma\right\rangle z/2\right)$$

where $\langle\Gamma\rangle$ is the mean Γ over the link.

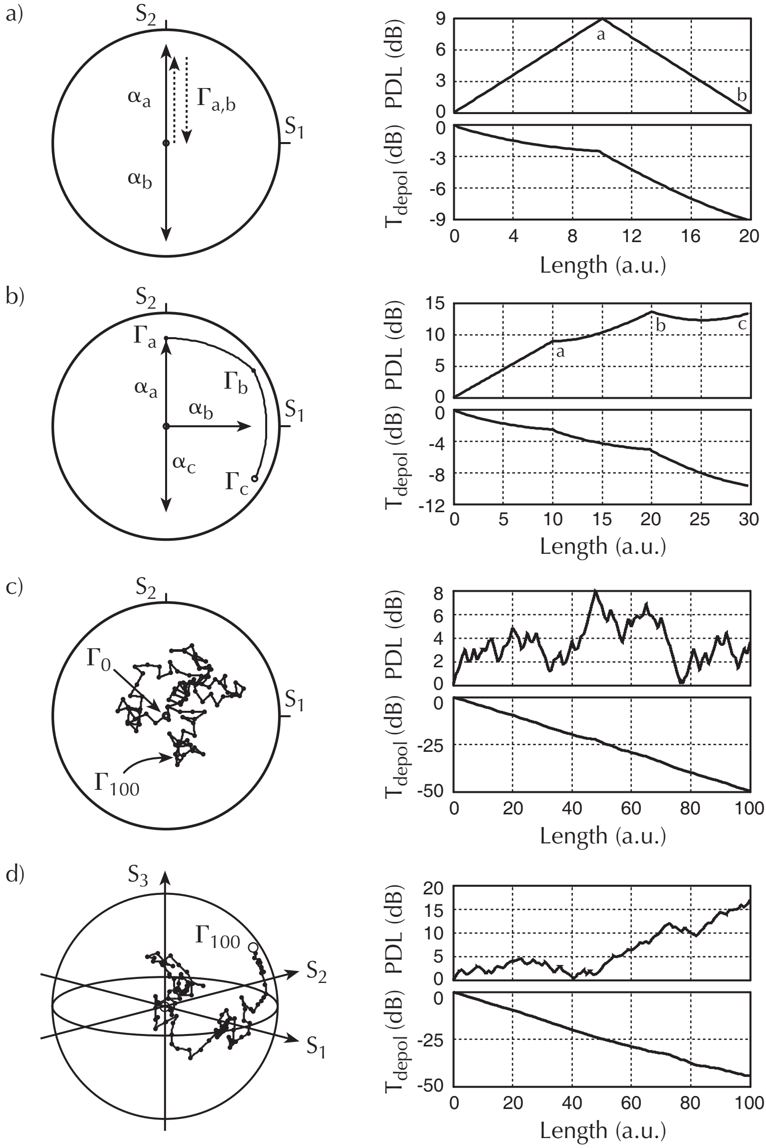

Fig. 8.6. PDL-vector evolution examples. a) Two orthogonal PDL segments $\vec{\alpha}_{1,2}$, 9 dB each. The PDL accumulates to a maximum of 9 dB through the first element and diminishes to zero at the termination of the second element. The insertion loss T_{depol} monotonically decreases to -9 dB. b) Three PDL segments $\vec{\alpha}_{1,2,3}$, 9 dB each, oriented at right angles. The cumulative PDL vector $\vec{\Gamma}$ tries to track the element vectors with increasing disparity. c) One hundred 1 dB randomly oriented PDL segments confined to the equator. The cumulative vector $\vec{\Gamma}$ does a random walk. The insertion loss decreases almost linearly (on log scale). d) One hundred 1 dB PDL segments with random orientation in all directions.

8.2 Polarization-Mode Dispersion

C. D. Poole and R. E. Wagner's seminal paper, entitled, "Phenomenological Approach to Polarisation Dispersion in Long Single-Mode Fibres" [52] marks the historical dividing line between the classical treatment of birefringence and the modern development of polarization-mode dispersion (PMD). In 1986, Drs. Poole and Wagner, while working at Bell Laboratories on fiber communications, created a characteristic description of the frequency-dependence of concatenations of birefringent elements, since recognized as the point of discovery of PMD. Few discoveries, however, are made in a vacuum – indeed a scan of the technical literature shows titles including the term "Polarization-Mode Dispersion" dating back at least to 1978 [2, 56] – and one is inclined to ask about the context which led to the breakthrough.

The year 1986 was before the advent of the fiber-based erbium-doped optical amplifier; the pinnacle of long-distance communications rested at the time on coherent communications. Coherent communications mixes a local oscillator (LO) with the received signal to pull the signal out of the noise. As early as the 1970s, birefringence of single-mode fiber was recognized and that the output polarization state would drift with temperature or abruptly shift when hit was well understood. Since the LO needs to be aligned to the signal polarization, the matter of polarization tracking and control was under intense investigation. Many papers can be found on these subjects throughout the 1980s. In this sense, polarization effects were treated in the time domain, for the speed and control of polarization reside there.

As head of a research department at Crawford Hill Research, Wagner tasked Poole to investigate the overall link properties and characteristics that needed to be understood in order to make coherent communications practicable. A Bell Labs colleague of Dr. Poole's at the time, N. S. Bergano, believed that the job would be completed somewhat quickly. Regarding polarization Dr. Bergano was to say, "its x, y, and a phase – how hard can it be?".

Part of Dr. Poole's work included building a relatively fast polarimeter to measure the speed of polarization change. When a fiber was terminated by the polarimeter and a probe diode laser was turned on, Poole noticed that the output polarization went through a transient but would settle out in a few minutes. Intrigued by this effect, Poole determined it was the diode laser ramping up in temperature which in turn produced a frequency sweep that was responsible for the changing output polarization. The unavoidable conclusion was that the output polarization state was frequency dependent. Determined to describe the effect as an eigenvalue problem, Poole found that the eigenvectors of a Jones transformation matrix always change to first-order in frequency. Nothing distinctive could be derived from that behavior. Poole and Wagner then asked if there exists an output state that is stationary to first-order in frequency – consequently the two Researchers discovered the principal characteristic of PMD.

The principal distinction between PMD research before and after Poole and Wagner's paper is that PMD enables a global description of the birefringence. Only local descriptions were available before 1986. Indeed, polarization optics had been studied for centuries, but always in the context of local birefringent behavior. Perhaps the closest earlier researchers got to the questions that encompass PMD are the several inventions of birefringent filters, including Lyot, Solc, Jones, Pancharatnam, and Harris. Even though, no one had put a global description together before and it is not too surprising that communications researchers were the first ones to make this observation.

With the advent of the optical amplifier, work on coherent communications went into decline. PMD, by contrast, has remained at the forefront of research ever since. Particularly important work includes the statistical treatment of PMD; the use of the statistics to develop link budgets for installed system designs; the measurement and mitigation of PMD in single-mode fiber; the measurement of PMD in components, installed fiber, operating links, and all manner of configurations; programmable PMD generation; and interaction of impairments such as PMD with PDL and chromatic dispersion, and PMD with nonlinear effects.

Moreover, the wealth of research developed in polarization tracking for coherent communications was redirected to active PMD compensation. The father of the optical PMD compensator is Fred Heismann, who through his work on lithium-niobate electro-optic polarization controllers [27, 28] demonstrated the first closed-loop PMDC [29]. Yet at the time of writing, the economics behind optical PMDCs appear unfavorable, even in light of proven high-quality, live-traffic-certified products [50]; and, simultaneously, lower-cost chip sets are being developed to make corrections electronically.

The following sections of this chapter cover the major highlights of the time, frequency, Fourier, and Stokes properties of PMD. These sections will be successful if they arm the reader with the tools necessary to read the literature and patents critically and informatively.

8.2.1 A PMD Primer

Polarization-mode dispersion is an optical effect present in concatenations of lossless birefringent elements. The following observation was first made by Poole [52]:

> Observation: There always exists an orthogonal pair of polarization states output from a birefringent concatenation which are stationary to first order in frequency. These two states are called the Principal States of Polarization (PSP). A differential delay exists between signals launched along one PSP and its orthogonal complement.

Based on this observation, the PMD vector, which has a length and a pointing direction in Stokes space, is defined as follows [20][1] :

> Definition Part 1: The pointing direction of the PMD vector is aligned to the slow PSP, the PSP that imparts more delay than the other.
>
> Definition Part 2: The length of the PMD vector is the differential-group delay between slow and fast PSP's.

The definition of the PMD vector follows immediately from the Poole observation, and that observation is clearly different from the usual input-to-output eigenstate definition for a birefringent concatenation.

The "usual" eigenstate of any concatenation is that polarization state which is the same at the output as at the input. (In contrast, the polarization state associated with the input and output PSP's are generally not the same.) The failure of the eigenstate description happens because the eigenstate almost always changes to first order with change in frequency. Since group delay requires a frequency derivative, the derivative of an eigenstate and its phase results in two terms: the derivative of the phase, which is delay, and the derivative of the eigenstate, which rotates the polarization. The lack of a stationary eigen-polarization prevents the separation of group delay from its eigenstate. In the context of eigenstates, differential-group delay is not uniquely determined.

Another way to express the need for a definition of the PMD vector is to say that a parallelism should exist for infinitesimal rotations of polarization state in length and in frequency. As seen many times earlier in this text, the differential change of output polarization state due to differential change in length of the last birefringent element in a concatenation is

$$\frac{d\hat{s}}{dz} = \beta_n \times \hat{s}$$

where β_n is the birefringent vector of the last element. This is the customary precession rule. What vector, then, does the output polarization precess about for differential change in frequency? That is, what is the vector V such that

$$\frac{d\hat{s}}{d\omega} = V \times \hat{s} \quad ?$$

Equivalent to a change in frequency would be a if all elements in the concatenation changed their length, or refractive index, or both, for the same frequency. That is, the optical phase for any section is

[1] Note: This definition of PMD follows that of Gordon and not that adopted by Poole. The Gordon definition aligns the PMD vector along the slow PSP and uses a right-hand coordinate system in Stokes space. A detailed comparison between the two systems is available [21].

$$\varphi = \frac{\omega}{c} nL$$

Proportional change of any parameter of the optical phase along an entire cascade leads to eigenstate change for first-order change in that parameter. For example, a change in temperature will change the refractive index, and eigenstates of the cascade with change to first order with it. This generalization of PMD was recognized by Shieh and Kogelnik [58]. It is customary, nonetheless, to use frequency as the underlying parameter.

It is important to observe that so far no specific context for PMD has been given other than lossless birefringent concatenations. PMD exists in optical fiber as well as birefringent crystals; it exists for any number of homogeneous sections, whether one or thousands; it exists for any type of statistical distribution, whether Maxwellian or not, and any delay within or accumulated across multiple sections. PMD also exists in the presence of PDL, although its description is more complicated. Whether PMD is a significant effect for an optical communications link or a single component depends on these particulars, but the particulars do not alter its definition.

It is not at all obvious that for any arbitrary lossless birefringent cascade of any length and composition a pair of principal polarization states exists. A good heuristic is to construct and to compare the differential-rotation rules for length and frequency, and to draw an analogy between the principal-state system and the eigenstate system.

To begin, an eigenstate ties the output to the input: the same polarization state must exist at both ends (at any frequency). Likewise, a principal state ties the output to the input but in a different way: the input polarization state must be oriented such that the output polarization state does not change to first order in frequency. Generally, the eigenstate and principal state are not the same. Figure 8.7 illustrates the eigenstates for a single homogeneous birefringent element. When an input polarization is oriented to the "fast" eigen-axis of the element (Fig. 8.7(a)), the refractive index seen by the light is the smaller of the two. The output polarization is the same as the input polarization, and a pulse of light exits at a certain time. Next, when the input polarization is oriented to the "slow" eigen-axis of the element (Fig. 8.7(b)), the refractive index seen by the light is the larger of the two. The output polarization is also the same as the input polarization, and an output pulse is delayed with respect to the fast-axis output pulse. The relative delay time is τ, and is called the differential-group delay (DGD). Note that when the input polarization is oriented either to the fast or slow axis, only one polarization state and one pulse is present at the output; this defines the eigenstate of the system.

In general an input polarization is not aligned to the birefringent axis of the element (Fig. 8.7(c)). In this case, the input state is projected onto the two orthogonal birefringent axes of the element and these projects propagate

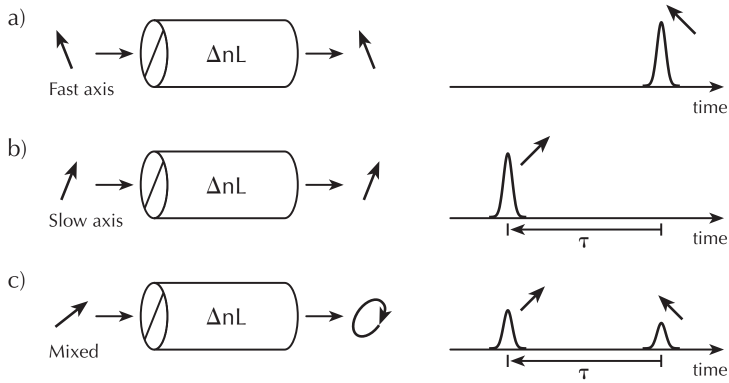

Fig. 8.7. Eigenstates and effect of single homogeneous birefringent element. a) Launch along fast axis gives an output polarization state that is the same as the input state. b) Launch along slow axis also leaves the input state unaltered. There is, however, a delay with respect to the fast axis. c) Arbitrary polarization state at the input is projected onto the two eigen-axes, and these two projects propagate independently.

independently to the output. Accordingly, an input pulse is split into two pulses, one delayed by τ with respect to the other. The relative intensity of the two pulses is dictated by the angle between the input polarization state and the birefringent axis of the element.

Next, consider two birefringent elements in cascade. The polarization evolution is best described in Stokes space. As in Fig. 8.8(a), consider one stage again. The input polarization $\hat{s}_{\text{in}}$ precesses about birefringent axis $\hat{r}_1$ to output state $\hat{s}_1(z_1)$, where the propagation axis is denoted as z. When a second stage is added (Fig. 8.8(b)), polarization state $\hat{s}_1(z_1)$ precesses about birefringent axis $\hat{r}_2$ to new output polarization $\hat{s}_2(z_2)$. Consider now a small increase in length Δz of the second stage. The output polarization comes from a continuation of the precession about $\hat{r}_2$ up to $z_2 + \Delta z$ (Fig. 8.8(c)). Thus the differential precession rule for change of length is $d\hat{s}/dz = \hat{r}_2 \times \hat{s}$.

Now, instead of an increment in length, consider an decrement in frequency.[2] At a first frequency ω, the output polarization from two stages is $\hat{s}_2(\omega)$ (Fig. 8.9(a)). When the frequency is decremented (Fig. 8.9(b)), the precession about $\hat{r}_1$ decreases, terminating the precession about $\hat{r}_1$ early. Accordingly, the radius of the precession circle centered on $\hat{r}_2$ is increased, and the precession angle about $\hat{r}_2$ is decreased due to the lower frequency. The output polarization is $\hat{s}_2(\omega - \Delta\omega)$. For this choice of input polarization state $\hat{s}_{\text{in}}$, the output polarization changes to first-order with frequency, Fig. 8.9(c). This

[2] In the following the frequency is decremented. The choice of increase or decrease in frequency is of no consequence; the decrease in frequency was selected only to make the illustrations clearer.

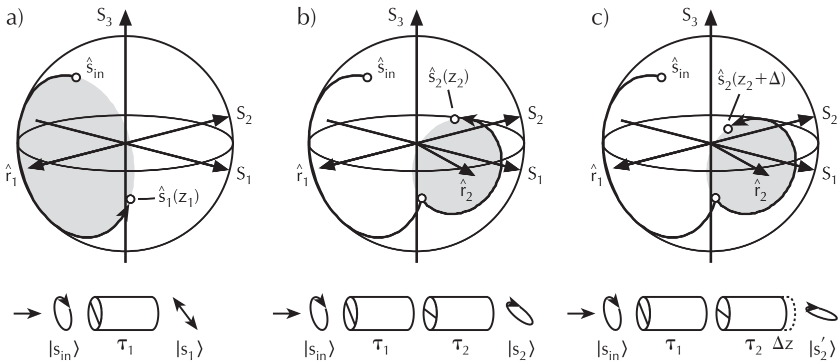

Fig. 8.8. Polarization evolution through one and two birefringent elements. a) Input state $\hat{s}_{\text{in}}$ precesses about birefringent axis $\hat{r}_1$ as the light travels along length z. b) The state output from the first stage is input to the second stage and precesses about $\hat{r}_2$. c) A small increase in length of the second stage continues the polarization precession to state $\hat{s}_2(z_2 + \Delta z)$ from $\hat{s}_2(z_2)$ about $\hat{r}_2$.

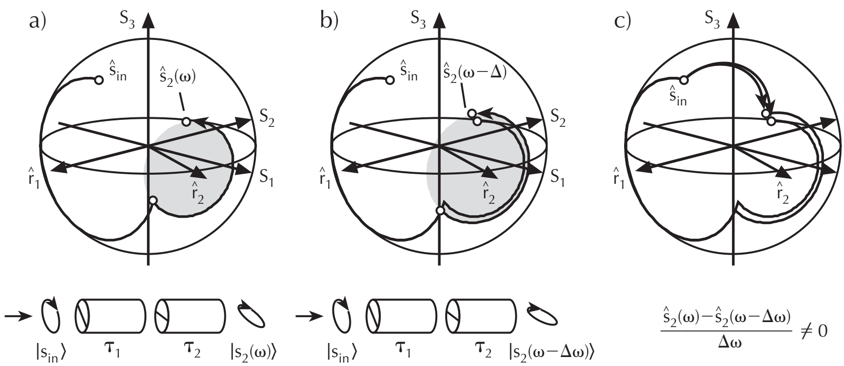

Fig. 8.9. Polarization evolution through two birefringent elements at two different frequencies. a) Evolution through two stages as in Fig. 8.8(b). b) A decrease in frequency reduces the precession through stages one and two. Reduced precession about $\hat{r}_1$ increases the radius of the precession circle centered on $\hat{r}_2$. c) For the same $\hat{s}_{\text{in}}$, the output polarization changes to first-order with frequency.

input state does not yield a principal state of polarization for the system at frequency ω.

A principal state of polarization for this system can be nonetheless found for a different input state.

Figure 8.10 shows the construction necessary for two equal-length stages. At frequency ω, input state $\hat{p}_{\text{in}}$ precesses about $\hat{r}_1$ until it reaches the equator. The polarization state then precesses about $\hat{r}_2$ to output state $\hat{p}_{\text{out}}$. Now, a decrement in frequency moves the intermediate polarization state $\hat{s}_1$ away

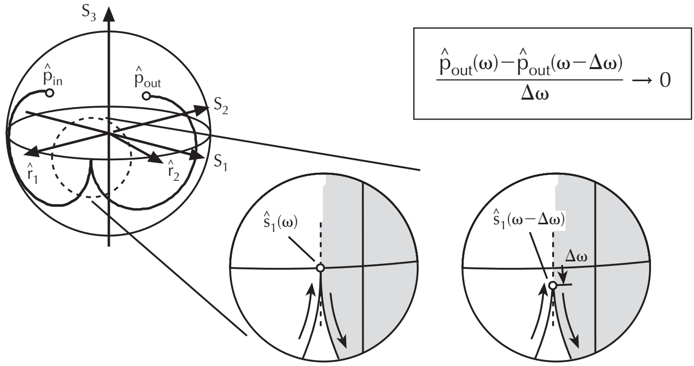

Fig. 8.10. Principal input and output states at ω for two equal-length birefringent stages. The insets show that a decrement in frequency does not change the radius of the precession circle about $\hat{r}_2$ to first order; and that decrease in precession about $\hat{r}_1$ is compensated by the requisite decrease in precession about $\hat{r}_2$ necessary to reach $\hat{p}_{\text{out}}$ to first order.

from the equator. However, as shown in the inset, the left-right motion of the polarization state as it precesses about $\hat{r}_1$ changes only to second order; the same holds for precession about $\hat{r}_2$. This is because the two precession circles share the same tangent line at their intersection. (That the two circles are tangent is the key point, the point of tangency happens to be at the equator in this example.) Accordingly, the precession circle centered on $\hat{r}_2$ does not change radius to first order. Second, the decrease in precession about $\hat{r}_1$ is compensated by a decrease in precession about $\hat{r}_2$ necessary to reach $\hat{p}_{\text{out}}$ at the output, to first order. Neither precession radii nor arc lengths change to first order with change in frequency. Therefore $\hat{p}_{\text{out}}$ is a principal state of the system. The stationary property is expressed as

$$\lim_{\Delta\omega \to 0} \frac{\hat{p}_{\text{out}}(\omega - \Delta\omega) - \hat{p}_{\text{out}}(\omega)}{\Delta\omega} = 0 \tag{8.2.1}$$

The output polarization state $\hat{p}_{\text{out}}$ is an output PSP. Its orthogonal state is also a PSP. The output principle states have corresponding input principal states, one of which is $\hat{p}_{\text{in}}$. The input and output principal states are related by the transformation matrix of the system:

$$|p_{\text{out}}\rangle = T\,|p_{\text{in}}\rangle$$

Unlike an eigenstate of T, the output PSP is generally not the same as the input PSP. However, the two output PSP's are orthogonal to one another, as are the two input PSP's.

Now, since the output PSP is stationary with frequency to first order, a group delay τ can be defined and is separable form the polarization state on which it travels. The existence of this PSP-dependent group delay was first

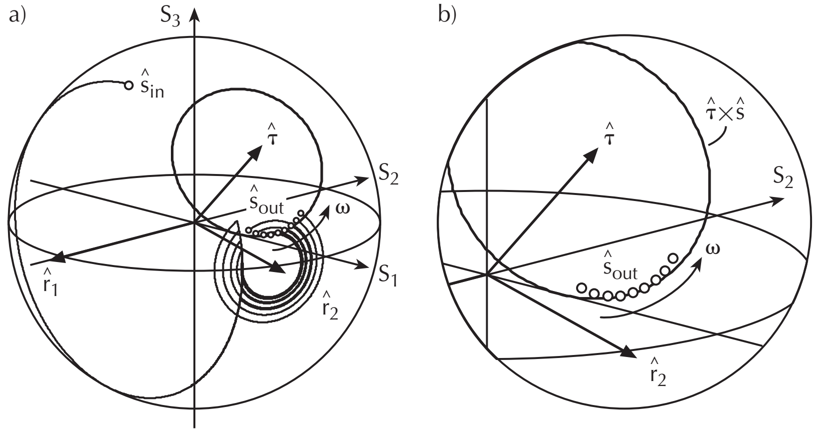

Fig. 8.11. The PMD vector $\vec{\tau}$ defines the precession rule in frequency for output polarizations. a) Evolution of polarization state through two stages from fixed $\hat{s}_{\rm in}$ for range of frequencies. Precession circle $\vec{\tau}\times\hat{s}_{\rm out}(\omega)$ is shown for comparison. b) Magnified view of output polarization state as function of frequency with precession circle. The two deviate only to second order.

experimentally demonstrated in the time-domain by Poole and Giles [51]. The above construction has not determined which of the two output PSP's is the slow axis and which is the fast axis, but nevertheless the frequency derivative of the phase of $\hat{p}_{\rm out}$ imparted by T yields the differential-group delay denoted by τ. The DGD is a length and therefore always a positive quantity. By construction, the PMD vector points along the direction of the slow PSP and has length τ; that is,

$$\vec{\tau} \equiv \tau\hat{p} \tag{8.2.2}$$

where $\hat{p}$ is the slow output PSP. The PMD vector defines the infinitesimal precession rule about which any output polarization state travels for small changes in frequency. That is,

$$\frac{d\hat{s}}{d\omega} = \vec{\tau}\times\hat{s} \tag{8.2.3}$$

where $\hat{s}$ is any polarization state at the output. This frequency-dependent precession rule was used by Curti *et al.* [1] as early as 1987 to measure the differential-group delay of a single-mode fiber.

Figure 8.11 illustrates a calculation of the output polarization states for nine different frequencies for an arbitrary but fixed input polarization state $\hat{s}_{\rm in}$. Overlayed on the figure is the precession circle $\vec{\tau}\times\hat{s}$, where the output PSP is defined by $\hat{p}_{\rm out}$ in Fig. 8.10. It is clear that the output states circle about $\vec{\tau}$ and deviate from the precession circle only to second-order.

That there is deviation at all between the output polarization states and the precession circle underscores a fundamental aspect of PMD. The PMD

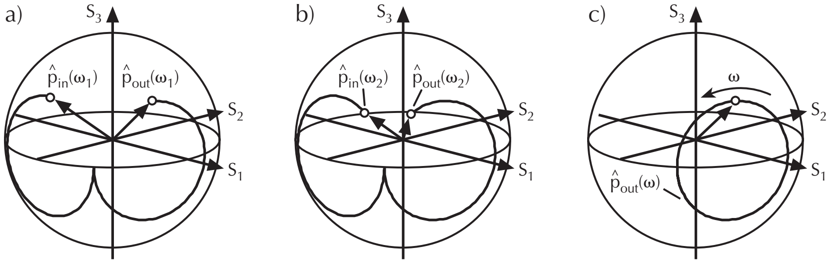

Fig. 8.12. The principal state of polarization changes with frequency. a) At ω_1 input and output PSP's are shown for a two-stage concatenation. b) At different frequency ω_2 new input and output PSP's are required to achieve a stationary output. c) The output PSP spectrum for all frequency. For two stages the spectrum is a circle. This spectrum is periodic.

vector is defined as a vector that is stationary to first order. However, the PMD vector is itself frequency dependent. The vector $\vec{\tau}$ in the infinitesimal precession rule (8.2.3) must be updated for each new frequency. One generally denotes the frequency dependence of the PMD vector as $\vec{\tau}(\omega)$.

Figure 8.12 illustrates the fundamental nature of PMD vector change. The PMD vector construction requires an output polarization that is stationary to first order with frequency for every frequency. Figure 8.12(a) shows the input and output PSP's at ω_1, the same frequency as in Fig. 8.10. However, at a higher frequency ω_2, more precession occurs about the birefringent axes. A new stationary state must account for this increase; this is shown in Figure 8.12(b). A locus of output PSP's is determined by a full frequency sweep. The locus is called the principal-state spectrum. The output PSP spectrum for two equal-length stages is shown in Figure 8.12(c). For two stages the PSP spectrum is always a circle in Stokes space. For more stages the spectrum shows more complicated motion.

That the PMD vector changes pointing direction with frequency is of great practical importance in optical communication systems. Generally there is also an associated change in the DGD with frequency. One should recognize that a principal state of polarization is as delicate and ephemeral as any state of polarization, and will change with any perturbation on the system.

The preceding work is now generalized to show the meaning and behavior of a PMD spectrum associated with any lossless system. Figure 8.13(a) illustrates the PMD vector in Stokes space for a particular frequency, $\vec{\tau}(\omega)$. The vector has a length $\tau(\omega)$ and a pointing direction $\hat{p}(\omega)$. The spectrum of this vector is composed of two components: the PSP spectrum (Fig. 8.13(b)), and the DGD spectrum (Fig. 8.13(c)). The PSP spectrum is a vector spectrum which indicates the principal state as a function of frequency. The DGD spectrum is a positive scalar spectrum which indicates the differential-group delay as a function of frequency. The PMD of any lossless system is entirely determined by these two spectra.

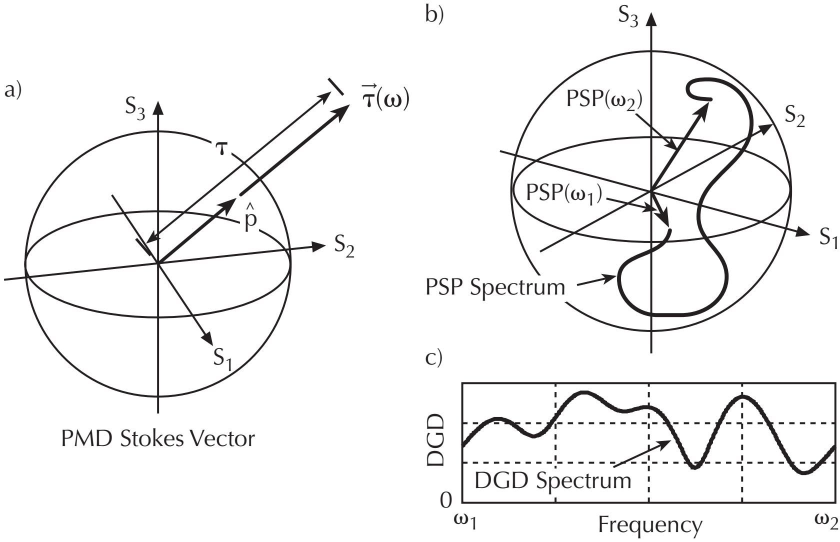

Fig. 8.13. The two components of a PMD spectrum. a) The PMD vector illustrated in Stokes space. The vector has a length τ and a pointing direction $\hat{p}$. These are functions of frequency. b) A PSP spectrum: $\hat{p}(\omega)$. The pointing direction changes with frequency on the unit sphere. c) A DGD spectrum: $\tau(\omega)$. The vector length, which is a positive scalar value, changes with frequency.

The connection between the two defining PMD spectra and the time domain is simple for narrow-band signals. Four narrow-band signals are illustrated in Fig. 8.14(a), having frequency centers at $\omega_1, \ldots, \omega_4$. The overlayed DGD spectrum determines the delay between the two orthogonal polarization components of each signal. For instance, at ω_1, the DGD spectrum has a high value which in turn imparts a large delay between the two components of the associated signal. At ω_4, however, the DGD spectrum has a small value, so the temporal delay for orthogonal components of the associated narrow-band signal is small. Note, however, that the DGD spectrum provides no information as to the relative weight between split components of the signals. That is left to the PSP spectrum.

The effect of the PSP spectrum is illustrated in Fig. 8.14(b). At a particular frequency, the relative weight between orthogonal polarization components is determined by the projection of the input state onto the PSP for that frequency. While the illustration is not rigorous, the projection for the signal centered at ω_1, which has a large differential delay, is about equal. In contrast, the projection for the signal centered at ω_4, which has a small differential delay, is lopsided. The change of relative weights between these two signals is due to the change of the PSP pointing direction with frequency. Note, however, that the PSP spectrum provides not information as to the differential delay between

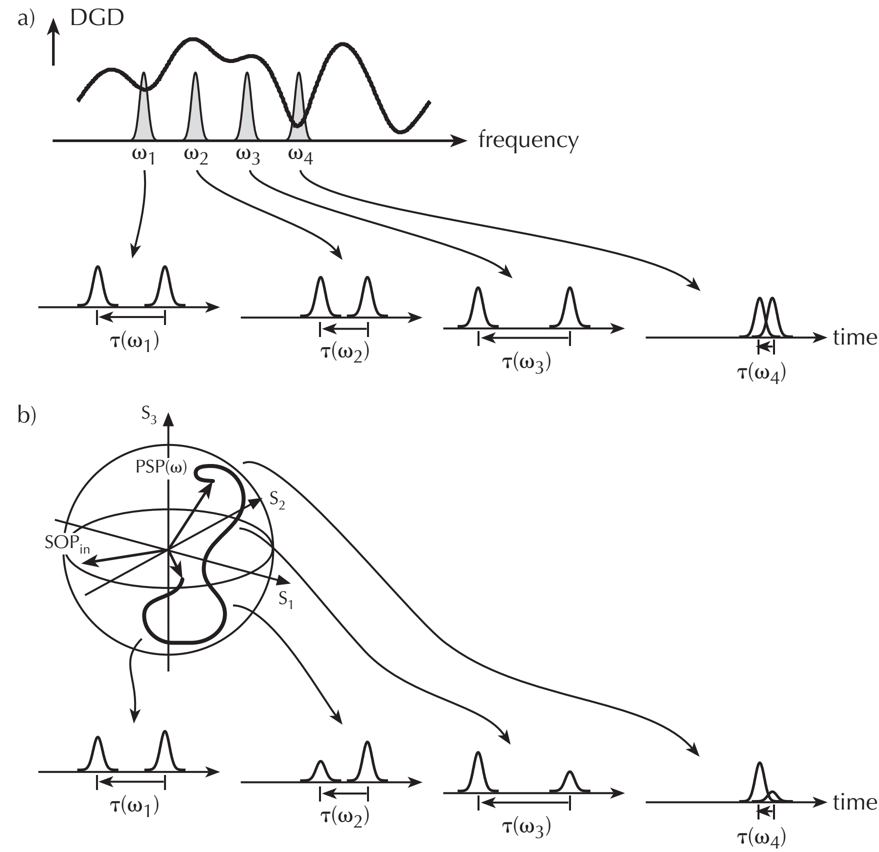

Fig. 8.14. Relation between frequency and time domains for various narrow-band signals affected by PMD in terms of the scalar and vector PMD spectra. a) The DGD spectrum determines the time delay between orthogonal polarization components of a narrow-band signal. For signal centered at ω_1, the time delay is $\tau(\omega_1)$; the figure illustrates a large differential delay. For signal centered at ω_4, the time delay is $\tau(\omega_4)$; the figure illustrates a small differential delay. b) The PSP spectrum determines the relative weight between orthogonal signal components for each frequency. The change in projection between the fixed input polarization and the PSP vector for different frequencies changes the relative weight between orthogonal signal components on each narrow-band signal.

split components of the signals. The DGD and PSP spectra are complementary and the two must be considered together.

More complicated pulse deformations occur when the PMD spectrum varies over the bandwidth of the signal. For the same PMD, the higher the data rate the broader the signal spectrum. In this case each spectral component can be analyzed as in Fig. (8.14), but then the interference between

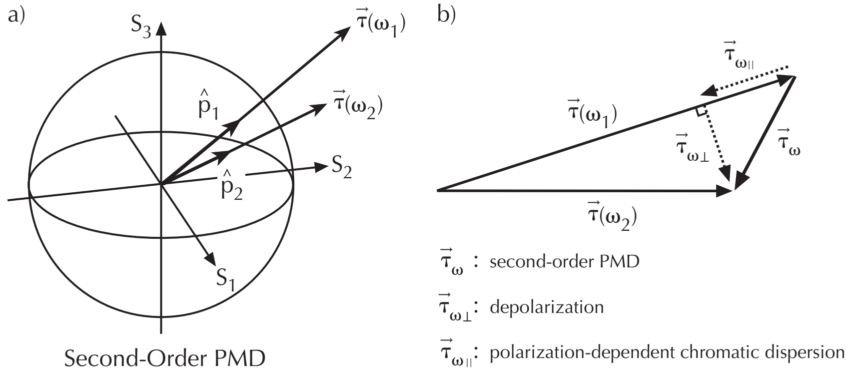

Fig. 8.15. Definition of the second-order PMD (SOPMD) vector $\vec{\tau}_\omega$. a) PMD vectors $\vec{\tau}(\omega_1)$ and $\vec{\tau}(\omega_2)$. The vectors have differential lengths and pointing directions. b) Second-order PMD is the vector difference $\vec{\tau}_\omega = (\vec{\tau}(\omega_2) - \vec{\tau}(\omega_1))/\Delta\omega$ as $\omega_2 - \omega_1 \to 0$. The SOPMD vector can be resolved on the $\vec{\tau}(\omega_1)$ axis into perpendicular and parallel components. The perpendicular component is called depolarization and the parallel component is called polarization-dependent chromatic dispersion (PDCD).

these components must be accounted for. A measure of pulse-deformation complexity is found by the degree of PMD change over a signal bandwidth. If the PMD changes little, then only "first-order" PMD affects the pulse. If the PMD changes a lot, then "higher-order" PMD affects become pronounced. The higher-order PMD effects generate complicated pulse deformations.

One higher-order effect that draws much attention is second-order PMD. Second-order PMD (SOPMD) is a vector generated by a first-order change of the PMD vector with frequency. The SOPMD vector is illustrated in Fig. 8.15. In general, the PMD vector $\vec{\tau}$ is different for different frequencies. Consider two vectors in particular: $\vec{\tau}(\omega_1)$ and $\vec{\tau}(\omega_2)$, where $\omega_2 - \omega_1$ is a small change. The vector difference is the SOPMD vector, denoted by $\vec{\tau}_w$. Like any vector, $\vec{\tau}_w$ has a length and a pointing direction. The pointing direction is generally neither parallel nor perpendicular to either first-order PMD vector, and the length is zero only when $\vec{\tau}$ pirouettes about itself or when there is only one birefringent element. The length of the SOPMD vector in relation to the DGD determines the significance of the second-order vector.

The SOPMD vector is resolved onto two components called the depolarization component and the polarization-dependent chromatic dispersion (PDCD) component (Fig. 8.15(b)). The depolarization component $\vec{\tau}_{\omega\perp}$ runs perpendicular to $\vec{\tau}(\omega_1)$ and indicates the rate which the pointing direction of the PMD vector changes. The PDCD component $\vec{\tau}_{\omega\|}$ runs parallel to $\vec{\tau}(\omega_1)$ and indicates the change of DGD with frequency. The length of the PDCD component is in fact the frequency derivative of the DGD spectrum at ω_1.

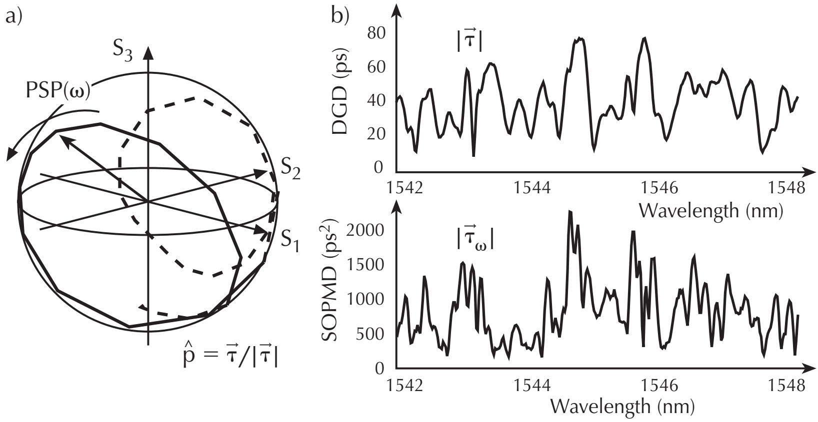

Fig. 8.16. Measurement of PSP, DGD, and magnitude-SOPMD spectra from a line of single-mode fiber. Courtesy D. Peterson, MCI [50].

The term "depolarization" has a connotation in the time domain, so why does SOPMD, best described in the frequency domain, have a depolarization component? The change in pointing direction of the PMD vector with frequency disperses a fixed input polarization into many states at the output. For each frequency there is one output state. When the signal is inverse Fourier transformed into the time domain, the dispersed output polarizations are folded into the time-domain signal so that each time interval contains many polarizations. A time average over these states reduces the degree-of-polarization as compared to the input; that is, the input state is depolarized by the PMD.

Second-order PMD is the frequency derivative of the PMD vector (8.2.2):

$$\vec{\tau}_\omega = \underbrace{\tau_\omega \hat{p}}_{\text{pdcd}} \quad + \quad \underbrace{\tau \hat{p}_\omega}_{\text{depolarization}} \tag{8.2.4}$$

where the two orthogonal vector components are identified in the expression. Statistically the depolarization component dominates the SOPMD vector.

Like first-order PMD, second-order PMD is itself dependent on frequency: $\vec{\tau}_\omega = \vec{\tau}_\omega(\omega)$. SOPMD therefore has a scalar and vector spectrum. Often the magnitude of first- and second-order PMD is plotted when comparing the two orders. Figure 8.16 shows measurements of the PSP, DGD, and magnitude SOPMD $|\vec{\tau}_\omega|$ as a function of frequency made on a line of single-mode fiber.

An alternative but equivalent view of PMD is to look at its response directly in the time domain. To do so, one looks at the impulse response of a cascade of birefringent sections. An impulse in time is a delta function whose

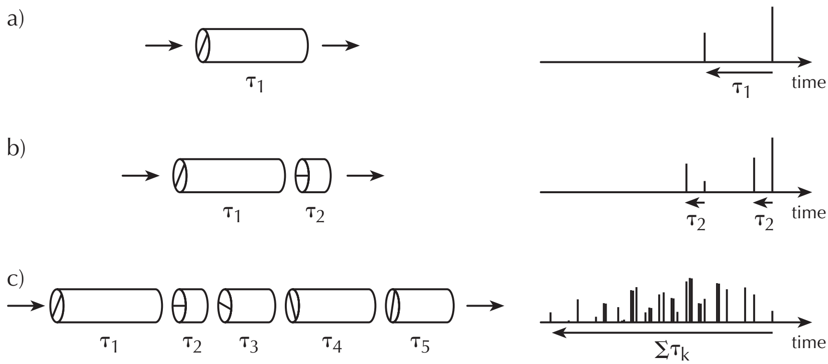

Fig. 8.17. Impulse response from one or more birefringent elements. The polarization of the impulses has been abstracted. a) Impulse response from one stage alone. An input impulse is split along the two birefringent axes and one pulse is delayed with respect to the other by τ_1. b) Impulse response from two stages. The two impulses from the first stage are each divided into two parts, the slow components are then delayed by τ_2. c) Impulse response from five stages generates 2^5 or 32 impulses. There is a first and last pulse, and the time response is FIR.

spectrum is uniform over all frequency. Since no two impulses can overlap unless they are precisely at the same time instant, no accounting for interference is required to determine the output response.

Figure 8.17 illustrates the impulse response from one, two, and five different-length birefringent elements. For one stage, Fig. 8.17(a), an impulse is split along the two birefringent axes and one impulse is delayed with response to the other by τ_1, the DGD of the element. When a second stage is added, Fig. 8.17(b), each impulse from the first stage is projected onto the birefringent axes of the second. The projection on the section stage is independent from the projection onto the first stage, so the relative impulse heights differ. The two components aligned to the slow axis of the second stage are delayed by τ_2, yielding in general four impulses. Note that in the figure the state of polarization for each impulse is not indicated but only its relative weight and time position are shown. Through a cascade of five stages, Fig. 8.17(c), there are five projections and five delays, resulting in 2^5 or 32 distinct impulses. The impulse response can be complicated, but a characteristic of the response is that it is finite in duration. In general, PMD is a linear effect that acts as a finite-impulse response (FIR) filter. The temporal extent of the impulse response is $\sum \tau_k$.

The temporal extent can be compared with a pulse duration of a signal to determine the impairment of PMD. When a signal pulse, such as a non-return-to-zero ONE, is much longer in time than the FIR response of the birefringent cascade, there is little effect of PMD on the pulse. However, when the time extent of the signal pulse and the birefringent cascade are comparable, the signal pulse can be significantly distorted. Strictly, the polarization-dependent

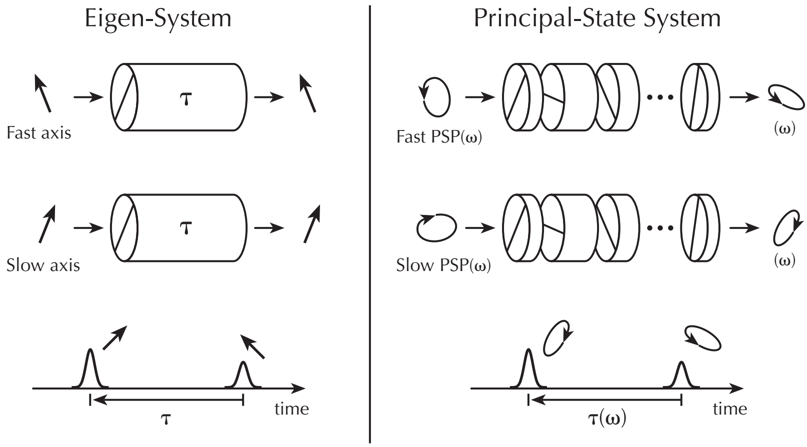

Fig. 8.18. Comparison between eigenstate and principal-state systems. Left: Eigensystem for a single birefringent stage. Input polarizations aligned to the two birefringent axes are output with no change in polarization. Temporally, one axis has a lower group index than the other, so the two eigenstates have a delay τ between them. Right: Principal-state system for an arbitrary birefringent cascade. For any frequency, there exists an input polarization such that the output polarization does not change to first-order in frequency. Along the pair of principal axes, there is a differential-group delay $\tau(\omega)$ between signals launched on orthogonal principal axes. Generally, the PSP and DGD change with frequency.

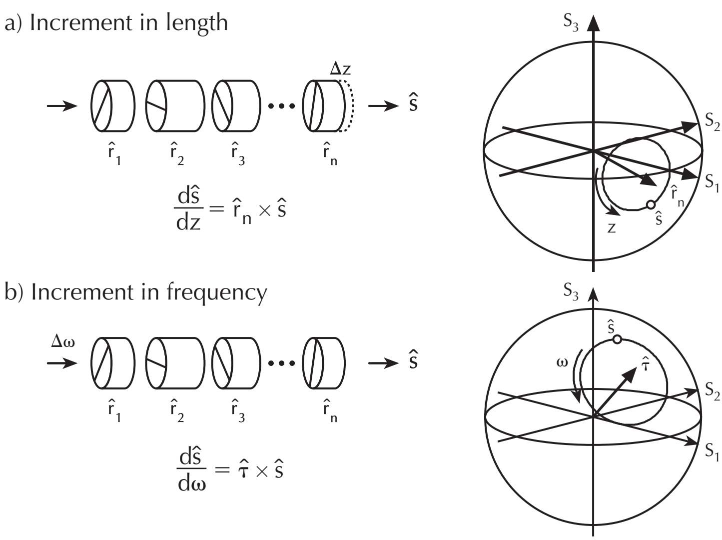

Fig. 8.19. Comparison of infinitesimal rotation for changes in length and frequency. a) Increment in length of the last element. The output polarization precesses about the birefringent axis of the last element. b) Increment in frequency. The output polarization precesses about the principal-state axis of the system.

convolution of the signal pulse with the FIR response of the cascade determines the shape of the output signal. Convolution in time is multiplication in frequency, and the frequency response of PMD has already been heuristically constructed.

To conclude this primer, Fig. 8.18 shows the analogy between the eigenstate system for a single birefringent element and the principal-state system for a birefringent cascade. Figure 8.19 illustrates the two infinitesimal rotation laws that govern PMD, one for change in length and the other for change in frequency. These rotation laws are derived and applied in the next sections in a more rigorous manner.

8.2.2 Fundamental Derivations

The output principal state vector is that vector which is stationary to first order in frequency. The principal state vector is well defined for a lossless birefringent concatenation of any length and composition. The issues at hand are to derive an equation whose eigenvectors are the principal states of the system and whose eigenvalues are the differential group delays along the PSP axes. Together the eigenvectors and values are used to define the PMD vector. This and the following section follows the spin-vector-based derivations set forth by Frigo [16], Gisin [19], and so well elucidated by Gordon and Kogelnik [20].

First, a remark about the analytic tools available for these PMD studies. The introduction of this chapter showed that while a Hermitian matrix fills all 16 entries of the corresponding Mueller matrix, a traceless Hermitian matrix fills only the lower right-hand 3×3 sub-matrix of the Mueller matrix. It was also shown in §2.1 that a unitary matrix also fills only the lower right-hand 3×3 sub-matrix of the corresponding Mueller matrix. Comparison of these matrices shows

$$\underline{H} \mapsto \mathbf{M}_{\underline{H}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix}, \quad \text{and} \quad U \mapsto \mathbf{M}_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix} \qquad (8.2.5)$$

It will be shown that the PMD operator is an identically traceless Hermitian operator. In Stokes space, PMD represents a vector having a length and pointing direction. Unitary matrices act on PMD matrices without scattering Mueller entries outside of the lower 3×3 sub-matrix. The action of a unitary matrix is to rotate the PMD matrix in Stokes space through similarity transforms: $\underline{H}_T = U\underline{H}U^\dagger$. A complete representation of vectors and rotational transformations can therefore be described solely within the O(3) group.

The Hermitian PMD operator and its traceless property is now derived. Table 8.2 lists the symbols used to describe PMD. An output polarization state is related to the input polarization state via the system's transformation matrix T. Recall that for a lossless system, $T = \exp(-j\phi_o)U$, where U is a unitary matrix. An arbitrary input is transformed at the output as

$$|t\rangle = \mathrm{e}^{-j\phi_o} U\, |s\rangle \tag{8.2.6}$$

The frequency derivative is

$$|t\rangle_{\omega} = \left(-j\phi_o' U + U_\omega\right) \mathrm{e}^{-j\phi_o}\, |s\rangle$$

Substitution of (8.2.6) into the frequency derivative yields a expression that relates the output polarization change in frequency to the output polarization itself:

$$|t\rangle_{\omega} = -j\left(\phi_o' + jU_\omega U^\dagger\right)|t\rangle \tag{8.2.7}$$

The derivative of the common phase is a delay, that is, $\phi_o' = \tau_o$. Physically this is the common delay experienced by both polarization states. By analogy one expects $jU_\omega U^\dagger$ to have units of delay as well.

To proceed further, the properties of $jU_\omega U^\dagger$ must be determined. Recall that a unitary matrix is defined as $UU^\dagger = U^\dagger U = I$. Taking the frequency derivative and multiplying both sides by j yields

$$jU_\omega U^\dagger = -jUU_\omega{}^\dagger$$

Notice that

$$\left(jU_\omega U^\dagger\right)^\dagger = -jUU_\omega{}^\dagger$$

Therefore $jU_\omega U^\dagger$ is Hermitian: its eigenvalues are real. The question is whether this Hermitian operator has trace or not. The answer is that $jU_\omega U^\dagger$ is identically traceless. One way to show this is by Taylor expansion. Given that the determinant of a unitary operator is $+1$ for all frequencies, one has

$$\det\left(U(\omega)\right) = \det\left(U(\omega+\delta\omega)\right) = +1$$

The determinant of the Taylor expansion of $U(\omega+\delta\omega)$ must be unity to first order in frequency. For $U(\omega+\delta\omega) \simeq U(\omega) + \delta\omega U_\omega$,

$$\det\left(U(\omega+\delta\omega)\right) = \det\left(I + \delta\omega U_\omega U^\dagger\right)\det\left(U\right)$$

In general, the determinant of a matrix A plus the identity matrix yields

$$\det\left(I + A\right) = 1 + \mathrm{Tr}(A) + \det(A)$$

For the present case,

$$\det\left(I + \delta\omega U_\omega U^\dagger\right) = 1 + \mathrm{Tr}(U_\omega U^\dagger)\delta\omega + \det(U_\omega U^\dagger)(\delta\omega)^2$$

The coefficient of the determinant is second order in frequency while that of the trace is first-order in frequency. In order for the determinant of $U(\omega+\delta\omega)$ to be unity to first-order, the trace must vanish. Therefore

Table 8.2. Symbol Definitions for PMD

Symbol	Expression	Definition
$\vec{\tau}$:	$\vec{\tau} = \tau\hat{p}$	PMD vector in Stokes space. Length τ, pointing direction $\hat{p}$
τ :	$\tau = \partial\varphi/\partial\omega$	Differential-group delay (DGD), $\tau \geq 0$
$\hat{p}, \hat{\tau}$:		Pointing direction of PMD vector, called principal state of polarization (PSP). PSP $\hat{p}$ is defined at output and aligned to slow PSP axis
φ :	$\varphi = \omega\Delta nL/c$	Birefringent phase
	$\varphi = \omega\tau$	As function of DGD
	$\varphi = \beta L$	As function of local birefringence β
$\vec{\tau}_n$:		Local PMD vector of n^{th} element
$\vec{\tau}(n)$:		Cumulative PMD vector through n^{th} element
$\langle\tau\rangle$:	$\langle\lvert\vec{\tau}(\omega)\rvert\rangle$	Mean PMD. Frequency average of DGD
$\langle\tau^2\rangle$:	$\langle\vec{\tau}(\omega)\cdot\vec{\tau}(\omega)\rangle$	Mean-square PMD. Frequency average of DGD^2

$$\mathrm{Tr}\left(jU_\omega U^\dagger\right) = 0 \tag{8.2.8}$$

A Hermitian matrix with zero trace has important properties. First, its eigenvalues are equal in magnitude and opposite in sign. Second, the SU(2) matrix is equivalent to a vector in O(3). That vector is determined by the eigenvalue equation for $jU_\omega U^\dagger$:

$$jU_\omega U^\dagger \,|p_\pm\rangle = \pm\,\lambda\,|p_\pm\rangle \tag{8.2.9}$$

where $|p_\pm\rangle$ are the eigenvectors of $jU_\omega U^\dagger$. The eigenvalues $\pm\lambda$ are defined as $\lambda \equiv \pm\,\tau/2$, and thus

$$jU_\omega U^\dagger \,|p_\pm\rangle = \pm\,\tau/2\,|p_\pm\rangle \tag{8.2.10}$$

Since the determinant is the product of the eigenvalues, $\det\left(jU_\omega U^\dagger\right) = -\tau^2/4$. Moreover, the determinant of a product of matrices is the product of the determinants, so the eigenvalues are

$$\tau = 2\sqrt{\det U_\omega} \tag{8.2.11}$$

Combining the eigenvalue equation (8.2.9) with (8.2.7), and choosing an output polarization along an eigenvector of $jU_\omega U^\dagger$, one finds

$$\begin{aligned} |p_\pm\rangle_\omega &= -j\left(\phi_o' + jU_\omega U^\dagger\right)|p_\pm\rangle \\ &= -j\left(\tau_o \pm \tau/2\right)|p_\pm\rangle \end{aligned}$$

The total group delay of the signals along the principal state axes is

$$\tau_g = \tau_o \pm \tau/2 \tag{8.2.12}$$

where τ_o is the common delay and $\pm\tau/2$ is the differential delay. The slow principal state is $|p_+\rangle$, with corresponding delay $\tau_o + \tau/2$, while the fast principal state is $|p_-\rangle$, with corresponding delay $\tau_o - \tau/2$. In the following, $|p_+\rangle$ is denoted simply by $|p\rangle$.

To verify that the output polarization $|p\rangle$ is stationary to first order for either principal state, the Jones vector is converted to a Stokes vector as

$$\begin{aligned}\vec{p}_\omega &= \langle p_\omega \,|\vec{\sigma}|\, p_\omega\rangle \\ &= j\tau_g \langle p\,|\vec{\sigma}|\,p\rangle - j\tau_g \langle p\,|\vec{\sigma}|\,p\rangle\end{aligned}$$

which is evidently zero.

In summary, the eigenvectors of $jU_\omega U^\dagger$ are the principal states of polarization of the system, and the eigenvalues are the differential group delays along the two axes. The PMD vector is defined to point along the slow output principal state and have a length of the total differential-group delay:

$$\vec{\tau} \equiv \tau\hat{p}, \quad \hat{p} \text{ is slow output PSP} \tag{8.2.13}$$

The Stokes vector interpretation of (8.2.13) is illustrated in Fig. 8.13(a).

8.2.3 Connection Between Jones and Stokes Space

That $jU_\omega U^\dagger$ is traceless Hermitian and has zero trace implies a connection between the SU(2) Jones space and the O(3) Stokes space (cf. §2.6). In particular, observe that for a Stokes vector $\vec{\tau}$, the following two eigenvalue equations are equal to within a factor of two:

$$\begin{aligned} jU_\omega U^\dagger\,|p\rangle &= (\tau/2)\,|p\rangle \\ (\vec{\tau}\cdot\vec{\sigma})\,|p\rangle &= \tau\,|p\rangle \end{aligned}$$

The spin-vector operator and $jU_\omega U^\dagger$ are clearly related:

$$jU_\omega U^\dagger = \frac{1}{2}\,(\vec{\tau}\cdot\vec{\sigma}) \tag{8.2.14}$$

The spin-vector can be used to determine how an arbitrary output polarization state changes with frequency. In Stokes space, the frequency change is

$$\hat{t}_\omega = (\langle t\,|_\omega)|\vec{\sigma}\,|t\rangle + \langle t\,|\,\vec{\sigma}(|t\rangle_\omega) \tag{8.2.15}$$

Employing the spin-vector form (8.2.14) and the eigenvalue equation (8.2.7), (8.2.15) generates

$$\begin{aligned}\hat{t}_\omega &= j\langle t\left|\left(\tau_o + \frac{1}{2}\,(\vec{\tau}\cdot\vec{\sigma})\right)\vec{\sigma}\right|t\rangle - j\langle t\left|\vec{\sigma}\left(\tau_o + \frac{1}{2}\,(\vec{\tau}\cdot\vec{\sigma})\right)\right|t\rangle \\ &= \frac{j}{2}\langle t\left|(\vec{\tau}\cdot\vec{\sigma})\,\vec{\sigma} - \vec{\sigma}\,(\vec{\tau}\cdot\vec{\sigma})\right|t\rangle \\ &= \langle t\left|\vec{\tau}\times\vec{\sigma}\right|t\rangle\end{aligned}$$

and thus

$$\frac{d\hat{t}}{d\omega} = \vec{\tau}\times\hat{t} \tag{8.2.16}$$

This is the infinitesimal frequency-change rule for an arbitrary output polarization state. The output state precesses about the PMD vector $\vec{\tau}$. Only if the output state is aligned or anti-aligned along $\vec{\tau}$ will its state not change with frequency. Illustrations of the precession rule in frequency are given by Figs. 8.11 and 8.19(b).

The PMD vector $\vec{\tau}$ is defined at the output of the system. There is a corresponding PMD vector at the input of the system. Denote $\vec{\tau}_t$ and $\vec{\tau}_s$ as the output and input PMD vectors, respectively. Since in general the output density matrix D_t of a unitary transformation is related to the input density D_s via $D_t = UD_sU^\dagger$, and the density operator is related to the spin-vector through (2.5.29) on page 56, the relation between input and output spin vectors is

$$(\vec{\tau}_t\cdot\vec{\sigma}) = U\,(\vec{\tau}_s\cdot\vec{\sigma})\,U^\dagger$$

Isolating $(\vec{\tau}_s\cdot\vec{\sigma})$ and substitution of (8.2.14) yields

$$\frac{1}{2}\,(\vec{\tau}_s\cdot\vec{\sigma}) = jU^\dagger U_\omega \tag{8.2.17}$$

It makes sense that the input PMD vector is determined by writing the unitary matrices U_ω and $U^\dagger$ in reverse order to that for the output PMD vector. Moreover, as the operators are unitary, the DGD for both the input and output PMD vectors is identical.

For the unitary transformation in Jones space $|t\rangle = T\,|s\rangle$, the equivalent transformation in Stokes space is

$$\hat{t} = R\hat{s} \tag{8.2.18}$$

where a vector form of the unitary operator R is given in (2.6.25) on page 68, repeated here due to its significance:

$$R = (\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) - \cos\varphi(\hat{r}\times)(\hat{r}\times) \tag{8.2.19}$$

The precession vector $\hat{r}$ of operator R points along an eigenvector of U and the precession angle $\varphi = \omega\tau$ is the angle of the complex eigenvalue of U.

A expression analogous to $jU_\omega U^\dagger$ exists in Stokes space. Consider a frequency change for (8.2.18):

$$\hat{t}_\omega = R_\omega\hat{s}$$

where the input polarization state is fixed in frequency. Substitution of $\hat{s} = R^\dagger \hat{t}$ gives

$$\hat{t}_\omega = R_\omega R^\dagger \hat{t}$$

By comparison with the previously derived infinitesimal rotation rule (8.2.16) one is led to the identification that

$$\vec{\tau}\times = R_\omega R^\dagger \tag{8.2.20}$$

The value of $R_\omega R^\dagger$ will be clear when deriving the PMD concatenation rules. Yet even at this point its use is significant. Consider again the input and output PMD vectors $\vec{\tau}_s$ and $\vec{\tau}_t$. It was already determined by (8.2.17) that the lengths of these two vectors are identical. Since in general $\hat{t} = R\hat{s}$, one can choose the input and output polarizations parallel to the PSP's of $jU_\omega U^\dagger$. Accordingly,

$$\vec{\tau}_t = R\,\vec{\tau}_s \tag{8.2.21}$$

That is, the input and output PMD vectors are related by R. How is the second-order PMD component, $\vec{\tau}_{t\omega}$, related to $\vec{\tau}_{s\omega}$? Taking the frequency derivative of (8.2.21),

$$\vec{\tau}_{t\omega} = R_\omega\,\vec{\tau}_s + R\,\vec{\tau}_{s\omega}$$

along with the substitution of (8.2.21) and (8.2.20) yields

$$\begin{aligned}\vec{\tau}_{t\omega} &= \vec{\tau}\times\vec{\tau} + R\,\vec{\tau}_{s\omega}\\ &= R\,\vec{\tau}_{s\omega}\end{aligned} \tag{8.2.22}$$

Both the first- and second-order PMD vectors transform from input to output through R. Higher-order frequency derivatives quickly become more complex.

The spin-vector form $(\vec{\tau}\cdot\vec{\sigma})$ also assists in the evaluation of the three Stokes components of the vector $\vec{\tau}$. In particular, (2.5.29) on page 56 gives

$$\vec{\tau}\cdot\vec{\sigma} = \begin{pmatrix} \tau_1 & \tau_2 - j\tau_3 \\ \tau_2 + j\tau_3 & -\tau_1 \end{pmatrix} \tag{8.2.23}$$

Likewise, (8.2.20) is identified with (2.6.29) on page 70 to gives

$$\vec{\tau}\times = \begin{pmatrix} 0 & -\tau_3 & \tau_2 \\ \tau_3 & 0 & -\tau_1 \\ -\tau_2 & \tau_1 & 0 \end{pmatrix} \tag{8.2.24}$$

When the unitary matrix is written in Cayley-Klein form, given by (2.4.28) on page 51, the matrix entries of $jU_\omega U^\dagger$ are identified with (8.2.23) as

$$\tau_1 = 2j\,(a_\omega a_\omega^* + b_\omega b_\omega^*) \tag{8.2.25a}$$

$$\tau_2 = 2\,\Im\mathrm{m}\,(a_\omega b - b_\omega a) \tag{8.2.25b}$$

$$\tau_3 = 2\,\Re\mathrm{e}\,(a_\omega b - b_\omega a) \tag{8.2.25c}$$

The DGD τ can be determined either from $\tau^2 = \tau_1^2 + \tau_2^2 + \tau_3^2$ or from (8.2.11). In either case the resulting expression is

$$\tau = 2\sqrt{a_\omega a_\omega^* + b_\omega b_\omega^*} \tag{8.2.26}$$

In comparison with (8.2.26) it is interesting to note that $\sqrt{aa^* + bb^*} = 1$. The frequency derivative of the Jones matrix elements brings down group-delay coefficients which are responsible for the differential-group delay of the system.

Finally, the spin-vector form is used to relate the eigenvector of U to the output PMD vector. The exponential form of the unitary operator

$$U = \exp\left[-j\varphi\left(\hat{r}\cdot\vec{\sigma}\right)/2\right] \tag{8.2.27}$$

has a frequency derivative of (cf. (2.6.19) on page 68)

$$U_\omega = -j\left(\varphi_\omega/2\right)\left(\hat{r}\cdot\vec{\sigma}\right)U - j\left(\hat{r}_\omega\cdot\vec{\sigma}\right)\sin\left(\varphi/2\right)$$

The product $jU_\omega U^\dagger$ is

$$jU_\omega U^\dagger = \varphi_\omega/2\left(\hat{r}\cdot\vec{\sigma}\right) + \left(\hat{r}_\omega\cdot\vec{\sigma}\right)\sin\left(\varphi/2\right)\ \left[I\cos\left(\varphi/2\right) + j\left(\hat{r}\cdot\vec{\sigma}\right)\sin\left(\varphi/2\right)\right]$$

Identification with (8.2.14) and use of spin-vector product identity (2.5.38) on page 57 generates the relation between the eigenvector $\hat{r}$ of U and the PMD vector $\vec{\tau}$ of $jU_\omega U^\dagger$:

$$\vec{\tau} = \varphi_\omega\hat{r} + \sin\varphi\,\hat{r}_\omega - \left(1-\cos\varphi\right)\,\hat{r}_\omega\times\hat{r} \tag{8.2.28}$$

As discussed in §8.2.1, the eigenvector of U generally changes to first order with frequency. That first-order effect is accounted for by $\hat{r}_\omega$ in (8.2.28). An important special case is when the PMD vector refers to a single homogeneous birefringent section. In this case $\hat{r}_\omega = 0$ and the PMD vector of the element is aligned to the birefringent axis: $\vec{\tau} = \varphi_\omega\hat{r}$. The frequency derivative of the phase is the group delay of the element, or $\varphi_\omega = \tau$.

8.2.4 Concatenation Rules for PMD

The PMD concatenation rules are very helpful to gain physical insight and intuition on how PMD behaves. The concatenation rules track the PMD vectors themselves as more and more sections are added to a system. Early work on PMD statistics clearly deals with the concatenation of many (indeed infinitely many) birefringent sections, but the discussion of PMD statistics is left for the next chapter. The four works from the technical literature that present concatenation rules are from Curti *et al.* [4], who presented a two-stage concatenation and numerically extended the work to large numbers; Gisin [19], who derived the recurrence relation for an arbitrary number of sections; Karlsson [38], who recast Gisin's recurrence relation in spin-vector form; and and Gordon and Kogelnik [20], who extended the prior work.

The concatenation rules are rules for adding two or more output PMD vectors. The rules could be written for the input PMD vectors just as well, but they are not for a mixture of input and output vectors.

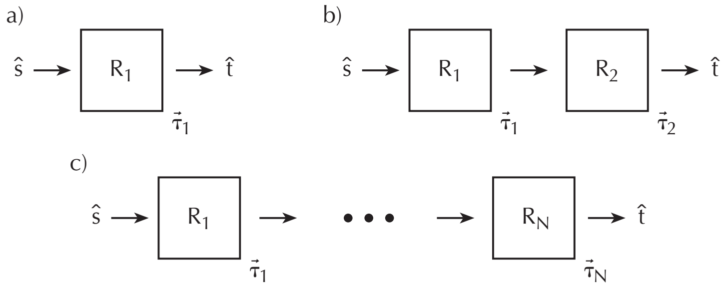

Fig. 8.20. Block diagrams of one, two, and N birefringent blocks in concatenation. It is assumed that the blocks are lossless. A birefringent block may have heterogeneous or homogeneous birefringence.

One Birefringent Section

Without even one section of birefringence there is no PMD at all. The PMD vector first appears after a single homogeneous birefringent section. The relation between a single birefringent element and the PMD vector is given in (8.2.28), where $\hat{r}_\omega = 0$. Thus,

$$\vec{\tau} = \tau\hat{r} \tag{8.2.29}$$

where $\hat{r}$ points in the direction of the slow birefringent axis of the element. The DGD is given by $\tau = \varphi_\omega$. $\vec{\tau}$ is the first-order PMD vector for the section. The second order vector is

$$\vec{\tau}_\omega = 0 \tag{8.2.30}$$

There is no second-order frequency dependence of the PMD vector. This is a unique case for PMD, as concatenations with two or more stages will always have finite $\vec{\tau}_\omega$ except, possibly, at particular frequencies when the second-order vector momentarily vanishes.

One Birefringent Block

A birefringent block is distinguished from a birefringent section in that the latter is homogeneous birefringence, i.e. with no intermediate polarization mode coupling, while the former can be birefringence of any composition and inhomogeneity. Any output polarization is related to the input polarization by $\hat{t} = R_1\hat{s}$, while the first- and second-order PMD of the block is denoted simply by $\vec{\tau}_1$ and $\vec{\tau}_{1\omega}$ (Fig. 8.20(a)). The PMD vector can be determined from $\vec{\tau}_1 \times = R_{1\omega}R_1{}^\dagger$.

Two Birefringent Blocks

Denote the PMD vector generated by a first birefringent block as $\vec{\tau}_1$ and by a second birefringent block as $\vec{\tau}_2$ (Fig. 8.20(b)). When these two blocks are concatenated the overall input-to-output transformation is $R = R_2 R_1$. The cumulative PMD vector $\vec{\tau}$ is related to the block transfer operators as

$$\begin{aligned}\vec{\tau}\times &= (R_2 R_1)_\omega (R_2 R_1)^\dagger \\ &= R_{2\omega} R_1 {R_1}^\dagger {R_2}^\dagger + R_2 R_{1\omega} {R_1}^\dagger {R_2}^\dagger \\ &= \vec{\tau}_2 \times + R_2 (\vec{\tau}_1 \times) {R_2}^\dagger \\ &= \vec{\tau}_2 \times + (R_2 \vec{\tau}_1) \times\end{aligned}$$

The last line is derived from the preceding one as $RvR^\dagger$ is a unitary transformation on v. The embedded expression for first-order PMD is

$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1 \tag{8.2.31}$$

The second-order expression comes from the frequency derivative

$$\vec{\tau}_\omega = \vec{\tau}_{2\omega} + R_2 \vec{\tau}_{1\omega} + R_{2\omega} \vec{\tau}_1$$

Using (8.2.31) to write $\vec{\tau}_1 = {R_2}^\dagger (\vec{\tau} - \vec{\tau}_2)$, the second-order expression reduces to

$$\vec{\tau}_\omega = \vec{\tau}_{2\omega} + R_2 \vec{\tau}_{1\omega} + \vec{\tau}_2 \times \vec{\tau} \tag{8.2.32}$$

The second-order expression for $\vec{\tau}_\omega$ is almost like that for $\vec{\tau}$ except for the additional $\vec{\tau}_2 \times \vec{\tau}$ on the right-hand side. The additional vector generates a pulling of the second-order cumulative vector $\vec{\tau}_\omega$ in a direction defined by $\vec{\tau}_2 \times \vec{\tau}$, which is orthogonal to both the local PMD component $\vec{\tau}_2$ and the cumulative first-order vector $\vec{\tau}$.

Since these derivations have applied to heterogeneous birefringent blocks, the block PMD vectors and the transformation operators R are generally a function of frequency. Accordingly, the first-order cumulative vector is more accurately written as

$$\vec{\tau}(\omega) = \vec{\tau}_2(\omega) + R_2(\omega) \vec{\tau}_1(\omega) \tag{8.2.33}$$

For small frequency changes, (8.2.33) establishes a precession rule. The PMD vector $\vec{\tau}_1$ precesses about the axis $\hat{r}_2$ of the second PMD block through angle φ_2. This is the effect of R_2 on $\vec{\tau}_1$. The rotated vector is then added to $\vec{\tau}_2$. As Gordon and Kogelnik point out, "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 the vectors" [20].

A significant special case of (8.2.33) is when R_2 and R_1 refer to homogeneous birefringent sections. In this case there is no frequency dependence of $\hat{r}_{1,2}$ or $\vec{\tau}_{1,2}$, and the precession behavior is more clear since the birefringent axes are fixed. The two-section precession rule (8.2.31) is expanded to

$$\vec{\tau} = \vec{\tau}_2 + \Big[(\hat{r}_2 \hat{r}_2 \cdot) + \sin\varphi_2 (\hat{r}_2 \times) - \cos\varphi_2 (\hat{r}_2 \times)(\hat{r}_2 \times) \Big] \vec{\tau}_1$$

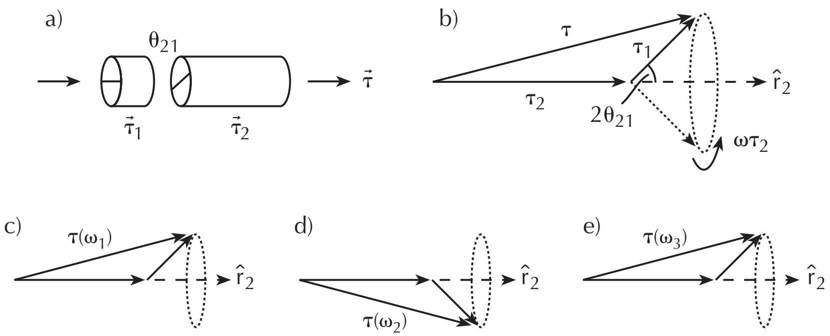

Fig. 8.21. PMD concatenation rule for two homogeneous birefringent sections. a) Concatenation of two birefringent sections having section DGD's of τ_1 and τ_2, and relative angle between birefringent axes θ_{21}. b) PMD vector addition. $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with retardance $\omega\tau_2$. The cumulative PMD vector $\vec{\tau}$ is the vector sum of the components. The pointing direction of $\vec{\tau}$ is the output PSP of the cascade. c–e) Component PMD vectors at three different frequencies. The PSP spectrum is periodic and the DGD spectrum is constant.

where $\hat{r}_2$ points in the direction of $\vec{\tau}_2$ and φ_2 is the birefringent phase of the second section. This motion and the associated physical construct is illustrated in Fig. 8.21(a,b).

Two homogeneous birefringent sections with mode-mixing at a well-defined junction are illustrated in Fig. 8.21(a). The mode-mixing angle is the difference between the angles of the birefringent axes of the two sections. The two-section concatenation rule is illustrated in Fig. 8.21(b), which is drawn in Stokes space. The base of $\vec{\tau}_1$ is jointed to the tip of $\vec{\tau}_2$. The angle between the two vectors is $2\theta_{21}$, twice the physical angle at the mode-mixing junction; this angle is fixed in frequency. When frequency is changed, the PMD vector of the first section $\vec{\tau}_1$ precesses about the axis of the second PMD vector, $\hat{r}_2$. The angle of precession is $\omega\tau_2$, which is solely dictated by the length of the second PMD vector. The precession rate is τ_2. Over all frequencies the tip of $\vec{\tau}_1$ traces a circle; the circular motion is periodic with a free-spectral range of $\mathrm{FSR} = 1/\tau_2$. Figures 8.21(c–e) illustrate this motion. At a first frequency the cumulative PMD vector $\vec{\tau}$ points upward; at a second frequency it points downward; and at a third frequency it points up again. Since $\vec{\tau}$ points in the direction of the slow output PSP, the circle traced by $\vec{\tau}_1$ in frequency is the PSP spectrum of the two-section concatenation. The length of the cumulative vector $\vec{\tau}$ is the vector sum of the components. In this case there are only two fixed-length components, so $\vec{\tau}$ completes the triangle rule and remains constant in length over frequency.

N Birefringent Blocks

The concatenation rule for N birefringent blocks is derived from repeated application of (8.2.31) and (8.2.32). Denoting $\vec{\tau}_k$ the PMD vector of the k^{th} block and $\vec{\tau}(k)$ the cumulative PMD vector through the k^{th} block, the cumulative first- and second-order PMD vectors are boot-strapped from the origin

$$
\begin{aligned}
&\vec{\tau}(1) = \vec{\tau}_1 && \vec{\tau}_\omega(1) = 0 \\
&\vec{\tau}(2) = \vec{\tau}_2 + R_2\vec{\tau}(1) && \vec{\tau}_\omega(2) = \vec{\tau}_2 \times \vec{\tau}(2) \\
&\vec{\tau}(3) = \vec{\tau}_3 + R_3\vec{\tau}(2) && \vec{\tau}_\omega(3) = \vec{\tau}_3 \times \vec{\tau}(3) + R_3\vec{\tau}_\omega(2) \\
&\qquad\vdots && \qquad\vdots \\
&\vec{\tau}(n) = \vec{\tau}_n + R_3\vec{\tau}(n-1) && \vec{\tau}_\omega(n) = \vec{\tau}_n \times \vec{\tau}(n) + R_n\vec{\tau}_\omega(n-1)
\end{aligned}
\tag{8.2.34}
$$

Another form for $\vec{\tau}(n)$ is to write recursively

$$\vec{\tau}(n) = \vec{\tau}_n + R_n\left(\vec{\tau}_{n-1} + R_{n-1}\left(\vec{\tau}_{n-2}\cdots R_3\left(\vec{\tau}_2 + R_2\vec{\tau}_1\right)\right)\cdots\right) \tag{8.2.35}$$

This form gives some physical insight. Starting from the beginning of the cascade, component vector $\vec{\tau}_1$ is placed on the tip of $\vec{\tau}_2$ and precesses about its axis at rate τ_2. This is the action of R_2. Together these two components are placed on the tip of $\vec{\tau}_3$ and they precess about the $\hat{r}_3$ axis at rate τ_3. This is the action of R_3. Keep in mind that while $\vec{\tau}_2$ and $\vec{\tau}_1$ precess about $\vec{\tau}_3$, $\vec{\tau}_1$ continues to precess about $\vec{\tau}_2$. The procedure is repeated through the n^{th} section.

Figure 8.22 illustrates the precession for three homogeneous birefringent sections in cascade. Unlike the two-section case, the motion of the tip of $\vec{\tau}_1$ is more complicated, tracing a “folded-eight” curve in Stokes space and having a frequency-dependent vector sum. The vector sum $\vec{\tau}$ is the cumulative PMD vector of the cascade.

Finally, a compact form of the cumulative first- and second-order PMD vectors is written as

$$\vec{\tau} = \sum_{k=1}^{n} R(n, k+1)\,\vec{\tau}_n \tag{8.2.36}$$

and

$$\vec{\tau}_\omega = \sum_{k=1}^{n} R(n, k+1)\left(\vec{\tau}_{n\omega} + \vec{\tau}_n \times \vec{\tau}(n)\right) \tag{8.2.37}$$

where

$$R(n, k) = R_n R_{n-1}\cdots R_k \tag{8.2.38}$$

Note that the evaluation of $\vec{\tau}_\omega$ requires the concurrent evaluation of $\vec{\tau}$. Expressions (8.2.36-8.2.37) offer a fast way to evaluate numerically the first- and second-order PMD vectors for an arbitrary cascade. The evaluation occurs directly in Stokes space and $\vec{\tau}_\omega$ is determined without a numerical derivative.

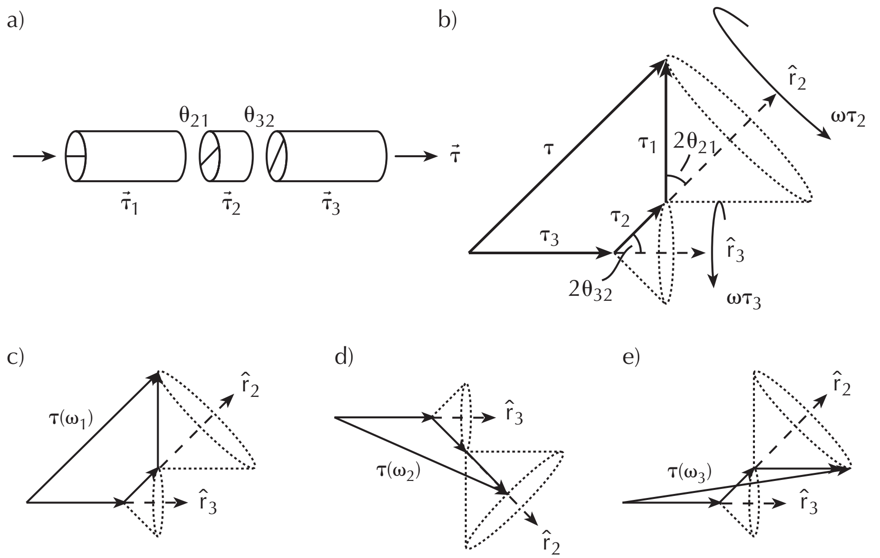

Fig. 8.22. PMD concatenation rule for three homogeneous birefringent sections. a) Concatenation of three birefringent sections having section DGD's of $\tau_{1,2,3}$ and relative angle between birefringent axes θ_{21} and θ_{32}. The drawing is for $\tau_1 = \tau_3 = 2\tau_2$. b) PMD vector addition. $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with retardance $\omega\tau_2$; combined, the vectors precess about $\vec{\tau}_3$ with retardance $\omega\tau_3$. The motion at the tip of $\vec{\tau}_1$ is more complicated than the two-section case, and the length of the cumulative PMD vector now changes with frequency. c–e) Component PMD vectors at three different frequencies. The PSP spectrum is periodic but more complicated than for two sections. The DGD spectrum varies with frequency and is also periodic.

An alternative method, the multiplication of Jones matrices where each matrix represents a homogeneous element and conversion to PMD vectors via (8.2.25-8.2.26), requires substantially more computations and requires evaluation at two frequencies in order to estimate $\vec{\tau}_\omega$.

8.2.5 PMD Evolution Equations

The PMD concatenation equations developed in the preceding abstracted the underlying birefringence and focused solely on how PMD vectors combine in length and precess in frequency. Studies of optical pulse propagation in fiber and statistical properties of PMD require an explicit connection between the cumulative PMD vector and the underlying local birefringence. There are at least two ways to derive the PMD evolution equation, one due to Poole *et al.* [53], and another due to Gisin *et al.* [19] and Gordon *et al.* [20]. The latter derivation type makes the connection between the PMD concatenation rules and the underlying birefringence. The Poole derivation is detailed below.

A note on notation. The PDL evolution equations denote $\vec{\alpha}(z)$ as the local PDL vector $\vec{\alpha}$ per unit length (cf. §8.1.4). For PMD, the local birefringence per unit length is customarily denoted β (as in $\exp(-j\beta z)$ where $\beta = \omega\Delta n/c$). The connection to the local PMD vector $\vec{\tau}_\ell$ is $\vec{\tau}_\ell = \vec{\beta}_\omega L$ where L is the length of the segment.

The Poole derivation starts with the established precession rules in length and frequency for an arbitrary polarization state:

$$\frac{\partial \hat{s}}{\partial z} = \vec{\beta} \times \hat{s}, \quad \text{and} \quad \frac{\partial \hat{s}}{\partial \omega} = \vec{\tau} \times \hat{s}$$

where $\vec{\beta}$ is the local birefringence vector and $\vec{\tau}$ is the cumulative PMD vector up to location z. Taking the frequency derivative of the first equation and the length derivative of the second gives

$$\frac{\partial^2 \hat{s}}{\partial z \partial \omega} = \vec{\beta}_\omega \times \hat{s} + \vec{\beta} \times \hat{s}_\omega$$
$$\frac{\partial^2 \hat{s}}{\partial \omega \partial z} = \vec{\tau}_z \times \hat{s} + \vec{\tau} \times \hat{s}_z$$

Under the assumption of continuity of the function $\hat{s}(z,\omega)$, the left-hand sides are equal. Using the vector identity $(\vec{\beta} \times \vec{\tau}) \times \hat{s} = \vec{\beta} \times (\vec{\tau} \times \hat{s}) - \vec{\tau} \times (\vec{\beta} \times \hat{s})$ results in the PMD evolution equation

$$\frac{\partial \vec{\tau}}{\partial z} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau} \tag{8.2.39}$$

The local birefringence changes the cumulative PMD vector in both an additive and multiplicative sense: additive through $\vec{\beta}_\omega$ and multiplicative through $|\beta\tau|$. Moreover, $\vec{\beta}$ drives the average direction of $\vec{\tau}_z$ while $\vec{\beta} \times \vec{\tau}$ drives $\vec{\tau}_z$ in a perpendicular direction.

Figure 8.23 illustrates the motion of $\vec{\tau}$ through two birefringent sections, where $\tau(z=0) = 0$. Through the first section the cross-product vanishes, so $\vec{\tau}$ directly tracks β_1. However, once the light enters the second section, the local birefringence and cumulative PMD vector are no longer parallel. The motion of the PMD vector is helical about a center axis $\vec{\beta}_2$. The $\vec{\beta}_{2\omega}$ term pulls the average direction of $\vec{\tau}$ along $\vec{\beta}_2$ while the $\vec{\beta} \times \vec{\tau}$ term drives the helical motion.

The physical interpretation of the motion of $\vec{\tau}$ in section β_2 is as follows. Recall that the length of the PMD vector is the differential-group delay. The DGD is, roughly, a measure of the number of full-wave slips between orthogonal polarizations that has occurred. For each accumulated full-wave slip along β_2 there is an increment of the DGD by the associated delay. For instance, at 1.55 μm a one-wave slip is a delay of about 5 fs. The projection of $\vec{\tau}$ onto the $\vec{\beta}_2$ axis is approximately the number of full-wave phase slips that has occurred in the section. Clearly the longer the section the longer the projected vector.

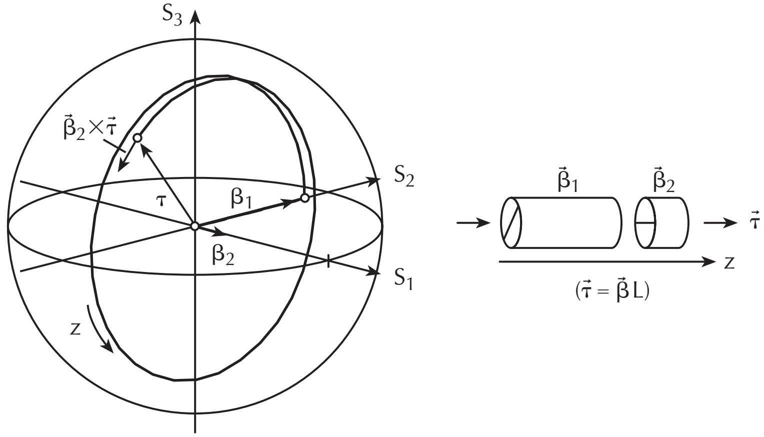

Fig. 8.23. Cumulative PMD vector $\vec{\tau}$ evolution through birefringent sections β_1 and β_2. The cumulative vector follows β_1 directly. Once the light enters β_2 the cross-product between local and cumulative vectors is non-zero and precession commences. $\vec{\tau}$ traces a helical path in the direction of β_2. The birefringence $\beta_{2\omega}$ is the driving term of the equation while $\beta_2 \times \vec{\tau}$ forces the cumulative vector in a direction perpendicular to β_2.

Since $\vec{\tau}$ is a Stokes vector as well, it must track the principal state of polarization as the light propagates. As the phase slips through one full wave the principal state (as well as the polarization state) precesses about $\vec{\beta}_2$. A full revolution is made for length Δz such that

$$\beta_2 \Delta z = 2\pi$$

If the helical motion had zero pitch then the motion of $\vec{\tau}$ about β_2 would be a pure precession. But given that the length of the PMD vector tracks the number of orthogonal-wave slips the helix pitch is greater than zero.

Figure 8.23 shows less than two full revolutions of $\vec{\tau}$ about $\vec{\beta}_2$. This corresponds to less than two full-wave phase slips, or less than 10 fs. A birefringent crystal or fiber segment that has a substantial DGD, say 1 ps, has about 200 revolutions of $\vec{\tau}$ in Stokes space at 1.55 μm. A 100 ps delay has accordingly some 20,000 revolutions. This is the origin of a order-of-magnitude distinction that is common when dealing with PMD. The DGD would have to be written to an accuracy better than one part in 20,000 in order to capture the fraction of a phase slip a long birefringent concatenation imparts. Yet only two or three significant figures are generally reported. Three significant figures leaves unresolved hundreds of revolutions for a 100 ps delay. It also leaves unresolved the fraction of a phase slip that occurs for even a 1 ps delay. But the fraction of a phase slip is essential information to properly track the PMD vector evolution. The criticality of the fractional phase slip is illustrated by the following example.

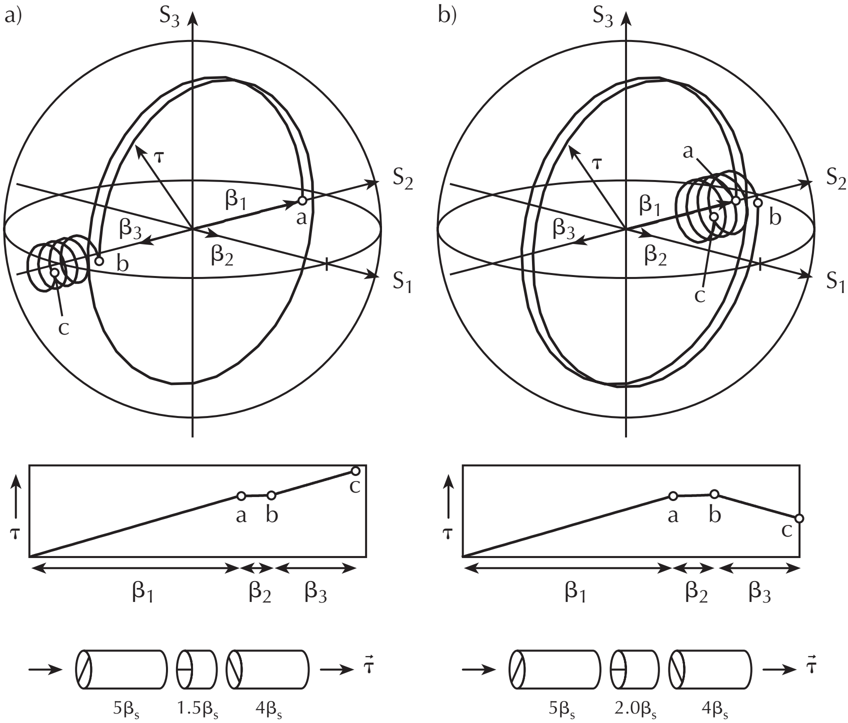

Fig. 8.24. The importance of residual birefringent phase on the evolution of the cumulative PMD vector. The three-segment cascades are identical but for the second segment: in (a) the segment imparts 1.5 revolutions while in (b) the segment imparts 2.0 revolutions. a) Trajectory of three-segment cascade where center segment has 1.5 birefringent phase slips. The cumulative DGD increases monotonically. a) Trajectory of three-segment cascade where center segment has 2.0 birefringent phase slips. The cumulative DGD decreases when it enters the third segment. The output PSP's of the two cascades are nearly in opposite directions.

Figure 8.24 shows the importance of the fractional phase slip, also known as the residual birefringent phase, in the evolution of the PMD vector. The figure illustrates the evolution through two concatenations of three segments each. The concatenations are almost the same; the only difference is that the second segment in Fig. 8.24(a) has 1.5 revolutions while the one in Fig. 8.24(b) has 2.0 revolutions. In the first sequence the PMD vector output from $\vec{\beta}_2$ points in the direction of $\vec{\beta}_3$; the cumulative PMD vector continues to increase in length through the third segment. In the second sequence, the extra half-wave rotation of $\vec{\tau}$ through the second segment orients the resultant PMD vector in the opposite direction from $\vec{\beta}_3$. Propagation through the third segment still pulls $\vec{\tau}$ toward $\vec{\beta}_3$ but does so first by decreasing the PMD vector length. The resultant cumulative vector length and pointing direction are very different for the two cases; yet the only underlying difference is a half-wave shift along the second birefringent segment. This discussion highlights the importance of the residual birefringent phase; this phase will present itself time and again in the context of the Fourier spectrum of the DGD and programmable PMD generation.

As a point of comparison to PDL evolution, the cumulative PMD vector can point in any direction in Stokes space even if the underlying birefringent vectors of the segments all lie in the same plane. For PDL, however, if the underlying PDL vectors all lie in the same plane, the cumulative PDL vector cannot leave that plane. This is illustrated in Fig. 8.6(c) on page 311. The difference stems from the $-(\vec{\alpha}\cdot\vec{\Gamma})\vec{\Gamma}$ term in the PDL evolution equation which pulls the cumulative PDL vector toward the local PDL vector, while the $\vec{\beta}\times\vec{\tau}$ term in the PMD evolution equation drives the cumulative PMD vector in a helical motion about the local birefringence. The helical motion is in a plane nearly orthogonal to the local birefringent vector (it has a small longitudinal component along the local vector). Hence the cumulative PMD vector will likely lie off of a plane of the local birefringent vectors.

8.2.6 Time-Domain Representation

The predominant representation of PMD thus far has been in the frequency domain. This is a natural consequence of the frequency-centric definition of the PMD vector. However, the parallel representation is the PMD impulse response in the time domain. The impulse response is generally a more difficult characteristic with which to make computations, but a richness of intuition is gained by understanding the parallels between the two domains.

Between the extremes of sine-wave response and impulse response lies the signal response of a communications channel, particularly the distortion imparted on a signal due to PMD. The signal response is fundamentally the convolution of the input waveform with the impulse response. What makes the calculation tricky is that co-polarized signal-image components that result from the convolution interfere coherently; the temporal location of the impulses matter to within a fraction of a wave. When in one case two co-polarized signal images are in phase and add constructively, a dephasing by π leads to destructive interference between the signal images. While the impulse weights change with changes in mode mixing, the temporal locations of the impulse response change only when the composition of the PMD concatenation changes. This implies that for a specific concatenation, the impulse response may extend well into the duration of a signal pulse; how the signal is distorted depends on which co-polarized signal images make constructive interference and which make destructive interference. In some cases the signal will look undistorted while in others it will look quite distorted. How the PMD impacts the signal depends on the expression of this coherent interference.

The following three subsections present three views of the time-domain response of a signal to PMD: simple distortions, signal-moments analysis, and impulse-moments analysis.

Signal Distortion

The impact of PMD on a signal is calculated by multiplying the waveform spectrum with the PMD spectrum or convolving the waveform with the PMD impulse response. Either way a reconciliation must be made as its natural to describe the signal waveform in the time domain and PMD in the frequency domain.

In general, the time- and frequency-domain representations of the input signal are

$$e_s(t) = f(t)\,|s\rangle\,, \quad \text{and} \quad E_s(\omega) = f(\omega)\,|s\rangle \tag{8.2.40}$$

where $f(t)$ is the waveform envelope of the electric field $e_s(t)$ and $|s\rangle$ is its input polarization state; and where $E_s(\omega)$ and $f(\omega)$ are their Fourier transform equivalents[3].

The output electric field is related to the input via the PMD transformation matrix $T(\omega)$:

$$\begin{aligned} E_t(\omega) &= T(\omega)E_s(\omega) \\ &= f(\omega)T(\omega)\,|s\rangle \end{aligned} \tag{8.2.41}$$

where $f(\omega)$ is a complex-valued function of frequency and $T(\omega)$ is an operator. In the time domain, the output waveform envelope is governed by the polarization transfer function

$$e_t(t) = \begin{pmatrix} f(t) * t_{11}(t) & f(t) * t_{12}(t) \\ f(t) * t_{21}(t) & f(t) * t_{22}(t) \end{pmatrix} |s\rangle \tag{8.2.42}$$

where $t_{ij}(t)$ is the inverse Fourier transform of the respective matrix element in $T(\omega)$. It is important to recognize that the waveform envelope and

[3] The Fourier transform pair used herein follows Haus [22]:

$$e(t) = \int_{\mathbb{R}} E(\omega)\mathrm{e}^{j\omega t}d\omega, \qquad E(\omega) = \frac{1}{2\pi}\int_{\mathbb{R}} e(t)\mathrm{e}^{-j\omega t}dt$$

$$W = \int_{\mathbb{R}} e^{\dagger}(t)e(t)dt = 2\pi\int_{\mathbb{R}} E^{\dagger}(\omega)E(\omega)d\omega$$

where the last equation is Parseval's theorem. Dirac delta functions are defined by the following integrals:

$$2\pi\,\delta(t-t') = \int_{\mathbb{R}} \mathrm{e}^{j\omega(t-t')}d\omega, \quad \text{and} \quad 2\pi\,\delta(\omega-\omega') = \int_{\mathbb{R}} \mathrm{e}^{j(\omega-\omega')t}dt$$

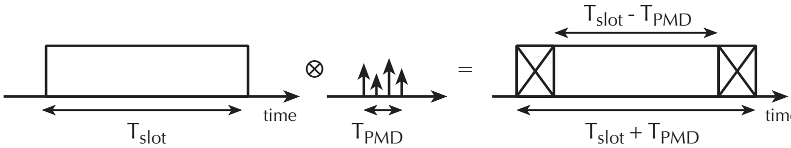

Fig. 8.25. Polarization-independent convolution of waveform envelope with PMD impulse response. The convolution broadens the output signal by T_{PMD} and reduces the duration of the undistorted portion of the signal to $T_{\mathrm{slot}} - T_{\mathrm{PMD}}$. The details of the distortion in the transition regions depend on the interference of co-polarized signal images. These details can be calculated only when the precise nature of the PMD and signal characteristic are known.

polarization state of the output field are not necessarily separable; that is, $e_t(t) = f_t(t)\,|t\rangle$ is not necessarily true. Rather, the waveform and polarization state are entangled. Such is the effect of depolarization: while each spectral component has a distinct polarization state, the inverse Fourier transform into the time domain folds the polarization states together so that on every time interval multiple polarization states exist; the degree of polarization is accordingly reduced (cf. §1.5.3).

The elements that affect the output waveform are present in (8.2.42): the input state $|s\rangle$, the impulse response of the PMD $t_{ij}(t)$, and its convolution with the waveform. It is hard to make generalizations about a system than can be arbitrarily complex. Instead, the following presents a few case studies to show different aspects of general PMD, first-order PMD, second-order PMD, and higher-order PMD.

- For an arbitrary birefringent concatenation with low overall PMD, signal distortion starts at the edges. Figure 8.25 illustrates schematically the effect of a signal "one" convolved with a simple PMD impulse response; the illustration is polarization independent but two of the four impulses are orthogonal to the others. Pulse images in the same polarization state coherently combine either constructively or destructively depending on the relative phase of the corresponding impulses. The regions of these combinations are at either transition edge of the signal. The temporal extent of the transition regions equals the width of the PMD impulse response T_{PMD}. Due to the convolution, the overall signal duration is increased by T_{PMD} and the duration of the center part of the signal is $T_{\mathrm{slot}} - T_{\mathrm{PMD}}$.

When there is a large number of impulses, and there are 2^N impulses for N birefringent segments, the gaussian distribution of impulses broadens the transition edges by an amount related to the standard deviation of the distribution. The calculation by Gisin and Pellaux [19], detailed in the third subsection, shows that the standard deviation of the impulse response equals the mean DGD.

- First-order PMD is distinct from all other forms because the output waveform is the sum of two identically shaped, orthogonally polarized, time-shifted

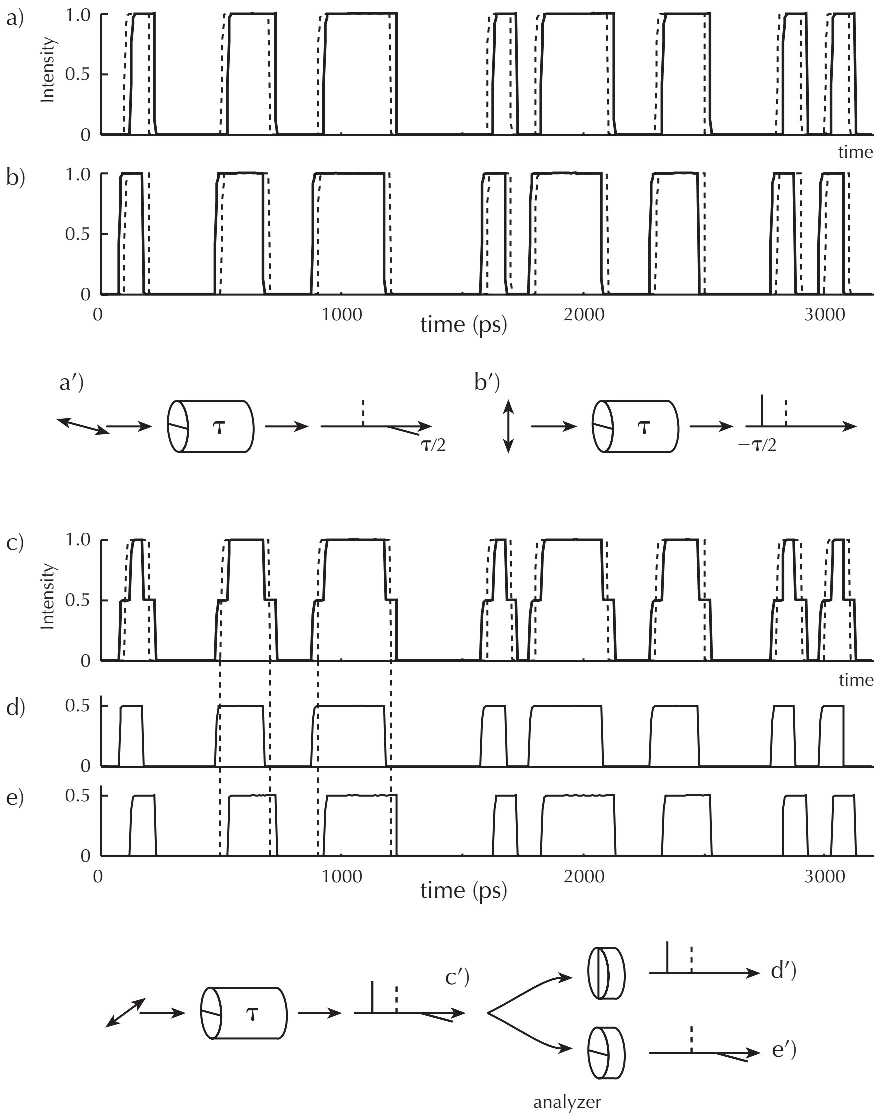

Fig. 8.26. Time response to first-order PMD. a–b) Output signal waveforms delayed and advanced by $\pm\tau/2$; the corresponding launch conditions are (a') and (b'). The output signals are undistorted. c) Distorted output waveform due to launch at 45° with respect to either birefringent axis. When analyzed by output polarizer aligned to either slow or fast axis, the original waveform is recovered, (d) and (e).

copies of the input. Pure first-order PMD comes only from a single birefringent section. Figure 8.26 illustrates the extrema conditions. When the input polarization state is aligned to either the slow or fast birefringent axis, Figs. 8.26(a') and (b') respectively, the output signal is either delayed or advanced with respect to the average delay, (a) and (b). All the light remains in a single polarization state.

Alternatively, when the input polarization state is equally divided between slow and fast axes (Fig. 8.26(c')), the signal is equally divided between the two axes and time-shifted relative to one another. The square-law detected output is distorted as shown in Fig. 8.26(c). Now, when an analyzer is placed at the output and aligned to the fast axis (d'), all the light from the slow axis is clipped; only the signal along the fast axis emerges. The shape of the analyzed output waveform is identical to the input waveform but with 3 dB less intensity (d). Similarly, when an analyzer is aligned to the slow axis (e'), the analyzed output waveform is also identical to the input but time-delayed by τ (e).

Finally, the polarization dependence of the distortion is evident from (a–c): launch along either birefringent axis leaves the signal undistorted, while the equally-mixed state launch maximally distorts the signal. The distortion is first-order launch-state dependent.

With this one example complete, the reader is warned that DGD is not PMD; DGD is one aspect of PMD. It is all too often forgotten or ignored that second-order and higher-order PMD have significant and characteristically different effects on a signal. Examples of significant activities that have been conducted under the "PMD is DGD" misconception are optical and electronic PMD compensator development; PMD emulator construction; PMD measurement; and PMD outage probability calculations. At least second-order PMD must be considered, and in fact the mean DGD of a link must also be folded into the analysis. Anything short of this richer set of considerations will likely render the calculation or product ineffective for industry applications.

- The first venture into second-order PMD (SOPMD) comes from two birefringent sections alone. These two birefringent sections impart depolarization as well as DGD onto the signal; the second component of SOPMD, the polarization-dependent chromatic dispersion (PDCD), is identically zero for two sections. The distortion of a signal due to second-order PMD is typified by overshoot and false floors. Moreover, the output waveform can *never* be resolved into two identical, time-shifted copies of the input, as was the case for first-order PMD alone.

Figure 8.26 showed that launch of a signal along a birefringent axis (or equivalently, an input PSP) into a one-stage system left the output signal undistorted. This is not true when SOPMD is present. Figure 8.27(a) shows the output distortion when the signal is launched along an input PSP. The output exhibits overshoot and false floors. When an analyzer aligned to either output PSP is placed after the two birefringent sections one observes light along both polarizations. Fig. 8.27(b) shows the analyzed components of the output signal, and (c) shows an excerpt. The evident effect is that the light

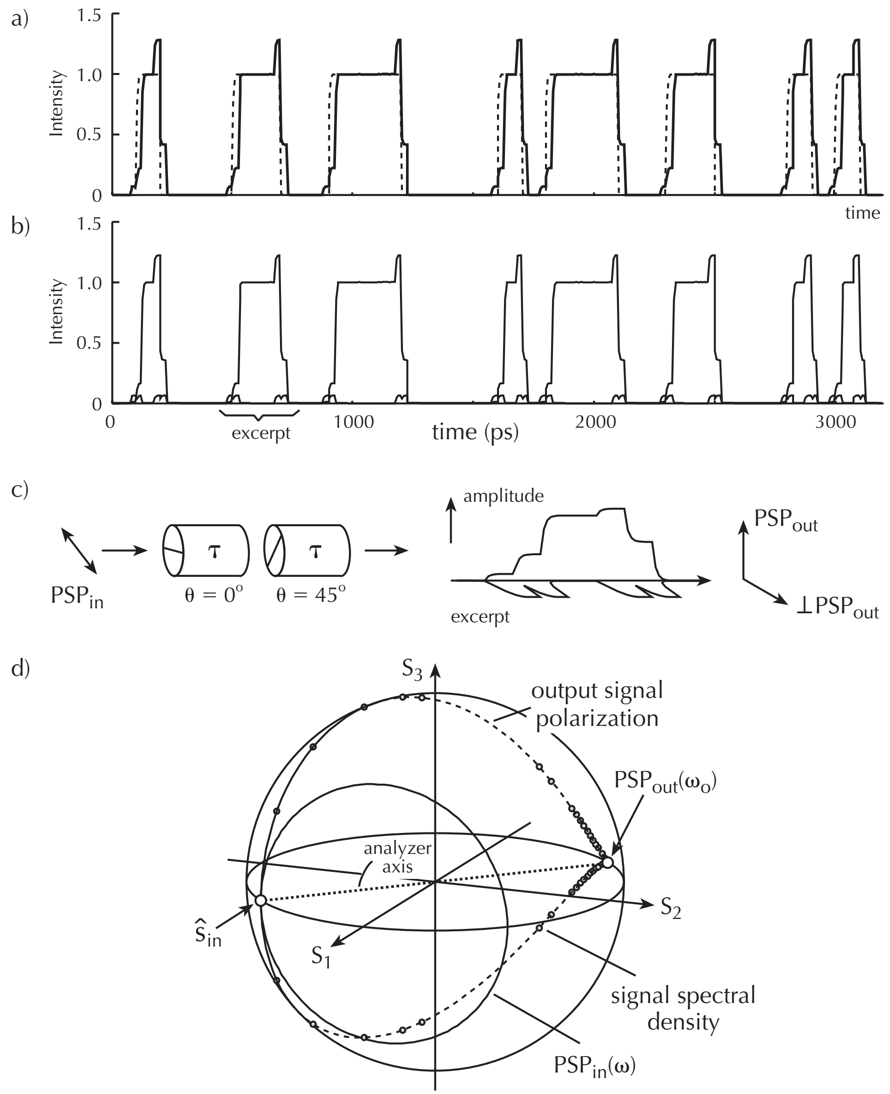

Fig. 8.27. Time-response to second-order PMD with only depolarization. a) Distorted output signal when launched state is aligned to the input PSP at a center frequency. b) Output signal analyzed by a polarizer aligned along the center-frequency output PSP and its orthogonal state. Light comes through the polarizer in both states; the output cannot be constructed from two time-shifted identical copies of the input. c) Blowup of excerpted signal, plotted in amplitude rather than intensity to highlight the relation of the waveforms. The transition edges of the input signal have in part been rotated down to the orthogonal polarization axis. d) Input launch state $\hat{s}_{\text{in}}$, input PSP spectrum, and output signal polarization spectrum.

on the orthogonal PSP axis coincides with the transition edges of the input waveform.

Along the transition edges the high-frequency components of the signal spectrum have their phases aligned; one can say the high-frequency Fourier components of the signal dominate at the edges, while the low-frequency components dominate at the centers of the "ones" and "zeros." With this in mind, the polarization spectra of the input PSP and output signal polarization are shown in Fig. 8.27(d). The launch polarization state is fixed in frequency, but the input PSP is not. At a center frequency the input PSP coincides with the launch state (here, not in general), but the PSP vector traces a circle in frequency. Accordingly, frequency components of the input signal are mapped to onto a locus of polarization states at the output; this is called polarization-state dispersion. The output signal spectrum is drawn in the figure. Overlaid with the output signal spectrum are small circles that indicate the amplitude of the signal spectrum. The signal spectrum is densely packed about $\mathrm{PSP}_{\mathrm{out}}(\omega_o)$, but at higher and lower relative frequencies the output polarization makes large excursions. Thus the high frequency components of the signal are misaligned to the output analyzer (when aligned to the output PSP) and come through. Those high-frequency components appear on the transition edges of the signal.

The impulse response of two birefringent sections is comprised of two impulses aligned along one axis and two impulses aligned along an orthogonal axis. This is shown in Fig. 8.28(a) for two equal delays. The output signal is the convolution of the input signal with this impulse response. The effect of interference, resulting in coherent addition or cancellation, of co-polarized signal images can now be well illustrated. The three columns in the center of Fig. 8.28 show how co-polarized signal images add for three different residual birefringent phases ϕ in the first delay section. When $\phi = 0°$, the impulses along the u-polarization have zero phase difference. Therefore when the convolved signal images overlap the underlying carriers are in phase and add (Fig. 8.28(c–d)). The phase relation is indicated by + signs. However, along the v-polarization the impulses are out of phase by 180°; when the signal images overlap the fields subtract, indicated by the − sign, but when they do not overlap the partial signal images come through (Fig. 8.28(d)). A square-law detector adds the orthogonal components in quadrature. The waveform in Fig. 8.28(e) for $\phi = 0°$ shows the result.

In the case when $\phi = 90°$, the phase difference between both pairs of co-polarized impulses is zero. In this case the signal images along the u-axis add as do the signal images along the v-axis. The resultant waveform is shown in Fig. 8.28(e). Finally, when $\phi = 180°$ the u-axis impulses are out of phase while the pair on the v-axis are in phase. The result is the mirror image of $\phi = 180°$ about τ_o.

Interference of co-polarized signal images plays a central role in how the PMD manifests itself on an input signal. A slight phase change clearly can change the enter shape of the signal. When the width of the PMD impulse

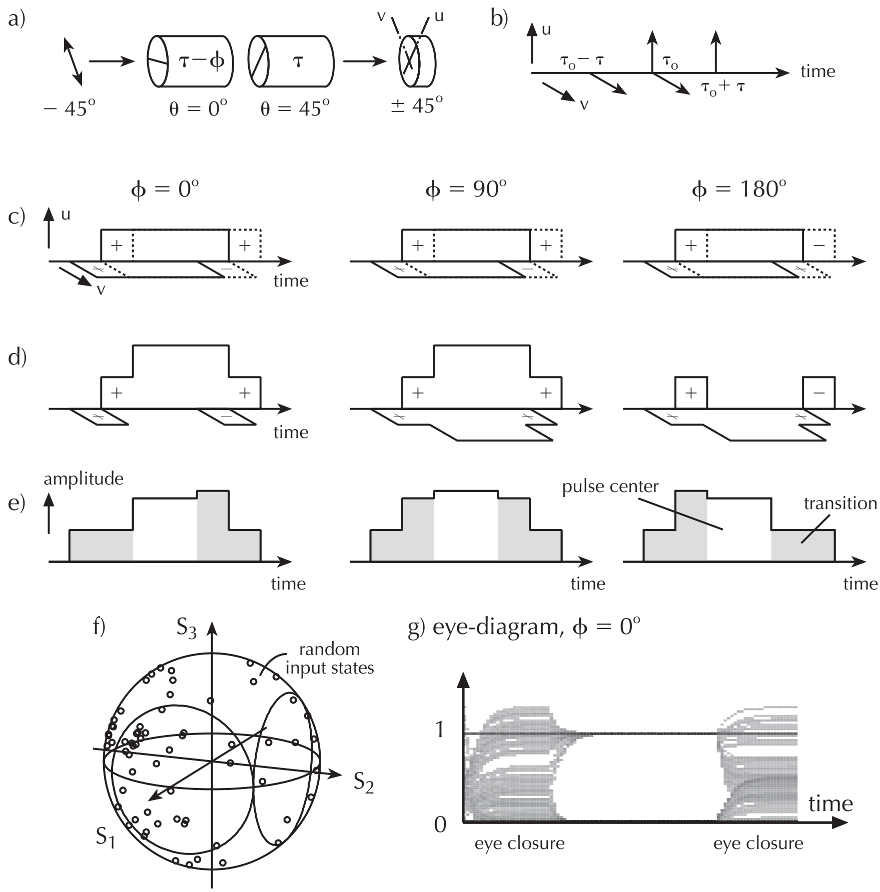

Fig. 8.28. Interference of an impulse response from two stages. a) Two stages generate an impulse with two co-polarized impulses along a u-axis and two more co-polarized impulses along a v-axis. c–e) Four signal images as convolved onto the impulse response. When two co-polarized impulses are aligned in phase the signal images add (denoted by $+$); when the same impulses are out of phase the signal images subtract (denoted by $-$). The complete output signal as measured by a square-law detector adds the coherently-added components in quadrature. f) Randomly generated launch states, and g) eye diagram of cumulative distortion.

response is on the order of the signal duration, the expression of coherent addition deep within a signal "one" or "zero" depends on this phase; when the phases cancel, the distortion will take place closer to the transitions, but when the phases add the distortion will be toward the center. In either case the temporal impulse locations do not change more than a fraction of a wave.

One aspect that is similar to first-order PMD is that signal distortion is critically dependent on launch state. Every launch state produces a different

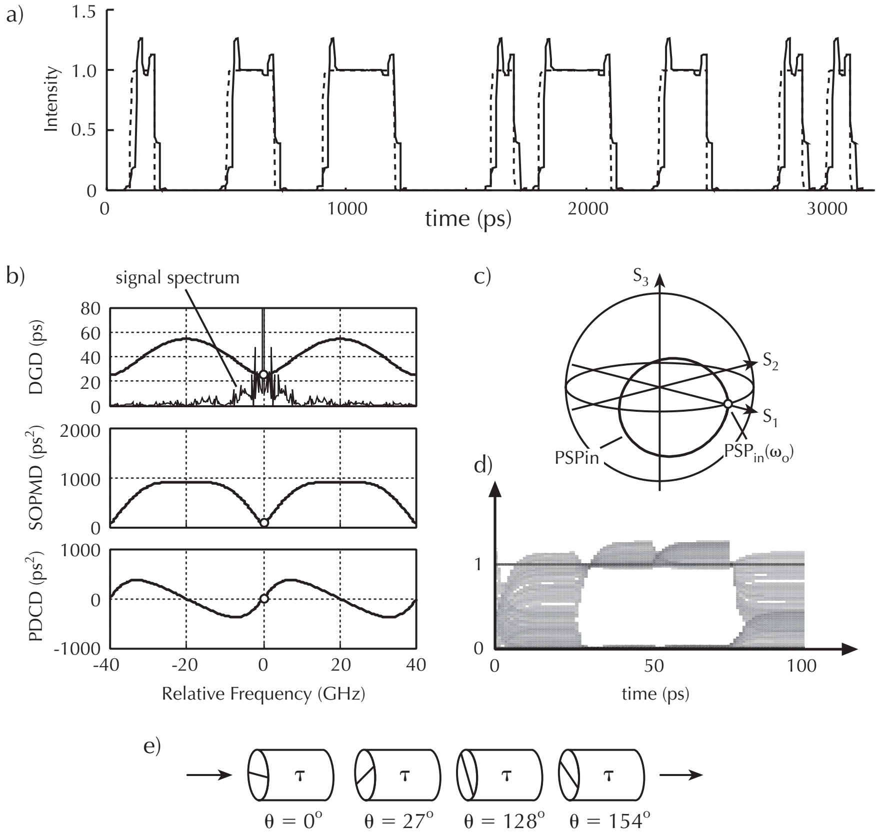

Fig. 8.29. Signal distortion for DGD with low average SOPMD. a) Comparison of distorted signal to input signal. Launch state is aligned to the input PSP at band center, indicated in (c). Four 25 ps delay sections are concatenated with mode-mixing angles shown in (e). Resultant scalar PMD spectra are shown in (b). For comparison, the signal spectrum is overlaid with the DGD spectrum. The eye diagram (d) is calculated from uniformly distributed random launch states.

distortion effect, although degeneracies may exist. Figure 8.28(g) shows the calculation of an eye diagram for 64 randomly and uniformly distributed input states. Since each delay section is 25 ps long, the outer 25 ps of the pulse are completely blurred. The overshoot is also evident. This gives a flavor why any measurement of bit-error rate or eye-margin penalty should be made while scrambling the input polarization and the measurement must last until the Poincaré sphere is reasonably covered by the input state.

• Extension beyond two birefringent stages leads toward anecdotal examples or a statistical treatment of PMD effects. A couple of important examples still remain to be shown, even though they are two of a subset in a larger class.

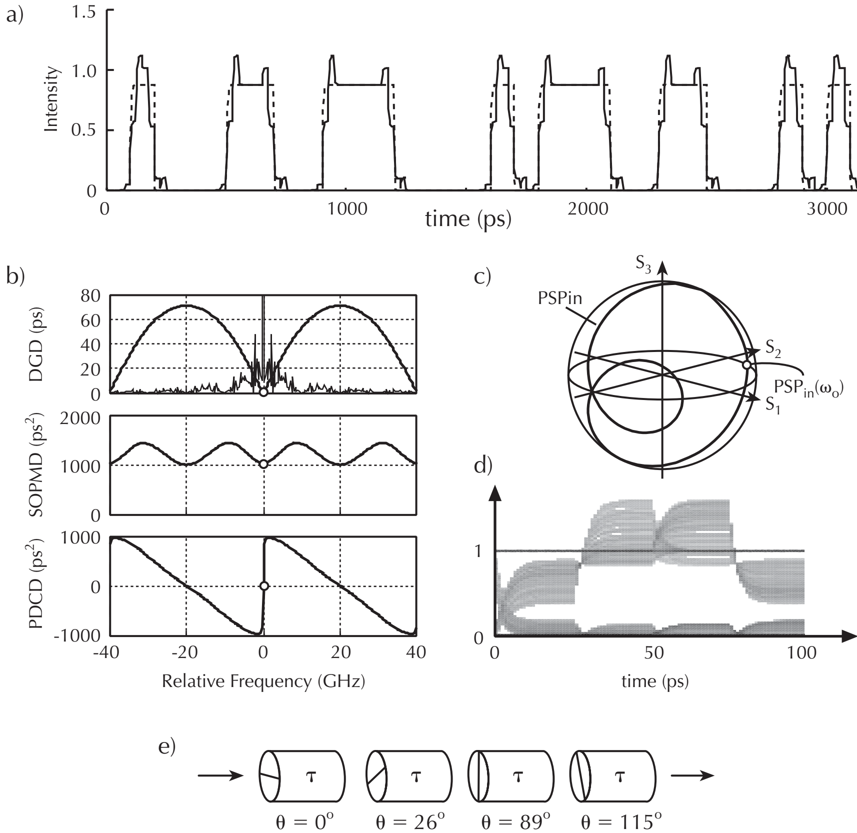

Fig. 8.30. Signal distortion for low average DGD with finite SOPMD. a) Comparison of distorted signal to input signal. Launch state is perpendicular to the input PSP at band center, the latter indicated in (c). Four 25 ps delay sections are concatenated with mode-mixing angles shown in (e). Resultant scalar PMD spectra are shown in (b). For comparison, the signal spectrum is overlaid with the DGD spectrum. The eye diagram (d) is calculated from uniformly distributed random launch states. The SOPMD creates large waveform distortions in the center region of the eye.

The examples are signal distortion for DGD with low average SOPMD, and for low average DGD with finite SOPMD. Figures 8.29 and 8.30 illustrate the two cases. Both calculations show that even when one PMD component is diminished with respect to the other, complicated distortions still occur.

The first example is DGD with low average SOPMD: (Fig. 8.29. The four birefringent sections, each 25 ps in this example, and their relative alignment is shown in (e). This example is special because the SOPMD vanishes at band center (b). The SOPMD can vanish when the output PSP pirouettes about a stationary point as the frequency changes. The PSP vector eventually stops its pirouette and continues on a course along the Poincaré sphere; the input PSP spectrum is shown in (c). The launch state chosen for this example is the input PSP at the pirouette position. The scalar spectra of the PMD condition is shown in (b)[4]; the signal spectrum is overlayed with the DGD spectrum for comparison. The output waveform is shown in (a); observe the large overshoots and false floors even though the SOPMD is significant only at higher waveform frequencies. An eye diagram for uniformly distributed random input launch states is shown in (d). Since the four section delay totals 100 ps, the width of the impulse response is 100 ps, 50 ps of which extends into the interior of the signal waveform. This is why the center region of the waveform is distorted. However, the particular location of the signal spectrum with respect to the PMD spectrum prevents the marginal impulses at either end of the PMD impulse response to have strong weight. Thus the eye center is not closed.

The second example is low average DGD with finite SOPMD: Fig. 8.30. Here, the DGD vanishes at band center, while the SOPMD remains significant. The DGD can vanish at one frequency when the component PMD vectors form a closed loop in Stokes space. Since in this case there are four component vectors, the closed loop is a square or rhombus. The depolarization must be finite in this case, or else the closed loop of PMD component vectors would not open back up (or close on itself in the first place). Nonetheless, even with the signal spectrum aligned to the vanishing point of the DGD, the signal distortion is significant. The launch state associated with the distortion in (a) is perpendicular to the input PSP at band center. The input PSP spectrum is shown in (c). Of particular interest is the eye diagram generated for a large number of randomly selected launch states. Even when the average DGD is low, significant distortion appears all across the pulse. This is caused by the marginal impulses in the PMD impulse response having significant weight. The eye center is still not closed, though, because the average DGD is low.

Signal Moments and Distortion

One can make some general statements about the effects of PMD by looking at the moments of the signal waveform in the time domain.

Karlsson calculates the first and second moments of a signal waveform when effected by PMD [38]. Gordon offers details of Karlsson's presentation [20]. The main results are that: 1) the differential-group delay is the maximum possible delay of a narrowband signal launched into the fast and slow input PSP's, all other launch conditions yield a first-moment less than the DGD; 2) there is a minimum, non-zero pulse broadening in the presence of second-order PMD principally due to the rotation of the PSP's in frequency.

The first- and second-moments of the waveform are calculated by

---

[4] Appendix C shows how to efficiently calculate the vector and scalar spectra for an arbitrary birefringent concatenation.

$$\langle t \rangle = \frac{1}{W} \int_{\mathbb{R}} t e^{\dagger}(t) e(t) dt = \frac{2j\pi}{W} \int_{\mathbb{R}} E^{\dagger}(\omega) E_{\omega}(\omega) d\omega \tag{8.2.43a}$$

$$\langle t^2 \rangle = \frac{1}{W} \int_{\mathbb{R}} t^2 e^{\dagger}(t) e(t) dt = \frac{2\pi}{W} \int_{\mathbb{R}} {E_{\omega}}^{\dagger}(\omega) E_{\omega}(\omega) d\omega \tag{8.2.43b}$$

The first moment at the input is simply

$$\begin{aligned} \langle t_s \rangle &= \frac{2j\pi}{W} \int_{\mathbb{R}} {E_s}^{\dagger} E_{s\omega} \, d\omega \\ &= \frac{2j\pi}{W} \int_{\mathbb{R}} f_{\omega}^*(\omega) f(\omega) \, d\omega \end{aligned}$$

If the waveform spectrum is real and symmetric then $\langle t_s \rangle = 0$. Since the output-field spectrum is $E_t = \exp(-j\phi_o) U E_s$, where ϕ_o is the common phase through the system and U is the unitary operator for the cascade, the integrand to $\langle t_t \rangle$ at the output is

$${E_t}^{\dagger} E_{t\omega} = -j {E_s}^{\dagger} \left(\tau_o + \tfrac{1}{2} (\vec{\tau}_s \cdot \vec{\sigma}) \right) E_s + {E_s}^{\dagger} E_{s\omega}$$

where $\vec{\tau}_s$ is the input principal state (recall $2jU_{\omega}U^{\dagger} = \vec{\tau}_s \cdot \vec{\sigma}$) and τ_o is the common delay through the system. Note that τ_o and $\vec{\tau}$ are both functions of frequency. The first moment of the output waveform is therefore

$$\begin{aligned} \langle t_t \rangle &= \langle t_s \rangle + \frac{2\pi}{W} \int_{\mathbb{R}} {E_s}^{\dagger} \left(\tau_o + \tfrac{1}{2} (\vec{\tau}_s \cdot \vec{\sigma}) \right) E_s \, d\omega \\ &= \langle t_s \rangle + \frac{2\pi}{W} \int_{\mathbb{R}} |f(\omega)|^2 \left(\tau_o + \tfrac{1}{2} (\vec{\tau}_s \cdot \hat{s}) \right) d\omega \end{aligned} \tag{8.2.44}$$

The mean signal delay τ_g through the medium is simply the difference of first moments:

$$\tau_g = \langle t_t \rangle - \langle t_s \rangle \tag{8.2.45}$$

This can be expressed in spectrally averaged form as

$$\tau_g = \left\langle \tau_o(\omega) \, |f(\omega)|^2 \right\rangle_{\omega} + \frac{1}{2} \left\langle \hat{s}(\omega) \cdot \vec{\tau}_s(\omega) \, |f(\omega)|^2 \right\rangle_{\omega} \tag{8.2.46}$$

where $\hat{s}(\omega)$ allows for the possibility of frequency dependence of the input state. Physically, the mean signal delay is the normalized spectral average of the common delay weighted by the waveform envelope, plus the spectral average of the input launch polarization as projected onto the input PSP, again weighted by the waveform envelope.

An important observation is that the phase of the waveform envelope $f(t)$ is eliminated in (8.2.46); initial chirp or chromatic dispersion of the pulse does not affect its average position at the output.

Consider two extreme cases for (8.2.46): monochromatic input, and any input to a single birefringence segment. For monochromatic input, the waveform input is a sine wave, so in frequency $f(\omega) = \delta(\omega - \omega_o)$. The mean signal delay for any concatenation is

$$\tau_g = \tau_o + \tfrac{1}{2}\,\tau\,(\hat{s}\cdot\hat{r}_s) \tag{8.2.47}$$

where each term is evaluated at ω_o. This expression is the main result which connects the mean signal delay to the spectral description of PMD. The mean delay at ω_o is $\tau_g = \tau_o \pm \tau/2$ when the launch state is parallel or perpendicular to the input PSP. Any intermediate launch condition produces a first-moment that is between these extrema. Accordingly, one can say that the DGD is the maximum delay at a particular frequency between the fast and slow axes of a cascade. This interpretation was used in the time-frequency correspondence figure (Fig. 8.18 on page 326).

When there is only one homogeneous birefringent element then τ_o and $\vec{\tau}_s$ are stationary with frequency; the mean signal delay has the same form as (8.2.47). This is an important connection to the impulse response of a cascade which is considered in the next section.

The pulse spreading between the output and input can be measured by the second moments of the signal. The pulse spread $\Delta\tau$ is defined as

$$\Delta\tau = \sqrt{(\langle t_t^2\rangle - \langle t_s^2\rangle) - (\langle t_t\rangle - \langle t_s\rangle)^2} \tag{8.2.48}$$

The second moment of the input waveform is

$$\langle t_s^2\rangle = \frac{2\pi}{W}\int_{\mathbb{R}} E_{s\omega}{}^{\dagger}E_{s\omega}d\omega = \frac{2\pi}{W}\int_{\mathbb{R}} |f_\omega(\omega)|^2\,d\omega \tag{8.2.49}$$

At the output, $E_{t\omega} = T_\omega E_s + TE_{s\omega}$. The difference between output and input second moments is

$$E_{t\omega}{}^{\dagger}E_{t\omega} - E_{s\omega}{}^{\dagger}E_{s\omega} = E_s{}^{\dagger}T_\omega{}^{\dagger}T_\omega E_s + E_s{}^{\dagger}T_\omega{}^{\dagger}TE_{s\omega} + E_{s\omega}{}^{\dagger}T^{\dagger}T_\omega E_s$$

The frequency derivative of the transformation matrix is

$$T_\omega = -j\mathrm{e}^{-j\phi_o}U\left(\tau_o + \tfrac{1}{2}\,\vec{\tau}_s\cdot\vec{\sigma}\right)$$

where the substitution $jU^{\dagger}U_\omega = \frac{1}{2}\,\vec{\tau}_s\cdot\vec{\sigma}$ was made. Since $jU^{\dagger}U_\omega$ is Hermitian, the matrix product is $T_\omega{}^{\dagger}T_\omega = \left(\tau_o + \frac{1}{2}\,\vec{\tau}_s\cdot\vec{\sigma}\right)^2$, or

$$T_\omega{}^{\dagger}T_\omega = \tau_o^2 + \tfrac{1}{4}\tau^2 + \tau_o\,(\vec{\tau}_s\cdot\vec{\sigma})$$

where the identity $(\vec{\tau}_s\cdot\vec{\sigma})\,(\vec{\tau}_s\cdot\vec{\sigma}) = \tau^2$ was used. Therefore

$$E_s{}^{\dagger}T_\omega{}^{\dagger}T_\omega E_s = |f|^2\left(\tau_o^2 + \tfrac{1}{4}\tau^2 + \tau_o\,(\vec{\tau}_s\cdot\hat{s})\right)$$

For the remaining two terms, one has $T_\omega{}^{\dagger}T = j\left(\tau_o + \frac{1}{2}\,\vec{\tau}_s\cdot\vec{\sigma}\right)$. The sum is

$$E_s{}^{\dagger}T_\omega{}^{\dagger}TE_{s\omega} + E_{s\omega}{}^{\dagger}T^{\dagger}T_\omega E_s = j\,(f^* f_\omega - f_\omega^* f)\left(\tau_o + \tfrac{1}{2}\,\vec{\tau}_s\cdot\hat{s}\right)$$

Expansion of the complex waveform envelope into amplitude and phase, $f(\omega) = a(\omega)\exp(-j\phi(\omega))$, one finds that

$$j\left(f^* f_\omega - f_\omega^* f\right) = 2\left|f\right|^2 \phi_\omega$$

where ϕ_ω is the frequency derivative of the waveform phase. Now that each integrand has been reduced, the complete second moment of the pulse spread is

$$\begin{aligned}\left\langle t_t^2\right\rangle - \left\langle t_s^2\right\rangle = \frac{2\pi}{W}\int_{\mathbb{R}} \left|f(\omega)\right|^2 \left(\tau_o^2 + \tfrac{1}{4}\tau^2 + \tau_o\left(\vec{\tau}_s\cdot\hat{s}\right)\right) d\omega \\ + \frac{4\pi}{W}\int_{\mathbb{R}} \left|f(\omega)\right|^2 \phi_\omega(\omega)\left(\tau_o + \tfrac{1}{2}\,\vec{\tau}_s\cdot\hat{s}\right) d\omega \quad (8.2.50)\end{aligned}$$

The first term on the right-hand side is similar in form to (8.2.46), where the waveform envelope intensity is the weighting factor to the spectral average of the delay components. Like the first moment, this term of the second moment does not depend on the phase of the waveform. The second term, however, includes the derivative of the waveform phase. Therefore, the pulse spread is related not only to the PMD but to the phase across the waveform as well. For example, for a linear delay across the waveform, $\phi_w(\omega) = \gamma\omega$, and the resultant ω multiplier in the integrand of the second term is equivalent to a time-derivative of the waveform intensity. The pulse spread then depends on the temporal details of the signal shape as well as the intensity spectrum.

Consider an example of one section of birefringence, the PMD τ is only first order, and the signal is real ($\phi_w(\omega) = 0$). The difference between output and input first and second moments is then

$$\begin{aligned}\left\langle t_t^2\right\rangle - \left\langle t_s^2\right\rangle &= \tau_o^2 + \tfrac{1}{4}\tau^2 + \tau_o\left(\vec{\tau}_s\cdot\hat{s}\right) \\ \left\langle t_t\right\rangle - \left\langle t_s\right\rangle &= \tau_o + \tfrac{1}{2}\left(\vec{\tau}_s\cdot\hat{s}\right)\end{aligned}$$

where $\hat{s}$ is the input launch state. Expansion of the frequency-independent dot product between launch state and input PMD vector gives $(\vec{\tau}_s\cdot\hat{s}) = \tau\cos\theta$, where θ is the Stokes angles between vectors. The pulse spread is therefore

$$\Delta\tau = \tfrac{1}{2}\,\tau\sin\theta \quad (8.2.51)$$

For comparison, the signal delay is

$$\tau_g = \tau_o + \tfrac{1}{2}\,\tau\cos\theta$$

Figure 8.31 illustrates the example. The maximum arrival-time deviation from τ_o is when the signal is launched along the fast or slow axis of the birefringent element (Fig. 8.31(a)). In this case $\cos\theta = \pm 1$. There is no pulse spreading with this condition. This is reasonable since all the energy is inserted into one eigenstate; no differential delay is experienced. However, if the launch is at 45° with respect to the input, then $\tau_g = \tau_o$ and the pulse spreading achieves its maximum at $\Delta\tau = \tau/2$ (Fig. 8.31(b)). Clearly when the DGD of the system approaches the time-slot duration of the signal then all communication can be lost.

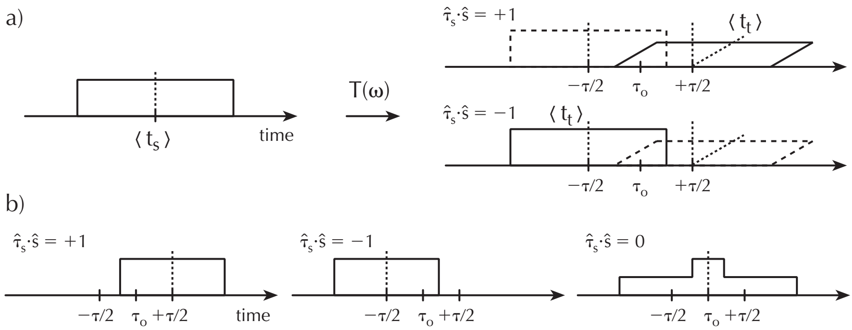

Fig. 8.31. Illustration of signal envelope through a very simple high-birefringence (first-order PMD only) system. a) When the signal is launched along a birefringent axis it is either advanced or delayed with respect to the isotropic travel time τ_o. b) Three launch conditions: aligned to slow axis, aligned to fast axis, and equally mixed between slow and fast. The waveform shapes represent amplitude and not polarization. The mixed launch imparts pulse spreading.

Consider an example of two sections of birefringence (cf. Fig. 8.21). This is the simplest form of second-order PMD: only depolarization exists. The output PMD vector is calculated from the concatenation rule: $\vec{\tau}_t = \vec{\tau}_2 + R_2\vec{\tau}_1$. The moment calculations use the input PSP so that the launch state $\hat{s}$ can be projected. The corresponding input PMD vector is $\vec{\tau}_s = \vec{\tau}_1 + {R_1}^\dagger\vec{\tau}_2$. The motion of the input PSP is illustrated in Fig. 8.21, but with the terms $\vec{\tau}_1$ and $\vec{\tau}_2$ interchanged. The input PSP traces a circle in Stokes space over frequency while the cumulative DGD remains fixed; the period of the circle is $\Delta\omega = 2\pi/\tau_1$. The mean delay of the output waveform is

$$\tau_g = \tau_o + \frac{2\pi}{W}\int_{\mathbb{R}} |f|^2 \, \tfrac{1}{2} \left(\vec{\tau}_s(\omega)\cdot\hat{s}\right) d\omega$$

Defining the integral

$$I_s = \frac{2\pi}{W}\int_{\mathbb{R}} |f|^2 \left(\hat{r}_s(\omega)\cdot\hat{s}\right) d\omega$$

where $\hat{r}_s$ is the pointing direction of the input PSP, itself a function of frequency, the mean delay is then recast as

$$\tau_g = \tau_o + \tfrac{1}{2}\,\tau I_s$$

The cumulative PMD τ was extracted from the integral I_s because its magnitude is fixed in frequency, as is the case for two birefringent stages. In a similar manner, the square of the pulse spreading at the output is

$$\begin{aligned}\Delta\tau^2 &= \tau_o^2 + \tfrac{1}{4}\,\tau^2 + \tau_o\tau I_s - \tau_g^2 \\ &= \tfrac{1}{4}\,\tau^2\left(1 - I_s^2\right)\end{aligned}$$

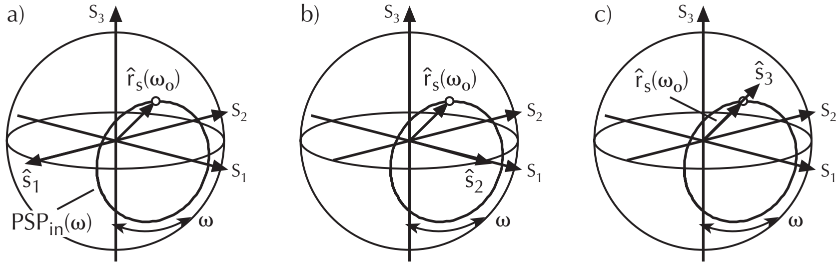

Fig. 8.32. Three launch conditions into a depolarizing system. Two birefringent sections generate precession of the input PMD vector about $\vec{\tau}_1$. a) Launch state for maximum pulse spread, $\hat{s}_1 = \pm\, \vec{\tau}_1 \times \vec{\tau}_2$. b) Intermediate launch state where $\hat{s}_2 \cdot \hat{r}_s$ is constant in frequency. c) Launch state for largest minimum pulse spread, $\hat{s}_3 = \hat{r}_s(\omega_o)$.

Since I_s enters the equation as a squared quantity, the pulse spread can only decrease with I_s. The value of I_s depends on the launch state at the input and the degree of rotation of $\hat{r}_s$. Figure 8.32 illustrates three launch states, the first and last states being the extrema. The maximum pulse spreading occurs when $I_s = 0$. State $\hat{s}$ can always be selected to drive I_s to zero. For a symmetric spectrum centered at ω_o and launch state $\hat{s}_1 = \pm\, \vec{\tau}_1 \times \vec{\tau}_2$ (Fig. 8.32(a)), where $\vec{\tau}_1$ and $\vec{\tau}_2$ are evaluated at ω_o, the product $\hat{r}_s \cdot \hat{s}_1$ is antisymmetric. Integral I_s therefore vanishes and the pulse spread is $\Delta\tau^2 = \tau^2/4$.

An interesting albeit non-extrema case is where the inner product is fixed in frequency. This occurs when $\hat{s}$ is aligned with $\vec{\tau}_1$ (Fig. 8.32(b)). In this case $I_s = \cos\theta$. Only when the two birefringent axes are aligned does the pulse spread reach zero. But this is simply the case of a single birefringent segment made up of two parts. The general case is when there is mode mixing between sections. Thus in the general $\hat{s}_2$ is not a state that produces minimum pulse spreading.

The launch state for minimum pulse spreading is illustrated in Fig. 8.32(c). In general, I_s will be less than unity, so the pulse always experiences non-zero minimum spread. This contrasts with the first-order PMD situation where launch along a PSP ensures zero spreading. The problem here is that the PSP's move with frequency, so there is no single PSP to launch into.

One can calculate the largest minimum pulse spread. This case coincides with maximum depolarization, which is when $\tau_1 = \tau_2$ and $\hat{r}_s \cdot \hat{s}_1 = 0$. The PMD vector at the input is

$$\vec{\tau}_s/\tau_1 = \hat{r}_1 - \sin\omega\tau_1\, \hat{r}_1 \times \hat{r}_2 - \cos\omega\tau_1\, \hat{r}_1 \times \hat{r}_1 \times \hat{r}_2$$

and the launch state is $\hat{s}_3 = (\hat{r}_1 + \hat{r}_2)/\sqrt{2}$. Since the length of the PMD vector is $\tau_s = \sqrt{2}\tau_1$, the frequency-dependence of the inner product is

$$\hat{s}_3 \cdot \hat{r}_s = \tfrac{1}{2}\left(1 + \cos\omega\tau_1\right)$$

In the regime where the bandwidth of the waveform is greater than $1/\tau_1$, then the frequency average of the inner product is driven to one-half. In this case the minimum pulse spread is approximately (ignoring details of the waveform spectrum) $\Delta\tau^2 \simeq (3/16)\tau^2$. The comparison between the two extrema is

$$\begin{aligned} &\text{maximum pulse spread:} \quad \hat{s} = \pm\, \vec{\tau}_1 \times \vec{\tau}_2, \quad \Delta\tau = \tau/2 \\ &\text{minimum pulse spread:} \quad \hat{s} = \hat{r}(\omega_o), \qquad \Delta\tau \to \sqrt{3}\,\tau/4 \end{aligned} \tag{8.2.52}$$

Nearly the same pulse spreading as found for the "best" and "worst" launch conditions. Clearly depolarization can have a strong effect on the waveform.

PMD Impulse Response

Gisin and Pellaux deduced the formal connection between the PMD impulse response and the DGD spectrum of a lossless birefringent cascade [19]. Their result shows that the root-mean-square (rms) DGD is equal to the standard deviation of the impulse response width. The derivation is elegant and yet result seems to be under represented in the subsequent literature.

The derivation constructs a recurrence relation for the frequency average of the mean-squared DGD, denoted $\langle \tau^2(\omega) \rangle$, and a separate recurrence relation for the second-moment of the impulse response, denoted $\langle h^2(t) \rangle$, from the same concatenation. The recurrence relations are then shown to be equivalent.

Consider a concatenation of N homogeneous birefringent sections, each section having DGD of τ_n and birefringent-vector orientation $\hat{r}_n$. In the frequency domain, the two-block PMD concatenation rule (8.2.33) on page 335 is written to relate the last element, element N, to the preceding $N-1$ elements as

$$\vec{\tau}(N) = \vec{\tau}_N + R_N \vec{\tau}(N-1) \tag{8.2.53}$$

where $\vec{\tau}(N)$ denotes the cumulative PMD through section N and $\vec{\tau}_N$ denotes the N^{th} local birefringent element. The cumulative vector $\vec{\tau}(N)$ is clearly a function of frequency ω while the local elements, such as $\vec{\tau}_N$, are not. The dot product of $\vec{\tau}$ with itself provides the DGD squared, in general $\tau^2 = \vec{\tau} \cdot \vec{\tau}$. The DGD squared for $\vec{\tau}(N)$ from recurrence relation (8.2.53) is

$$\tau^2(N) = \tau_N^2 + \tau^2(N-1) + 2\vec{\tau}_N \cdot \vec{\tau}(N-1)$$

The DGD squared spectrum, $\tau^2(N;\omega)$, is then averaged over all frequency to find its mean-square. The average is written

$$\left\langle \tau^2(N) \right\rangle_\omega = \tau_N^2 + \left\langle \tau^2(N-1) \right\rangle_\omega + 2 \left\langle \vec{\tau}_N \cdot \vec{\tau}(N-1) \right\rangle_\omega$$

The value τ_N^2 comes through the average as its just a number. The last term on the right-hand side may be further reduced by employing (8.2.53) to the $N-1$ term:

$$\begin{aligned} \langle \vec{\tau}_N \cdot \vec{\tau}(N-1) \rangle_\omega &= \langle \vec{\tau}_N \cdot (\vec{\tau}_{N-1} + R_{N-1}\, \vec{\tau}(N-2)) \rangle_\omega \\ &= \vec{\tau}_N \cdot \vec{\tau}_{N-1} + \vec{\tau}_N \cdot \langle R_{N-1}\, \vec{\tau}(N-2) \rangle_\omega \end{aligned} \tag{8.2.54}$$

The evaluation of the last term on the right-hand side is at the heart of the derivation. Expansion of the last term to include the rotation operator R_{N-1} gives

$$\begin{aligned}\langle R_{N-1}\,\vec{\tau}(N-2)\rangle_\omega = &\ \langle \hat{r}_{N-1}\hat{r}_{N-1}\cdot\vec{\tau}(N-2)\rangle_\omega \\ &+ \langle \sin(\omega\tau_{N-1})\,\hat{r}_{N-1}\times\vec{\tau}(N-2)\rangle_\omega \\ &- \langle \cos(\omega\tau_{N-1})\,\hat{r}_{N-1}\times\hat{r}_{N-1}\times\vec{\tau}(N-2)\rangle_\omega \end{aligned} \tag{8.2.55}$$

The last two frequency-average terms include sine and cosine terms that average to zero. Even though $\vec{\tau}(N-2)$ itself generally varies with frequency, for a concatenation with enough segments the frequency average will eventually drive these terms to zero. In contrast, the disposition of the first term on the right-hand is not so clear, so it remains. Insertion of (8.2.55) into (8.2.54) and some manipulation produces

$$\langle \vec{\tau}_N\cdot\vec{\tau}(N-1)\rangle_\omega = \tau_N\,(\hat{r}_N\cdot\hat{r}_{N-1})\,(\tau_{N-1}+\langle \hat{r}_{N-1}\cdot\vec{\tau}(N-2)\rangle_\omega)$$

The complete recurrence relation in the frequency domain is therefore

$$\begin{aligned}\left\langle \tau^2(N)\right\rangle_\omega = &\ \tau_N^2 + \left\langle \tau^2(N-1)\right\rangle_\omega \\ &+ 2\,\tau_N\,(\hat{r}_N\cdot\hat{r}_{N-1})\,(\tau_{N-1}+\langle \hat{r}_{N-1}\cdot\vec{\tau}(N-2)\rangle_\omega)\end{aligned} \tag{8.2.56}$$

The form of this expression has fully identified the effect of $\vec{\tau}_N$ on $\vec{\tau}(N-1)$. It is this recurrence relation that will be compared to the impulse-response relation.

Now for the impulse response. Each impulse has a position in time and a weight. As there is no loss, the sum of all the weights remains fixed regardless of the number of sections in the concatenation. The time-position and weight of each impulse depends on the path the light takes. If there are n sections, there are 2^n paths and 2^n impulses at the output. Each one has to be enumerated to construct the impulse response. The time position of the k^{th} pulse with respect to the common delay through the concatenation is

$$t_k = \epsilon_1(k)\tau_1 + \epsilon_2(k)\tau_2 + \ldots + \epsilon_N(k)\tau_N \tag{8.2.57}$$

where $\epsilon_i = \pm 1$ depending whether the impulse travels along the slow or fast axis of the i^{th} segment[5] To enumerate each path only once the binary equivalent of the decimal index is useful. If for each index integer $k \in [0, N-1]$ the integer is converted to its binary form $k \mapsto b_0 b_1 b_2 \cdots b_N$, then the path selector ϵ is defined by $\epsilon_i = 1 - 2b_i$. The path selector is further indexed by k so that $\epsilon_i(k)$ is generated by the i^{th} binary digit of the binary representation of the index k.

[5] The factor of one-half is dropped to conform with the PMD vector definition $\vec{\tau} = \tau\hat{p}$ rather than $\vec{\tau} = \tau/2\hat{p}$.

The weight of the k^{th} impulse is determined by the projection of adjacent birefringent axes and whether the impulse is going between fast axis in one segment to the fast axis in the next, or between fast and slow, slow and fast, or slow and slow. The weight of the impulse is

$$w(k) = \frac{1}{2} \prod_{n=2}^{N} \frac{1}{2} \left(1 + \epsilon_{n-1}(k)\, \epsilon_n(k)\, \hat{r}_{n-1} \cdot \hat{r}_n \right) \tag{8.2.58}$$

where the first one-half factor comes from a 45° launch into the first element, that is $\hat{s}_{\text{in}} \cdot \hat{r}_1 = 0$. The weights are normalized such that

$$\sum_{k=1}^{2^N} w(k) = 1$$

The impulse response of the birefringent concatenation is

$$h(t, N) = \sum_{k=1}^{2^N} w(k)\, \delta(t - t_k)$$

where $\delta(t - t_o)$ is the dirac delta function that has zero value everywhere but for t_o. In the following the explicit time dependence of h will be dropped. The average position of the impulse response relative to the common delay is zero:

$$\langle h(N) \rangle_t = \int_{\mathbb{R}} t h(t, N) dt = \sum_{k=1}^{2^N} w(k)\, t_k = 0 \tag{8.2.59}$$

The pulse train is symmetric about $t = 0$. The second moment of $h(t, N)$ requires more work. The second moment is

$$\left\langle h^2(N) \right\rangle_t = \int_{\mathbb{R}} t^2 h(t, N) dt = \sum_{k=1}^{2^N} w(k)\, t_k^2 \tag{8.2.60}$$

The square of the k^{th} time position is expanded with (8.2.57) to give

$$\begin{aligned} t_k^2 &= \sum_{i=1}^{N} \sum_{j=1}^{N} \epsilon_i(k)\, \epsilon_j(k)\, \tau_i\, \tau_j \\ &= \sum_{i=1}^{N} \tau_i^2 + 2 \sum_{i=1}^{N} \sum_{j=i+1}^{N} \epsilon_i(k)\, \epsilon_j(k)\, \tau_i\, \tau_j \end{aligned}$$

where the first term on the right-hand side of the second line is the sum along the diagonal of the $N \times N$ matrix while the second term is the sum over the upper triangle of the matrix. Substitution back into (8.2.60) gives

$$\left\langle h^2(N)\right\rangle_t = \sum_{i=1}^{N} \tau_i^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} \tau_i\,\tau_j \sum_{k=1}^{2^N} \epsilon_i(k)\,\epsilon_j(k)\,w(k)$$

where the normalization of $w(k)$ was applied to the first right-hand-side term. The sum over k in the second term has an significant simplification. The binary-weight product $\epsilon_i(k)\epsilon_j(k)$ changes sign when counting over k with a frequency that depends on j. For instance with $N = 4$, $\epsilon_i(k)\epsilon_j(k)$ changes sign for every increment of k when $j = 4$, but changes sign for every two increments when $j = 3$. Concurrently, $\epsilon_{n-1}(k)\epsilon_n(k)$ changes sign with a rate when counting over k related to n. The combined result is that all terms of $\epsilon_{n-1}(k)\epsilon_n(k)$ that change sign at a counting-rate faster than $\epsilon_i(k)\epsilon_j(k)$ vanish from the sum over k while the remaining terms add to a unity coefficient. That is, all $n > j$ terms vanish.

The second-moment of $h(N)$ simplifies to

$$\left\langle h^2(N)\right\rangle_t = \sum_{i=1}^{N} \tau_i^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} \tau_i\,\tau_j\ (\hat{r}_i\cdot\hat{r}_{i+1})\cdots(\hat{r}_{j-1}\cdot\hat{r}_j)$$

To construct a recurrence relation, the sum $\left\langle h^2(N)\right\rangle_t$ must be related to the partial sum $\left\langle h^2(N-1)\right\rangle_t$. Recall that $\left\langle h^2(N)\right\rangle_t$ can be viewed as the element sum over a symmetric $N \times N$ matrix, the first right-hand-side term being the sum along the diagonal and the second term being the sum of the upper triangle. The difference between $\left\langle h^2(N)\right\rangle_t$ and $\left\langle h^2(N-1)\right\rangle_t$ is therefore the element sum of the N^{th} column. Accordingly, the recurrence relation is

$$\left\langle h^2(N)\right\rangle_t = \tau_N^2 + \left\langle h^2(N-1)\right\rangle_t + 2\sum_{i=1}^{N-1} \tau_i\tau_N\,(\hat{r}_1\cdot\hat{r}_2)\cdots(\hat{r}_{N-1}\cdot\hat{r}_N)$$

or more succinctly,

$$\left\langle h^2(N)\right\rangle_t = \tau_N^2 + \left\langle h^2(N-1)\right\rangle_t + 2\tau_N S_N \tag{8.2.61}$$

where

$$S_N = (\hat{r}_N\cdot\hat{r}_{N-1})\,(\tau_{N-1} + (\hat{r}_{N-1}\cdot\hat{r}_{N-2})\,(\tau_{N-2}+\cdots)) \tag{8.2.62}$$

Comparison of the frequency-average recurrence relation (8.2.56) and the time-average recurrence relation (8.2.61) shows that $\left\langle \tau^2(N)\right\rangle_\omega = \left\langle h^2(N)\right\rangle_t$. The second-moment of the impulse response equals mean-square of the DGD spectrum. This result of Gisin is the formal tie between the impulse and frequency response of birefringent concatenation.

Figure 8.33 shows three concatenation calculations in both the frequency and time domains. The number of segments is four, six, and eight. The left-hand figures plot $\tau(\omega)$ as a function of frequency; the right-hand figures plot $h(t)$, or the impulse response of the cascade. The impulse response is

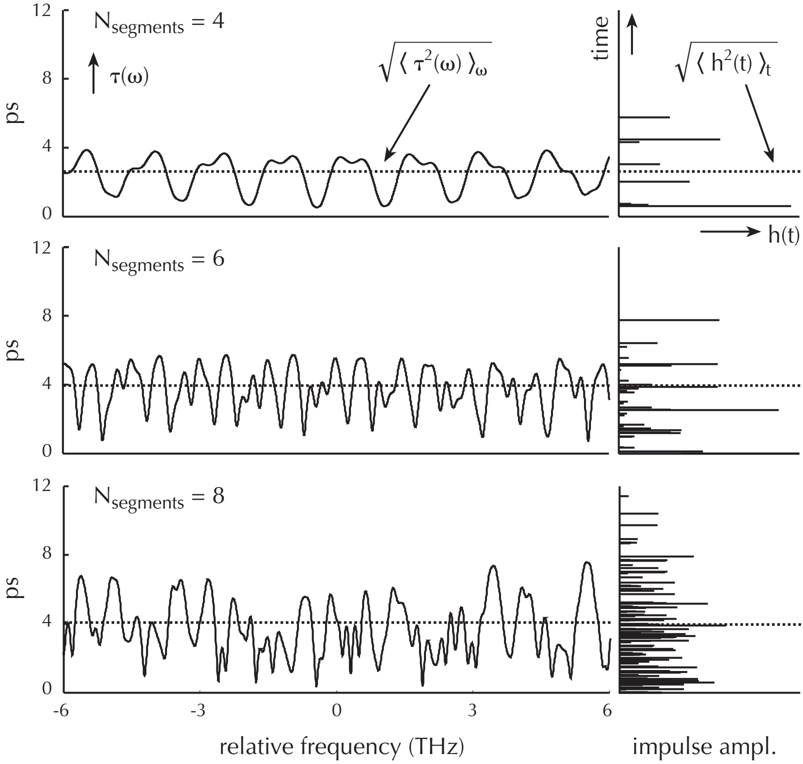

Fig. 8.33. Root-mean-square of DGD spectrum equals the standard deviation of its impulse response for lossless birefringent cascades of various lengths (after [19]). Calculated DGD spectra (left) and impulse responses for $t \geq 0$ (right) are shown for $N_{\mathrm{seg}} = 4, 6, 8$. The dotted line shows the calculated rms DGD (left) and standard deviation (right) of the respective representations. The amplitude of the impulses as illustrated is the square-root of $w(k)$ to indicate all impulses more clearly. The birefringent vector orientations are randomly selected over a uniform distribution on the Poincaré sphere; the DGD values are randomly selected on a Maxwellian distribution where the underlying iid Gaussian distributions have a standard deviation of unity.

symmetric about $t = 0$ because $\hat{s}_{\mathrm{in}} \cdot \hat{r}_1$ was set to zero – the plots shows only the positive side of the response. The calculated rms DGD in frequency and the standard deviation of the impulse spread in time are indicated by the dotted lines. It is remarkable how quickly the two averages converge as the number of segments increases.

Armed with the Gisin and Pellaux result a full circle can be closed on the triad of time-domain analyses of the preceding. The Signal Distortion section that started on page 343 covered the temporal extent of the PMD impulse response and the interference of co-polarized signal images that result from the

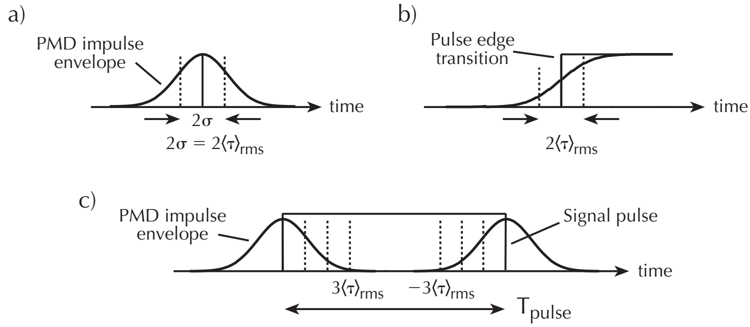

Fig. 8.34. Pulse-width broadening due to PMD. a) For a long, sufficiently random link, the PMD impulse response converges in distribution to a gaussian. Gisin and Pellaux show that the standard deviation is the rms DGD of the link $\sqrt{\langle \tau^2 \rangle}$. b) Leading transition edge is broadened by the impulse response. c) Three standard deviations covers 99.9% of the gaussian. The center of a pulse can be distorted when three standard deviations of the impulse response equal half the pulse width.

convolution. In that section only simple impulse responses were considered. Now, the effects on a signal by the most general of impulse responses can be intuited.

For a long birefringent concatenation sufficiently random, the impulse response envelope converges in distribution to a gaussian shape; this is illustrated in Fig. 8.34(a). As a measure, the width of the gaussian is twice its standard deviation; but it is now known that the standard deviation is the rms DGD of the link. When a transition edge is convolved by the gaussian impulse response, the edge is characteristically broadened by the gaussian width which is twice the rms DGD (Fig. 8.34(b)). Moreover, one standard deviation covers only 84% of the envelope; two and three standard deviations include 97.7% and 99.9% of the gaussian envelope – the transition edges are likely to be broadened further than one standard deviation.

Since the impulse response is the inverse Fourier transform of the entire PMD spectrum, the impulse response does not depend on the particular DGD, SOPMD, and higher-order PMD experienced by the signal spectrum. The signal carrier may be frequency shifted to observe different aspects of the PMD spectrum, but the only effect on the convolution is how the signal images interfere. Therefore, the overlap between a signal "one" and the impulse-response gaussian envelope at the transition edges (Fig. 8.34(c)), remains fixed for any carrier frequency. The tails of the gaussians have reasonable probability three to four standard deviations into the signal; therefore some conditions of interference can excite distortion this far into the signal pulse. One can say roughly that when $T_{\text{pulse}}/2 \simeq 3 \langle \tau \rangle_{\text{rms}}$, the center of the pulse can be effected by PMD. In the next chapter it is shown that for long fibers $\langle \tau \rangle \simeq 1.085 \langle \tau \rangle_{\text{rms}}$, that is, the mean and rms with within 10% of one another. As a measure of complete channel loss due to distortion, the metric of

$$T_{\text{pulse}} \simeq 3\,\langle\tau\rangle \qquad \text{Metric of complete distortion}$$

is commonly used in the industry. For instance, a 10 Gb/s NRZ communication link has 100 ps time slots. The signal can be completely distorted when running on a fiber with accrued mean PMD of $\langle\tau\rangle \simeq 33$ ps.

Concluding Remarks on Time-Domain Studies

The preceding has intended to familiarize the reader with the cornerstone features of the time-domain representation. Many researchers have gone further than this text in the study of particular aspects. Heffner offers several papers on the time/frequency equivalence of PMD and particularly the impact of input signal chirp and coherence on the output distortion [23–25]. The moments analysis of Karlsson is extended by Shieh [57]. The correspondence between time and frequency domains in the study of second-order PMD is principally given by Gisin *et al.* [3], and separately by Penninckx *et al.* [13, 14]. The impact of higher orders of PMD has been studied by Gisin [30] and separately by Kogelnik *et al.* [40]. The degree of polarization as a function of PMD and optical data spectrum is analyzed by Nezam *et al.* [46].

Many researchers have studied the Jones-matrix representation of SOPMD and PMD in general. Significant works come from Eyal [9], Kogelnik [41], and Vincetti [47] for SOPMD; and Heismann [26], Penninckx [49], and Vincetti [12, 48] for general PMD.

8.2.7 Fourier Analysis of the DGD Spectrum

The Fourier analysis of the DGD spectrum offers yet another view into the nature of PMD. This Fourier analysis is not the PMD impulse response nor its Fourier transform into the associated vector and scalar spectra, but strictly a Fourier analysis of one scalar spectrum. The analysis can be extended to include the magnitude-SOPMD scalar spectrum. The purpose of this analysis is twofold. First, to go beyond second-order PMD, one can take higher and higher derivatives of the PMD vector, as reported by Kogelnik [40], or one can look at the degree of variation over a frequency window, as first analyzed by Poole and Favin [54]. In the former case, increased precision is found but over an infinitesimally narrow bandwidth; in the latter case, approximation is made but how the PMD spectrum overlays with the finite-width spectrum of a signal can be better understood.

The Poole and Favin analysis showed that in the long-length regime the expected number of level crossings $\langle N_m\rangle$ of the DGD spectrum through a given level over an interval $\Delta\omega$ is

$$\langle N_m\rangle = \frac{\langle\tau\rangle\,\Delta\omega}{4}$$

The number of level crossings increase linearly with the mean DGD of the link. Therefore the variation of the DGD spectrum must increase linearly with mean DGD as well. For instance, if one conservatively wants to place a channel between two adjacent level crossing, one would take $\langle N_m \rangle = 1$ and estimate the maximum $\langle \tau \rangle$ as $\langle \tau \rangle_{\max} \simeq 4/\Delta\omega$. For a channel bandwidth of $\Delta f = 20$ GHz, the maximum $\langle \tau \rangle$ is $\langle \tau \rangle_{\max} \simeq 33$ ps.

The following analysis is presented for the short-length regime [5, 6, 64]. Its extension to the long-length regime may be possible but to date this has not been done. The purpose of presenting this analysis is to emphasize the origin of oscillatory variation in the DGD spectrum and to exhibit the spin-vector formalism used to arrive at the conclusions. In light of the Gisin and Pellaux time-response derivation based on recurrence relations, it is believed a similar approach can be used to extend the Fourier analysis into the long-length regime.

The problem at hand is similar to that of angular momentum. Analyses of angular momentum relate to coupled spinning objects or particles and the total overall momentum. Often one looks for the probability density of the overall angular momentum given all possible orientations of the component spins. Alternatively, the extrema can be determined. The quantized angular momentum analysis of coupled subatomic spins, such as electron and nuclear spins, determines the quantized levels of the total angular momentum and the state densities.

PMD concatenations are similar because component PMD vectors precess, or spin, about the axes of adjacent component vectors as frequency is swept. While the probability density of the overall PMD pointing direction or PMD vector length can be evaluated, the focus of the present analysis is to determine the Fourier components embedded in the variation of the PMD vector length when frequency is swept. The embedded Fourier components depend only on the delays of the component PMD vectors and not on their relative orientation or the frequency. The amplitude and phase of the Fourier components, however, do determine on these details.

A note on nomenclature. The Fourier content determined in the following refers to the oscillatory rate of a DGD spectrum, that is, the frequency of variation. The use of "frequency" as related to Fourier content differs from the use of "frequency" as related to the carrier frequency of a probe signal that measures the PMD. The optical carrier frequency is completely different from the frequencies of the Fourier content of the DGD spectrum.

The following analysis builds DGD spectra and the respective Fourier components from two, three, and four birefringent stages. The concatenation rules for each are illustrated in Fig. 8.35 and the corresponding DGD spectra and Fourier analysis are illustrated in Fig. 8.36.

The simplest concatenation is that of two vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ (Fig. 8.35(a)). The angle between the two vectors in the diagram is determined by the mode mixing angle between the stages and is not a function of frequency. The output vector is

$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1 \tag{8.2.63}$$

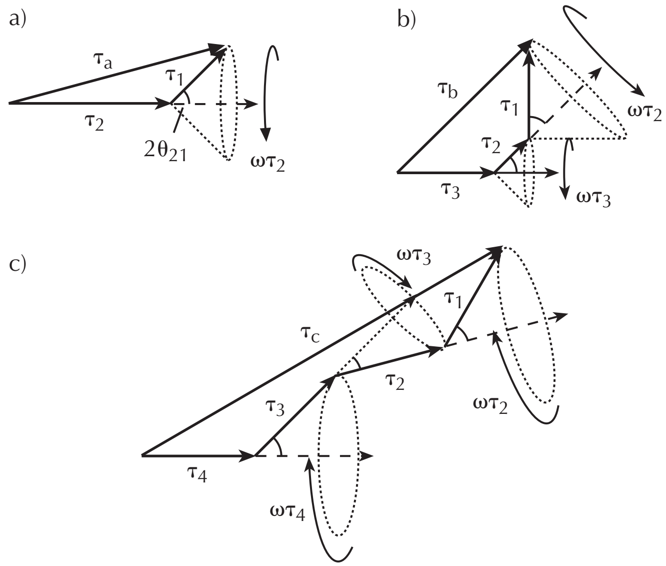

Fig. 8.35. Component PMD vector concatenations for two, three, and four stages. a) Two stages. Angle θ_{21} is determined by the mode mixer and is frequency independent. Vector $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ at rate τ_2. The PMD vector is the sum of its component vectors; the length τ_a is the DGD. b) Three stages, two different precessions: $\omega\tau_2$ and $\omega\tau_3$. c) Four stages, three different precessions: $\omega\tau_2$, $\omega\tau_3$, and $\omega\tau_4$.

where $\vec{\tau}_k = \tau_k \hat{r}_k$, $k = 1, 2$. Component $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with birefringent phase $\varphi = \omega\tau_2$; the free-spectral range is $\text{FSR} = 1/\tau_2$. A note on the order of precession. The figure shows $\vec{\tau}_1$ precessing about $\vec{\tau}_2$, although in the concatenation $\vec{\tau}_1$ comes first. Physically, the cumulative PMD vector $\vec{\tau}$ is defined at the output. Looking from the output toward the input, one sees $\vec{\tau}_2$ immediately and $\vec{\tau}_1$ through the aperture of the second element that generates $\vec{\tau}_2$. When the frequency is changed, the appearance of Stokes orientation $\vec{\tau}_1$ is rotated due to the birefringence of $\vec{\tau}_2$, precisely in the same way a polarization state is altered due to the birefringence of $\vec{\tau}_2$. This gives the precession of $\vec{\tau}_1$ about $\vec{\tau}_2$.

The magnitude-squared of the DGD spectrum is

$$\vec{\tau} \cdot \vec{\tau} = \tau_2^2 + \tau_1^2 + 2\tau_2\tau_1 \, \hat{r}_2 \cdot \hat{r}_1 \tag{8.2.64}$$

The dot product in the last term,

$$\hat{r}_2 \cdot \hat{r}_1 = \cos\theta_{21} \, , \tag{8.2.65}$$

is frequency independent and θ_{21} is a Stokes angle. Since there is no frequency dependence in the DGD spectrum, the spectrum can be characterized as

$$\vec{\tau} \cdot \vec{\tau} = a_0 \tag{8.2.66}$$

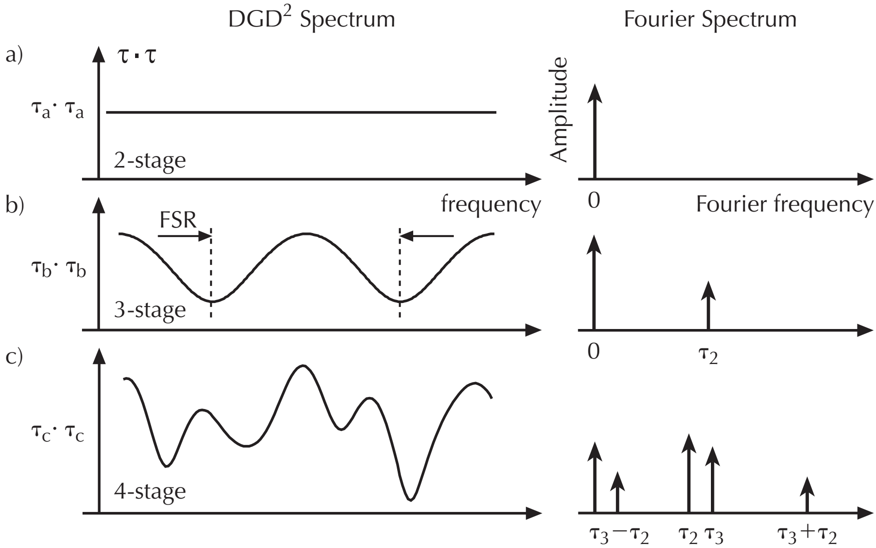

Fig. 8.36. Magnitude-square DGD spectra and associated Fourier analysis corresponding to precession diagrams in Fig. 8.35. The magnitude-square DGD spectra plot $\tau \cdot \tau$ as a function of carrier frequency. The Fourier analyses show the Fourier frequencies that are present in the respective $\tau \cdot \tau$ spectra. Only Fourier amplitudes are shown, although the Fourier phases are not necessarily zero. a) Two-stage spectrum has no oscillatory Fourier components and is governed by (8.2.66). Only a DC Fourier component is present. b) Three-stage spectrum has single oscillatory component with well-defined FSR and is governed by (8.2.70). The Fourier spectrum has a DC plus a single oscillatory component at τ_2; only the center stage dictates the frequency of oscillation. c) Four-stage spectrum has four oscillatory components, governed by (8.2.75). The Fourier spectrum has one constant plus four oscillatory components, including sum and different terms. Only the center two stage delays contribute to the oscillation.

where a_0 is a real number and the Fourier content is only DC, see Fig. 8.36(a).

The next case is the concatenation of three component vectors $\vec{\tau}_1$, $\vec{\tau}_2$, and $\vec{\tau}_3$ (Fig. 8.35(b)). As illustrated, the birefringent axes between the stages are not aligned, which results in mode mixing. The resultant PMD vector is

$$\vec{\tau} = \vec{\tau}_3 + R_3\vec{\tau}_2 + R_3R_2\vec{\tau}_1 \tag{8.2.67}$$

The figure shows the motion of the three vector components. Vector $\vec{\tau}_1$ precesses about the $\vec{\tau}_2$ axis with birefringent phase $\varphi_2 = \omega\tau_2$. Vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ combined precess about the $\vec{\tau}_3$ axis with birefringent phase $\varphi_3 = \omega\tau_3$. The length and pointing direction of $\vec{\tau}$ exhibit a more complicated motion than that of the two-stage example and are in general frequency dependent. The magnitude-squared of the DGD spectrum is

$$\begin{aligned}\vec{\tau}\cdot\vec{\tau} = {} & \tau_3^2+\tau_2^2+\tau_1^2+\\ & 2\tau_3\tau_2\,\hat{r}_3\cdot\hat{r}_2+2\tau_2\tau_1\,\hat{r}_2\cdot\hat{r}_1+\\ & 2\tau_3\tau_1\,\hat{r}_3\cdot(R_2\hat{r}_1)\end{aligned} \tag{8.2.68}$$

In light of (8.2.65), the first five terms on the right-hand side are frequency independent. The last term, however, generates one non-DC Fourier component. The last term expands to

$$\begin{aligned}\hat{r}_3\cdot(R_2\hat{r}_1) = {} & \cos\theta_{32}\cos\theta_{21}+\\ & \sin\varphi_2\,\hat{r}_3\cdot(\hat{r}_2\times\hat{r}_1)-\cos\varphi_2\,\hat{r}_3\cdot(\hat{r}_2\times\hat{r}_2\times\hat{r}_1)\end{aligned} \tag{8.2.69}$$

The last two terms on the right-hand side add to yield a single oscillatory term governed by φ_2 (see Appendix A). Combining Eqs. (8.2.65), (8.2.69), and using the identity

$$A\cos\phi\pm B\sin\phi=\sqrt{A^2+B^2}\,\cos\left(\phi\mp\tan^{-1}B/A\right)$$

yields the Fourier form of the magnitude-squared DGD spectrum:

$$\vec{\tau}\cdot\vec{\tau}=a_0+a_1\cos(\varphi_2-\xi) \tag{8.2.70}$$

where, as before, a_0 and a_1 are real numbers that are independent of frequency. The spectrum is periodic where the periodicity is determined solely by the center section τ_2, see Figs. 8.36(b). The Fourier-component phase shift ξ is determined from (8.2.69):

$$\xi=\tan^{-1}\left[\frac{\hat{r}_3\cdot(\hat{r}_2\times\hat{r}_1)}{\hat{r}_3\cdot(\hat{r}_2\times\hat{r}_2\times\hat{r}_1)}\right] \tag{8.2.71}$$

where in general the phase changes as the relative angles between PMD components change. Only when all three birefringent axes lie in the same plane, which leads to $\hat{r}_3\cdot(\hat{r}_2\times\hat{r}_1)=0$, does phase shift ξ vanish.

The coupling of Fourier-component phase shift ξ to the mode mixing angles $\hat{r}_2\cdot\hat{r}_1$ and $\hat{r}_3\cdot\hat{r}_2$ is an interesting effect that is illustrated in Fig. 8.37. Both three-stage examples in the figure show a range of DGD spectral shapes when the center section is rotated as indicated. When the birefringent axes $\hat{r}_{1,2,3}$ of all three sections lie in the same plane, such as the equator, then Fourier-component phase shift ξ is identically zero. In this case the frequency location of the maximum DGD value does not change even though the shape of the spectrum changes (Fig. 8.37(a)). However, when any one of the birefringent axes lies out of the plane of the other two then coupling between phase shift ξ and mode mixing occurs (Fig. 8.37(b)). This effect is observable, for instance, when a zero-order quarter-wave waveplate is inserted to either side of the center element. In a communication link, the bulk components such as isolators can make the apparent birefringent axes fall outside of a common plane.

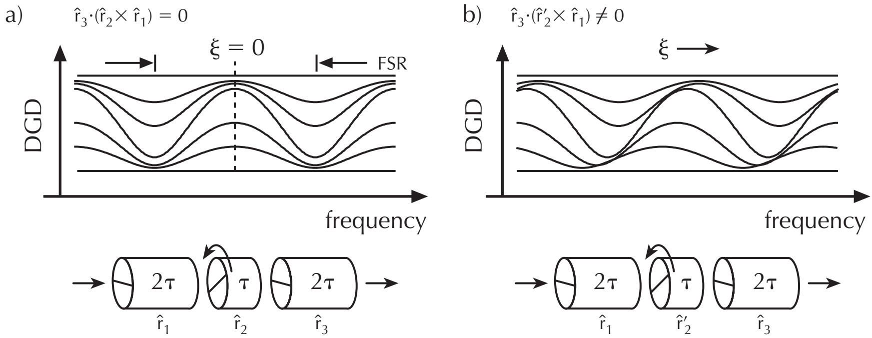

Fig. 8.37. Fourier-component phase shift ξ as decoupled (a) and coupled (b) to mode mixing. a) Locus of DGD spectra for a three-stage system as the center section is rotated. All three birefringent axes lie in the same plane in Stokes space. The frequency location of the maximum DGD is fixed for each spectrum. b) Locus of DGD spectra when one birefringent axes lies out of the plane defined by the other two. Fourier-component phase shift ξ is coupled to the mode mixing.

The following analysis is simplified by assuming all birefringent axes lie in the same plane. This is called the co-planar assumption. Violation of this assumption will not change the location or number of Fourier frequencies, only their phase.

The concatenation of four sections $\vec{\tau}_1$, $\vec{\tau}_2$, $\vec{\tau}_3$, and $\vec{\tau}_4$ is illustrated in Fig. 8.35(c). The resultant PMD vector is

$$\vec{\tau} = \vec{\tau}_4 + R_4\vec{\tau}_3 + R_4R_3\vec{\tau}_2 + R_4R_3R_2\vec{\tau}_1 \tag{8.2.72}$$

The motion of the vector sum is more complicated yet. Vector $\vec{\tau}_1$ precesses about the $\vec{\tau}_2$ axis with phase $\varphi_2 = \omega\tau_2$; vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ combined precess about $\vec{\tau}_3$ with phase $\varphi_3 = \omega\tau_3$; and vectors $\vec{\tau}_{1,2,3}$ combined precess about $\vec{\tau}_4$ with phase $\varphi_4 = \omega\tau_4$. The magnitude-squared DGD spectrum takes the form

$$\begin{aligned} \vec{\tau}\cdot\vec{\tau} = &\sum_{k=1}^{4}\tau_k^2 + 2\sum_{k=1}^{3}\tau_{k+1}\tau_k\ \hat{r}_{k+1}\cdot\hat{r}_k + \\ &2\sum_{k=1}^{2}\tau_{k+2}\tau_k\ \hat{r}_{k+2}\cdot(R_{k+1}\hat{r}_k) +\ 2\tau_4\tau_1\ \hat{r}_4\cdot(R_3R_2\hat{r}_1) \end{aligned} \tag{8.2.73}$$

The first term on the right-hand side of (8.2.73) is scalar; the second term, identified with (8.2.65), generates no frequency-dependent terms; and the third term, identified with (8.2.69), generates the frequency-dependent terms $\cos\varphi_2$ and $\cos\varphi_3$ (assuming coplanar mode-mixing vectors). The last term generates additional frequency-dependent components. That term expands to

$$\begin{aligned}\hat{r}_4\cdot(R_3R_2\hat{r}_1) = &\cos\theta_{43}\cos\theta_{32}\cos\theta_{21}\\ &-\ \hat{r}_3\cdot(\hat{r}_2\times\hat{r}_2\times\hat{r}_1)\cos\theta_{43}\cos\varphi_2\\ &-\ \hat{r}_4\cdot(\hat{r}_3\times\hat{r}_3\times\hat{r}_2)\cos\theta_{21}\cos\varphi_3\\ &+\ \hat{r}_4\cdot(\hat{r}_3\times\hat{r}_2\times\hat{r}_1)\sin\varphi_3\sin\varphi_2\\ &+\ \hat{r}_4\cdot(\hat{r}_3\times\hat{r}_3\times\hat{r}_2\times\hat{r}_2\times\hat{r}_1)\cos\varphi_3\cos\varphi 2\end{aligned} \tag{8.2.74}$$

The mixing products $\sin\varphi_3\sin\varphi_2$ and their cosine complements resolve themselves into sum and difference terms, e.g.

$$2\sin\varphi_3\sin\varphi_2 = \cos(\varphi_3-\varphi_2)-\cos(\varphi_3+\varphi_2)$$

The Fourier content of a four-stage concatenation therefore takes the form

$$\begin{aligned}\vec{\tau}\cdot\vec{\tau} = a_0 + a_1\cos\omega\tau_2 + a_2\cos\omega\tau_3 +&\\ a_3\cos\omega(\tau_3-\tau_2) + a_4\cos\omega(\tau_3+\tau_2)&\end{aligned} \tag{8.2.75}$$

where, as before, coefficients a_k are real, frequency-independent numbers that are determined solely by the mode mixing between stages. An exemplar spectra and its Fourier spectrum are illustrated in Figs. 8.36(c).

Several observations about (8.2.75) are made. First, the Fourier content of the magnitude-squared DGD spectrum is determined only by the center two delay stages. In general the first and last section for any number of stages do not contribute to the Fourier content; this observation is proven below. Second, sum and difference Fourier terms are present in addition to the Fourier components associated with the center two stages. Here there are four Fourier components in total. Third, there is a Fourier-component generator function that generates these components. The generator function $\mathcal{G}(N)$ is

$$\mathcal{G}(N) = \hat{r}_N\cdot R(N,2)\,\hat{r}_1 \tag{8.2.76}$$

where the cumulative operator $R(N,2)$ is defined by (8.2.38) on page 337. The first four generator functions are

$$\begin{aligned}\mathcal{G}(1) &= 1\\ \mathcal{G}(2) &= g_0\\ \mathcal{G}(3) &= g_0 + g_1\cos\varphi_2\\ \mathcal{G}(4) &= g_0 + g_1\cos\varphi_2 + g_2\cos\varphi_3 +\\ &\qquad g_3\cos(\varphi_3-\varphi_2) + g_4\cos(\varphi_3+\varphi_2)\end{aligned} \tag{8.2.77}$$

where the value g_k for one generator function has no relation to the value g_k of another generator function.

As a last part to this section, the absence of Fourier components generated from the first and last stages is shown [61]. The independence of the first stage has already been demonstrated: it is clear from any of the vector diagrams in Fig. 8.35 that no vector precesses about $\vec{\tau}_1$, so there is no sensitivity of $\vec{\tau}$ to $\omega\tau_1$. The insensitivity to $\vec{\tau}_N$ is proven by separating the last stage from the preceding $N-1$ stages:

$$\begin{aligned}\vec{\tau}\cdot\vec{\tau} &= \left(\vec{\tau}_N + R_N\vec{\tau}_{(N-1)}\right)\cdot\left(\vec{\tau}_N + R_N\vec{\tau}_{(N-1)}\right)\\ &= \tau_N^2 + \tau_{(N-1)}^2 + 2\tau_N\tau_{(N-1)}\ \hat{r}_N\cdot\hat{r}_{(N-1)}\end{aligned} \tag{8.2.78}$$

Since the last term on the right-hand side has the form $\mathcal{G}(2)$, there is in fact no φ_N Fourier component generated by the last stage. Geometrically this makes sense because rotation about $\vec{\tau}_N$ pirouettes the remaining vector structure, changing its pointing direction but not its length.

8.3 Combined Effects of PMD and PDL

When birefringent elements and partial polarizers are interspersed along a concatenation, the resulting polarization effects are more complicated than PDL or PMD alone. The resulting effects form a superset of PDL and PMD: communication impairments due to combined PDL and PMD can be more severe than either isolated effect. Since there is no separate name for the combined phenomena, "PMD and PDL" here denotes the aggregate effect.

The principal results from combined PMD and PDL are

1. The principal-states of polarization are not orthogonal to one another.
2. The output polarization state does not follow simple precessional motion as a function of frequency.
3. The polarization-dependent loss is frequency dependent. A new term, differential-attenuation slope (DAS) is introduced to characterize this effect.

These results conspire to create anomalous distortions. For instance, the differential-group delay of a PMD and PDL combination can be greater than the sum of the individual delays [32]. Or, a pulse can suffer spreading even with zero net DGD [30]. Moreover, neither an optical PDL nor PMD compensator can be perfectly realized.

The original work on PMD and PDL is by Frigo [16], who derived the equation of motion of the output polarization vector as function of frequency, albeit in the context of coupled-mode equations. Eyal [10] later recast the Frigo equation into unit-vector form to derive the complex motion of the output polarization state with frequency. The authors Gisin, Huttner, and Geiser [18, 31] discovered many anomalous effects due to the interaction of PMD and PDL. R. C. Jones also contributed on this subject: his 1941 paper proves a theorem, theorem 3, that any number of PMD and PDL elements (although he called them retarders and partial polarizers) can be replaced with just four elements – two PMD elements, one PDL element, and one optically active polarization rotator [36]. This theorem has practical use when separating PMD and PDL effects.

Further research on the interaction of PMD and PDL in an optical communications link has been reported by Mollenauer [62, 63], Feced [11], and Eyal [8]. A very interesting measurement method to test for maximum eye-opening excursion has been invented by Kuperman *et al.* [42].

Three principal equations are derived in the following: the change of output polarization state with frequency; the change of output polarization state through propagation; and the cumulative PDL vector equation of motion. Table 8.3 compares the expressions for pure PMD and those including PDL.

8.3.1 Frequency-Dependence of the Polarization State

In general, the transformation matrix T between input and output polarization states is neither unitary nor Hermitian in the presence of combined PMD and PDL. Due to the birefringence in the medium, T is frequency dependent: $|t\rangle = T(\omega)\,|s\rangle$. As before, for $|s\rangle$ fixed in frequency, the frequency-change of the output state is $|t\rangle_\omega = T_\omega\,|s\rangle$. As long as a perfect polarizer is not placed between input and output, T^{-1} exists and the change of output state is expressed as

$$|t\rangle_\omega = T_\omega T^{-1}\,|t\rangle \tag{8.3.1}$$

Even without any particular reference to PMD and PDL parameters, the matrix $T_\omega T^{-1}$ can always be decomposed into a Hermitian, skew-Hermitian, and trace components, as was shown by (2.5.72) on page 61. The decomposition of $T_\omega T^{-1}$ takes the form

$$jT_\omega T^{-1} = \frac{1}{2}\left(\vec{\Omega}_r\cdot\vec{\sigma} + j\,\vec{\Omega}_i\cdot\vec{\sigma} + a_0 I\right) \tag{8.3.2}$$

The factor of j in front of $T_\omega T^{-1}$ is added to keep the notation parallel to the Hermitian operator $jU_\omega U^\dagger$ defined for the description of PMD. The scalar a_0 is in general complex; the real component is the common phase through the system and the imaginary part is the common loss. The vectors $\vec{\Omega}_r$ and $\vec{\Omega}_i$ are real-values Stokes vectors. These vectors relate to the system birefringence and differential attenuation, but not is a straight-forward way. The decomposition parameters are related to $jT_\omega T^{-1}$ via (2.5.68) on page 61 for the trace and (2.5.73) for the real and imaginary parts.

Since T is not necessarily unitary, the length of the output Stokes vector $\vec{t}$ is generally not the same as the input vector $\vec{s}$. One must be careful to distinguish the output vector length t and the unit vector $\hat{t}$. With this remark in mind, the Stokes vector $\vec{t}_\omega$ is calculated in the same way as (8.2.15) on page 330. This makes

$$\vec{t}_\omega = \langle t\left|\left(T_\omega T^{-1}\right)^\dagger\vec{\sigma} + \vec{\sigma}\left(T_\omega T^{-1}\right)\right|t\rangle \tag{8.3.3}$$

Substitution of $T_\omega T^{-1}$ (8.3.2) into (8.3.3) generates the Frigo equation of motion:

$$\frac{d\vec{t}}{d\omega} = \vec{\Omega}_r \times \vec{t} + a_{i,0}\,\vec{t} + t\,\vec{\Omega}_i \tag{8.3.4}$$

where $a_{i,0}$ is the imaginary part of the trace of $T_\omega T^{-1}$. This equation describes a complex behavior of the output Stokes vector $\vec{t}$. The first term on the right-hand side generates a precessional motion: $\vec{t}$ precesses about $\vec{\Omega}_r$ as a function of frequency. In the absence of PDL, $\vec{\Omega}_r = \vec{\tau}$, the PMD vector. In the presence of PDL, $\vec{\Omega}_r$ includes PDL as well as PMD terms. The second term on the right-hand side describes the growth or decay of $\vec{t}$ along its own axis. The imaginary part of the trace of $T_\omega T^{-1}$ governs this behavior. Also, these first two terms run perpendicular to one another. Finally, the third term on the right-hand side pulls $\vec{t}$ toward $\vec{\Omega}_i$. The pulling behavior has been seen before in §8.1.2 in regard to PDL. In sum, there are three distinct axes along which the output state is changed.

There is a competition setup between $\vec{\Omega}_r$ and $\vec{\Omega}_i$. If the former is the dominant term, then it acts to retard the growth or decay of $\vec{t}$ by generating a motion perpendicular to $\vec{t}$. If the latter term dominates, then $\vec{t}$ grows or decays without bound.

Further insight is found by decomposing (8.3.4) into coupled unit-vector and vector-length equations of motion [15]. The decomposition requires two identifications. First, by definition $\vec{t} = t\hat{t}$, so the frequency derivative is

$$\frac{dt\hat{t}}{d\omega} = t\frac{d\hat{t}}{d\omega} + \hat{t}\frac{dt}{d\omega}$$

Note that the first term is perpendicular to $\vec{t}$ while the second term is parallel (that the first term is perpendicular is a consequence of $\hat{t}$ being a unit vector). Second, the vector $\vec{\Omega}_i$ is decomposed into components parallel and perpendicular to $\vec{t}$:

$$\begin{aligned}\vec{\Omega}_i &= \vec{\Omega}_{i,\parallel} + \vec{\Omega}_{i,\perp}\\ &= \left(\vec{\Omega}_i \cdot \hat{t}\right)\hat{t} - \left(\vec{\Omega}_i \cdot \hat{t}\right)\hat{t} + \vec{\Omega}_i\left(\hat{t}\cdot\hat{t}\right)\\ &= \left(\vec{\Omega}_i \cdot \hat{t}\right)\hat{t} + \left(\hat{t} \times \vec{\Omega}_i\right) \times \hat{t}\end{aligned}$$

Substitution into the equation of motion makes

$$t\frac{d\hat{t}}{d\omega} + \hat{t}\frac{dt}{d\omega} = t\,\vec{\Omega}_r \times \hat{t} + (a_{i,0}t)\hat{t} + t\,\left(\vec{\Omega}_i \cdot \hat{t}\right)\hat{t} + t\,\left(\hat{t} \times \vec{\Omega}_i\right) \times \hat{t}$$

The parallel and perpendicular components must separately satisfy this equation. Therefore the decomposition produces the two coupled equations

$$\frac{dt}{d\omega} = a_{i,0}t + \left(\vec{\Omega}_i \cdot \hat{t}\right)t \tag{8.3.5a}$$

$$\frac{d\hat{t}}{d\omega} = \vec{\Omega}_r \times \hat{t} - \left(\vec{\Omega}_i \times \hat{t}\right) \times \hat{t} \tag{8.3.5b}$$

These equations exhibit interesting properties. First, unit-vector equation is not coupled to the vector-length equation. Therefore these equations can be solved in sequence: first solve for $\hat{t}$, then for t. Second, the unit-vector equation of motion has both double- and triple-vector products. This result was first published by Eyal [10]. Third, the change of length equation is driven explicitly by $a_{i,0}$ and $\vec{\Omega}_i$, both directly related to PDL. If the PDL were zero, these terms would be absent and the vector length would be invariant. With PDL present, t changes with frequency – this is the origin of the differential-attenuation slope, or DAS. Even when the common loss generated by $a_{i,0}$ is transformed out, the differential attenuation embedded in $\vec{\Omega}_i$ induces DAS.

8.3.2 Non-Orthogonality of PSP's

The principal states of polarization for PMD and PDL are found in the same way as the PSP's are for pure PMD. Recall from (8.2.9) on page 329 that the PSP's are defined by the eigenvalue equation of the operator $jU_\omega U^\dagger$. When this equation is satisfied, the output polarization state is stationary to first-order in frequency.

In an entirely analogous way, the eigenvalue equation for $jT_\omega T^{-1}$ is defined. Substitution of the spin-vector form of $jT_\omega T^{-1}$ into (8.3.1) yields

$$|t\rangle_\omega = -\frac{j}{2}\left(\vec{\Omega}_r \cdot \vec{\sigma} + j\,\vec{\Omega}_i \cdot \vec{\sigma}\right)|t\rangle - j(a_0/2)\,|t\rangle \tag{8.3.6}$$

To make the output state stationary, the spin-vector operator must collapse to a complex scalar value:

$$\vec{\Omega} \cdot \vec{\sigma}\,|\tilde{p}_\pm\rangle = \pm\lambda\,|\tilde{p}_\pm\rangle \tag{8.3.7}$$

where $\vec{\Omega} = \vec{\Omega}_r + j\vec{\Omega}_i$. That the eigenvalues λ are equal and opposite is a direct consequence of the zero trace of $\vec{\Omega}$. Moreover, since the operator $\vec{\Omega} \cdot \vec{\sigma}$ is non-Hermitian, the eigenvalues are in general complex. Gisin *et al.* identify the real and imaginary parts of the eigenvalues as [31]

$$\lambda = \tau + j\eta \tag{8.3.8}$$

The real part τ is the familiar differential-group delay magnitude; the imaginary part η is the differential-attenuation slope (DAS), which is the frequency derivative of the differential attenuation along the two eigenvectors.

The eigenvalue λ and the operator $\vec{\Omega} \cdot \vec{\sigma}$ are related by

$$\lambda^2 = \vec{\Omega} \cdot \vec{\Omega} \tag{8.3.9}$$

This expression may be confirmed by solving $\det(\vec{\Omega} \cdot \vec{\sigma} - \lambda I) = 0$ and is analogous to the case for pure PMD where $\vec{\tau} \cdot \vec{\tau} = \tau^2$.

In the presence of PDL, the PSP's are not orthogonal (except, possibly, in transient or pathological cases). This is due to the complex value of the operator $\vec{\Omega}$. The overlap γ^2 of the two eigenvectors can be computed in Stokes space from the dot-product $\hat{\tilde{p}}_+ \cdot \hat{\tilde{p}}_-$, see (2.5.65) on page 60. The calculation is simplified by rearranging (8.3.7) so that the operator has unit length and the eigenvalues are real:

$$\hat{\Omega} \cdot \vec{\sigma} \, |\tilde{p}_\pm\rangle = \pm \, |\tilde{p}_\pm\rangle$$

where $\hat{\Omega} = \vec{\Omega}/\lambda$, $\hat{\Omega} = \vec{w}_r + j\vec{w}_i$, and $\hat{\Omega} \cdot \hat{\Omega} = 1$. From this eigenvalue equation two auxiliary equations are computed, the first by multiplying the equation by $\langle \tilde{p}_\pm \, |$ and the second by $\langle \tilde{p}_\pm \, | \vec{\sigma}$:

$$\hat{\Omega} \cdot \langle \tilde{p}_\pm \, |\vec{\sigma}| \, \tilde{p}_\pm \rangle = \pm \, \langle \tilde{p}_\pm \, |\tilde{p}_\pm \rangle$$
$$\langle \tilde{p}_\pm \, |\vec{\sigma}(\hat{\Omega} \cdot \vec{\sigma})| \, \tilde{p}_\pm \rangle = \pm \, \langle \tilde{p}_\pm \, |\vec{\sigma}| \, \tilde{p}_\pm \rangle$$

Conversion to Stokes space gives

$$\hat{\Omega} \cdot \hat{p}_\pm = \pm \, 1, \quad \text{and} \quad \hat{\Omega} + j\hat{\Omega} \times \hat{p}_\pm = \pm \, \hat{p}_\pm \tag{8.3.10}$$

where the tilde has been removed for brevity. Now, since $\hat{\Omega} \cdot \hat{\Omega} = 1$, the imaginary part of the dot product must vanish: this requires $\vec{w}_r \cdot \vec{w}_i = 0$. Therefore an orthogonal group of three (unnormalized) axes may be constructed, $(\vec{w}_r, \vec{w}_i, \vec{w}_r \times \vec{w}_i)$, and $\hat{p}_\pm$ can be projected onto this basis:

$$\hat{p}_\pm = c_r \vec{w}_r + c_r \vec{w}_i + c_\times \left(\vec{w}_r \times \vec{w}_i \right) \tag{8.3.11}$$

The real-valued coefficients are isolated through the dot products

$$c_r \vec{w}_r \cdot \vec{w}_r = \vec{w}_r \cdot \hat{p}_\pm$$
$$c_i \vec{w}_i \cdot \vec{w}_i = \vec{w}_i \cdot \hat{p}_\pm$$
$$c_\times \left(\vec{w}_r \times \vec{w}_i \right) \cdot \left(\vec{w}_r \times \vec{w}_i \right) = \left(\vec{w}_r \times \vec{w}_i \right) \cdot \hat{p}_\pm$$

Given that $\hat{\Omega} \cdot \hat{p}_\pm = \pm \, 1$, the first two coefficients are $c_r = 1/\left(\vec{w}_r \cdot \vec{w}_r \right)$ and $c_i = 0$. The third coefficient is evaluated from the dot-product and the second auxiliary equation (8.3.11)

$$c_\times = \frac{\left(\vec{w}_r \times \vec{w}_i \right) \cdot \hat{p}_\pm}{w_r^2 w_i^2} \longrightarrow c_\times = \frac{1}{w_r^2}$$

The normalized eigenvectors in Stokes space are then [31]

$$\hat{p}_\pm = \frac{\pm \, \vec{w}_r + \vec{w}_r \times \vec{w}_i}{\vec{w}_r \cdot \vec{w}_r} \tag{8.3.12}$$

Using the fact that $\hat{\Omega} \cdot \hat{\Omega}^* = w_r^2 + w_i^2$ and $\vec{\Omega} \cdot \vec{\Omega}^* = |\lambda|^2$, the overlap of the eigenvectors is computed as

$$\gamma^2 = \frac{1 + \hat{p}_+ \cdot \hat{p}_-}{2} = \frac{w_r^2 + w_i^2 - 1}{w_r^2 + w_i^2 + 1}$$
$$= \frac{|\vec{\Omega}|^2 - |\lambda|^2}{|\vec{\Omega}|^2 + |\lambda|^2} \tag{8.3.13}$$

Clearly when $\vec{\Omega}$ is real, the case for pure PMD, the overlap integral vanishes. However, addition of any PDL at all pulls the two PSP's away from an orthogonal orientation.

8.3.3 PMD and PDL Evolution Equations

There are two evolution equations to derive, both being extensions of the pure PMD and pure PDL case. First, the evolution of the complex operator $\vec{\Omega}$ as a function of length is derived; the analogue to this equation is (8.2.39) on page 339, although here a different derivation is employed. Second, the evolution of the cumulative PDL $\vec{\Gamma}$ is derived; the analogue is (8.1.27) on page 310. In both cases, the combined effects of birefringence and PDL are accounted for.

Earlier, the evolution equation for $\vec{\tau}$ was derived by combining the partial derivatives of the output state $\vec{t}$ with respect to both length and frequency. This is the Poole method. The present situation is more difficult because both the vector direction and length vary with length and frequency. Instead, the method of Gisin *et al.* is used [18]. Their method is similar to that used in §8.1.4 except that sections are taken as discrete rather than in the continuum limit.

Given a transformation T such that $|t\rangle = T\,|s\rangle$, the transformation is partitioned into N homogeneous birefringent and lossy sections: $T = A_N T_N$, where A_N is the common loss and T_N represents the product of transformation matrices through N sections, $T_N = T_n T_{n-1} \dots T_1$. Capital subscripts denote section products and lower-case subscripts denote particular sections. The terms A_N and T_N are the discrete analogue to the continuous expression (8.1.22) on page 309.

To account for birefringence and loss, the spin-vector operator for each section is written as

$$T_n = \exp\left(\frac{(-jw\vec{\tau}_n + \vec{\alpha}_n)\cdot\vec{\sigma}}{2}\right) \tag{8.3.14}$$

This definition of T_n is not totally general because the vector direction of the loss and birefringence are aligned; this is a reasonable model because the origin of differential loss and birefringence (in the perturbation regime) is likely due to the same disturbance. A shorthand variable $\vec{g}$ is defined as $\vec{g}_n = -j\omega\vec{\tau}_n + \vec{\alpha}_n$, and $\vec{g} = g\hat{g}$.

The last element of the concatenation is separated from the remaining by writing $T_N = T_n T_{N-1}$. Given $T_{\omega,N} = T_{\omega,n}T_{N-1} + T_n T_{\omega,N-1}$, the operator $T_{\omega,N}T_N^{-1}$ may be written in incremental form as

$$T_{\omega,N}T_N^{-1} = \frac{-j\left(\vec{\tau}_n\cdot\vec{\sigma}\right)}{2} + \exp\left(\frac{\vec{g}_n\cdot\vec{\sigma}}{2}\right) T_{\omega,N-1}T_{N-1}^{-1} \exp\left(\frac{-\vec{g}_n\cdot\vec{\sigma}}{2}\right)$$

With the identification $T_{\omega,k}T_k^{-1} = -j/2\,(\vec{\Omega}_k\cdot\vec{\sigma})$, the operator is rewritten as

Table 8.3. Comparison of Pure PMD and Entangled PMD + PDL

PMD	**PMD + PDL**
$jU_\omega U^\dagger \lvert p\rangle = \pm\tau/2 \lvert p\rangle$	$jT_\omega T^{-1} \lvert\tilde{p}\rangle = \pm\lambda/2 \lvert\tilde{p}\rangle + a_0/2 \lvert\tilde{p}\rangle$
$jU_\omega U^\dagger = \frac{1}{2}(\vec{\tau}\cdot\vec{\sigma})$	$jT_\omega T^{-1} = \frac{1}{2}(\vec{\Omega}\cdot\vec{\sigma}) + a_0/2$
$\left(jU_\omega U^\dagger\right)^\dagger = jU_\omega U^\dagger$	$\left(jT_\omega T^{-1}\right)^\dagger \neq jT_\omega T^{-1}$
$\vec{\tau}, \tau \to$ real	$\vec{\Omega}, \lambda \to$ complex
$\tau \to$ DGD	$\lambda = \tau + j\eta \to$ DGD + DAS
$\tau^2 = \vec{\tau}\cdot\vec{\tau}$	$\lambda^2 = \vec{\Omega}\cdot\vec{\Omega}$
$\hat{p} = \langle p \lvert\vec{\sigma}\rvert p\rangle \to$ PSP	$\hat{p} = \langle p \lvert\vec{\sigma}\rvert p\rangle \to$ PSP
$(1 + \hat{p}_+\cdot\hat{p}_-)/2 = 0$	$(1 + \hat{p}_+\cdot\hat{p}_-)/2 \geq 0$
$\partial\vec{\tau}/\partial z = \vec{\beta}_\omega + \vec{\beta}\times\vec{\tau}$	$\partial\vec{\Omega}/\partial z = \vec{\beta}_\omega + (\vec{\beta} + j\vec{\alpha})\times\vec{\tau}$

$$\vec{\Omega}_N\cdot\vec{\sigma} = \vec{\tau}_n\cdot\vec{\sigma} + \exp\left(\frac{\vec{g}_n\cdot\vec{\sigma}}{2}\right)\left(\vec{\Omega}_{N-1}\cdot\vec{\sigma}\right)\exp\left(\frac{-\vec{g}_n\cdot\vec{\sigma}}{2}\right)$$

Use of the complex spin-vector operator expansion (2.5.8) on page 63, the relevant spin-vector identities, and identification of the embedded equation gives

$$\vec{\Omega}_N = \vec{\tau}_n + \left(\vec{\Omega}_{N-1}\cdot\hat{g}_n\right)\hat{g}_n + \\ \cosh g_n\left(\vec{\Omega}_{N-1} - \left(\vec{\Omega}_{N-1}\cdot\hat{g}_n\right)\hat{g}_n\right) - j\sinh g_n\left(\vec{\Omega}_{N-1}\times\hat{g}_n\right)$$

To make the differential operator $\partial_z\vec{\Omega}_N$, a small increment of length is characterized by a small magnitude g_n, and the birefringence and PDL are written in per-length form as $\vec{\tau}(z)$ and $\vec{\alpha}(z)$. Moreover, recognize that $\tau(z) = \Delta n/c = \beta_\omega$. The resultant equation of motion for $\vec{\Omega}$ is

$$\frac{\partial\vec{\Omega}}{\partial z} = \vec{\beta}_\omega + \left(\vec{\beta} + j\vec{\alpha}(z)\right)\times\vec{\Omega} \qquad (8.3.15)$$

In comparison with the pure PMD evolution equation (8.2.39), PDL adds to the curl of $\vec{\Omega}$ and drives the vector to a complex quantity. Li and Yariv have worked out the analytic solutions (8.3.15) in [43].

Regarding the cumulative PDL vector $\vec{\Gamma}$, the equation of motion is derived in the same way as that shown in §8.1.4 but with the transformation operator (8.3.14) substituted for that in (8.1.22). The equations of motion for $\vec{\Gamma}$ and the transmission of depolarized light are

$$\frac{d\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - \left(\vec{\alpha} \cdot \vec{\Gamma}\right) \vec{\Gamma} \tag{8.3.16a}$$

$$\frac{d\,T_{\text{depol}}}{dz} = \left(\vec{\alpha} \cdot \vec{\Gamma} - \alpha\right) T_{\text{depol}} \tag{8.3.16b}$$

While the depolarized transmission equation is the same once PMD is included, the cumulative PDL equation has a new term: the $\vec{\beta} \times \vec{\Gamma}$ generates a rotation of $\vec{\Gamma}$ about the local birefringence vector $\vec{\beta}$. This rotation is to be expected since linear birefringence always generates precessional motion in Stokes space.

8.3.4 Separation of PMD and PDL

In 1941 R.C. Jones showed that most any Jones matrix generated by any number of retarders and partial polarizers can always be reconstructed with two retarders and one partial polarizer such that

$$\mathbf{J}(\omega) = U(\omega)P(\omega)V(\omega) \tag{8.3.17}$$

where P represents a partial polarizer and U and V are unitary matrices [36]. In general each matrix is a function of optical frequency. The partial polarizer is a Hermitian matrix; accordingly it has real eigenvalues and perpendicular eigenvectors. Such a matrix can be decomposed in $H = S\Lambda S^{\dagger}$, where S is a matrix whose columns are the eigenvectors of H and Λ is a diagonal matrix whose entries are the corresponding eigenvalues.

The unitary operator to the left or right of P can be absorbed in the following way: decompose P and absorb one of its neighbors into a unitary matrix:

$$\begin{aligned}\mathbf{J}(\omega) &= US\Lambda S^{\dagger}V \\ &= UV(S^{\dagger}V)^{\dagger}\Lambda(S^{\dagger}V) \\ &= \tilde{U}(\omega)\tilde{P}(\omega)\end{aligned} \tag{8.3.18}$$

where $\tilde{U} = UV$ and $\tilde{P} = (S^{\dagger}V)^{\dagger}\Lambda(S^{\dagger}V)$. Likewise, $\mathbf{J} = \tilde{P}'\tilde{V}'$.

Consider a link composed of PMD and PDL. The transformation matrix can be written $T(\omega) = P(\omega)U(\omega)$. At any particular frequency all of the differential attenuation is concentrated in matrix P: the eigenvectors of P determine the Stokes direction of $\vec{\Gamma}$ and its eigenvalues are the maximum and minimum transmission. The PDL component can be isolated from T by taking advantage of the Hermitian properties of P:

$$TT^{\dagger} = P^2 = S\Lambda^2 S^{\dagger} \tag{8.3.19}$$

The eigenvectors of $TT^{\dagger}$ point in the direction of the cumulative PDL. Also, since the eigenvalues of P are real, the entries in Λ^2 must be positive. Note that the magnitude and direction of the PDL vector is in general frequency dependent; the dependence is governed by the birefringence of the link which is concentrated in U. This shows that even with the decomposition PU, the PDL and PMD remain entangled.

The unitary matrix U is found once the PDL matrix P is calculated from $TT^\dagger$:

$$U(\omega) = P^{-1}(\omega)T(\omega) \tag{8.3.20}$$

Given U it is tempting to calculate the PMD properties from $jU_\omega U^\dagger$. For small PDL this form of U provides a correction to T for an the investigator who wants to isolate the PMD effects. Both Shtengel and Karlsson have reported using this correction [33, 39]. Although suitable to remove perturbations, one should keep in mind that τ generated from $jU_\omega U^\dagger$ is not the same as τ generated from $T_\omega T^{-1}$. Huttner *et al.* define an effective PMD τ_{eff} for $jU_\omega U^\dagger$ to highlight the fact that τ and τ_{eff} are two different quantities [31].

Table 8.4. Table of Important SOP, PDL, and PMD Relations

Evolution Equations

SOP:

$$\frac{d\hat{s}}{dz} = \vec{\beta} \times \hat{s}$$

PDL:

$$\frac{d\,\vec{\Gamma}}{dz} = \vec{\alpha} - \left(\vec{\alpha} \cdot \vec{\Gamma}\right)\vec{\Gamma}$$

$$\frac{d\,T_{\text{depol}}}{dz} = \left(\vec{\alpha} \cdot \vec{\Gamma} - \alpha\right) T_{\text{depol}}$$

PMD:

$$\frac{d\vec{\tau}}{dz} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau}$$

$$\frac{d\vec{\tau}_\omega}{dz} = \vec{\beta}_{\omega\omega} + \vec{\beta} \times \vec{\tau}_\omega + \vec{\beta}_\omega \times \vec{\tau}$$

$$\frac{d\hat{s}}{d\omega} = \vec{\tau} \times \hat{s}$$

PMD+PDL:

$$\frac{d\,\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - \left(\vec{\alpha} \cdot \vec{\Gamma}\right)\vec{\Gamma}$$

$$\frac{d\,T_{\text{depol}}}{dz} = \left(\vec{\alpha} \cdot \vec{\Gamma} - \alpha\right) T_{\text{depol}}$$

$$\frac{d\hat{s}}{d\omega} = \vec{\Omega}_r \times \hat{s} - \left(\vec{\Omega}_i \times \hat{s}\right) \times \hat{s}$$

Defining Expressions

SOP:

$$|t\rangle = U\,|s\rangle$$

PDL:

$$T_p = \langle t\,|t\rangle = \langle s\,|\,P^\dagger P\,|s\rangle$$

$$\rho_{\text{dB}} \equiv 10\log_{10}\left(\frac{T_{\max}}{T_{\min}}\right)$$

PMD:

$$jU_\omega U^\dagger\,|p_\pm\rangle = \pm\,\tau/2\,|p_\pm\rangle$$

$$jU_\omega U^\dagger = \frac{1}{2}\left(\vec{\tau} \cdot \vec{\sigma}\right)$$

$$\vec{\tau}\times = R_\omega R^\dagger$$

PMD+PDL:

$$\vec{\Omega} \cdot \vec{\sigma}\,|\tilde{p}_\pm\rangle = \pm\left(\tau + j\eta\right)|\tilde{p}_\pm\rangle$$

PMD Concatenation

$$\vec{\tau} = \sum_{k=1}^{n} R(n, k+1)\,\vec{\tau}_n$$

$$\vec{\tau}_\omega = \sum_{k=1}^{n} R(n, k+1)\left(\vec{\tau}_{n\omega} + \vec{\tau}_n \times \vec{\tau}(n)\right)$$

$$R(n,k) = R_n R_{n-1} \cdots R_k$$

References

1. D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurement of the group-delay difference between the principal states of polarization on a low-birefringent terrestrial fiber cable," *Optics Letters*, vol. 12, no. 10, pp. 844–846, 1987.
2. A. J. Barlow, "Birefringentce and polarization mode dispersion in spun single mode fibers," *Applied Optics*, vol. 20, no. 17, p. 2962, 1981.
3. P. Ciprut, B. Gisin, N. Gisin, R. Passy, J. Weid, F. Prieto, and C. W. Zimmer, "Second-order polarization mode dispersion: Impact on analog and digital transmissions," *Journal of Lightwave Technology*, vol. 16, no. 5, pp. 757–771, May 1998.
4. F. Curti, B. Daino, Q. Mao, F. Matera, and C. G. Someda, "Concatenation of polarization dispersion in single-mode fibres," *Electronics Letters*, vol. 14, no. 4, pp. 290–291, 1989.
5. J. N. Damask, "Methods to construct programmable PMD sources, Part I: Technology and theory," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 997–1005, Apr. 2004.
6. J. N. Damask, P. R. Myers, A. Boschi, and G. J. Simer, "Demonstration of a coherent PMD source," *IEEE Photonics Technology Letters*, vol. 15, no. 11, pp. 1612–1614, Nov. 2003.
7. E. Desurvire, *Erbium-Doped Fiber Amplifiers, Principles and Applications.* Hoboken, New Jersey: Wiley-Interscience, 2002.
8. A. Eyal, D. Kuperman, O. Dimenstein, and M. Tur, "Polarization dependence of the intensity modulation transfer function of an optical system with PMD and PDL," *IEEE Photonics Technology Letters*, vol. 14, no. 11, pp. 1515–1517, Nov. 2002.
9. A. Eyal, W. K. Marshall, M. Tur, and A. Yariv, "Representation of second-order polarization mode dispersion," *Electronics Letters*, vol. 35, no. 19, pp. 1658–1659, 1999.
10. A. Eyal and M. Tur, "A modified poincare sphere technique for the determination of polarization-mode dispersion in the presence of differential gain/loss," in *Tech. Dig., Optical Fiber Communications Conference (OFC'98)*, San Jose, CA, Feb. 1998, paper ThR1, p. 340.
11. R. Feced, S. J. Savory, and A. Hadjifotiou, "Interaction between polarization mode dispersion and polarization-dependent losses in optical communication links," *Journal of the Optical Society of America B*, vol. 20, no. 3, pp. 424–433, Mar. 2003.
12. E. Forestieri and L. Vincetti, "Exact evaluation of the Jones matrix of a fiber in the presence of polarization mode dispersion of any order," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1898–1909, 2001.
13. C. Francia, F. Bruyére, D. Penninckx, and M. Chbat, "PMD second-order effects on pulse propagation in single-mode optical fibers," *IEEE Photonics Technology Letters*, vol. 10, no. 12, pp. 1739–1741, Dec. 1998.
14. C. Francia and D. Penninckx, "Polarization mode dispersion in single-mode optical fibers: Time impulse response," *IEEE Internation Conference on Communications*, vol. 3, no. 6-10, pp. 1731–1735, June 1999.
15. N. Frigo, private communication, 2003.
16. ——, "A generalized geometric representation of coupled mode theory," *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 11, pp. 2131–2140, 1986.

17. N. Gisin, "Statistics of polarization dependent loss," *Optics Communications*, vol. 114, pp. 399–405, Feb. 1995.
18. N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Optics Communications*, vol. 142, pp. 119–125, Oct. 1997.
19. N. Gisin and J. P. Pellaux, "Polarization mode dispersion: Time versus frequency domains," *Optics Communications*, vol. 89, pp. 316–323, May 1992.
20. J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: http://www.pnas.org
21. ——, "PMD fundamentals: Polarization mode dispersion in optical fibers; Appendix B: Relation between PMD vectors t and w," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000, supplemental Appendix. [Online]. Available: http://www.pnas.org
22. H. A. Haus, *Waves and Fields in Optoelectronics.* Englewood Cliffs, New Jersey: Prentice–Hall, 1984.
23. B. L. Heffner, "Single-mode propagation of mutual temporal coherence: Equivalence of time and frequency measurements of polarization-mode dispersion," *Optics Letters*, vol. 19, no. 15, pp. 1104–1106, Aug. 1994.
24. ——, "Optical pulse distortion measurement limitations in linear time invariant systems, and applications to polarization mode dispersion," *Optics Communications*, vol. 115, pp. 45–51, Mar. 1995.
25. ——, "Influence of optical source characteristics on the measurement of polarization-mode dispersion of highly mode-coupled fibers," *Optics Letters*, vol. 21, no. 2, pp. 113–115, Jan. 1996.
26. F. Heismann, "Accurate Jones matrix expansion for all orders of polarization mode dispersion," *Optics Letters*, vol. 28, no. 11, p. 20132015, Nov. 2003.
27. F. Heismann and M. S. Whalen, "Fast automatic polarization control system," *IEEE Photonics Technology Letters*, vol. 4, no. 5, pp. 503–505, May 1992.
28. F. Heismann, "Analysis of a reset-free polarization controller for fast automatic polarization stabilition in fiber-optic transmission systems," *Journal of Ligthwave Technology*, vol. 12, no. 4, pp. 690–699, Apr. 1994.
29. F. Heismann, D. A. Fishman, and D. L. Wilson, "Automatic compensation of first order polarization mode dispersion in a 10 gb/s transmission system," in *European Conference on Optical Communication (ECOC'98)*, vol. 1, Sept. 1998, pp. 529–530.
30. B. Huttner, C. D. Barros, B. Gisin, and N. Gisin, "Polarization-induced pulse spreading in birefringent optical fibers with zero differential group delay," *Optics Letters*, vol. 24, no. 6, pp. 370–372, Mar. 1999.
31. B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 2, pp. 317–329, Mar. 2000.
32. B. Huttner and N. Gisin, "Anomalous pulse spreading in birefringent optical fibers with polarization-dependent loss," *Optics Letters*, vol. 22, pp. 504–507, Apr. 1997.
33. E. Ibragimov, G. Shtengel, and S. Suh, "Statistical correlation between first and second order PMD," *Journal of Lightwave Technology*, vol. 20, no. 4, pp. 586–590, 2002.

34. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-12: Examinations and measurements - Polarization dependence of attenuation of a single-mode fibre optic component: Matrix calculation method*, International Electrotechnical Commission Std. IEC 61 300-3-12, 1997. [Online]. Available: https://www.iec.ch/
35. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-2: Examinations and measurements - Polarization dependence of attenuation in a single-mode fibre optic device*, International Electrotechnical Commission Std. IEC 61 300-3-2, 1999. [Online]. Available: https://www.iec.ch/
36. R. Jones, "A new calculus for the treatment of optical systems, Part II. proof of three general equivalence theorems," *Journal of the Optical Society of America*, vol. 31, no. 7, pp. 493–499, July 1941.
37. I. P. Kaminow, "Polarization in optical fibers," *IEEE Journal of Quantum Electronics*, vol. QE-17, no. 1, pp. 15–22, 1981.
38. M. Karlsson, "Polarization mode dispersion-induced pulse broadening in optical fibers," *Optics Letters*, vol. 23, no. 9, pp. 688–690, May 1998.
39. M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *Journal of Lightwave Technology*, vol. 18, no. 7, pp. 941–951, 2000.
40. H. Kogelnik, L. E. Nelson, and J. P. Gordon, "Emulation and inversion of polarization-mode dispersion," *Journal of Lightwave Technology*, vol. 21, no. 2, pp. 482–495, 2003.
41. H. Kogelnik, L. Nelson, J. P. Gordon, and R. Jopson, "Jones matrix for second-order polarization mode dispersion," *Optics Letters*, vol. 25, no. 1, pp. 19–21, 2000.
42. D. Kuperman, A. Eyal, O. Mor, S. Traister, and M. Tur, "Measurement of the input states of polarization that maximize and minimize the eye opening in the presence of PMD and PDL," *IEEE Photonics Technology Letters*, vol. 15, no. 10, pp. 1425–1427, Oct. 2003.
43. Y. Li and A. Yariv, "Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses," *Journal of the Optical Society of America B*, vol. 17, no. 11, pp. 1821–1827, Nov. 2000.
44. A. Mecozzi and M. Shtaif, "Signal to noise ratio degradation caused by polarization dependent loss and the effect of dynamic gain equalization," *Journal of Lightwave Technology*, 2004, accepted for publication.
45. C. Menyuk, D. Wang, and A. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," *IEEE Photonics Technology Letters*, vol. 9, no. 9, pp. 1247–1249, Sept. 1997.
46. S. M. R. M. Nezam, J. E. McGeehan, and A. E. Willner, "Theoretical and experimental analysis of the dependence of a signals degree of polarization on the optical data spectrum," *Journal of Lightwave Technology*, vol. 22, no. 3, pp. 763–772, Mar. 2004.
47. A. Orlandini and L. Vincetti, "A simple and useful model for Jones matrix to evaluate higher order polarization-mode dispersion effects," *IEEE Photonics Technology Letters*, vol. 13, no. 11, pp. 1176–1178, 2001.
48. ——, "Comparison of the Jones matrix analytical models applied to optical system affected by high-order PMD," *Journal of Lightwave Technology*, vol. 21, no. 6, pp. 1456–1464, 2003.

49. D. Penninckx and V. Morenas, “Jones matrix of polarization mode dispersion,” *Optics Letters*, vol. 24, no. 13, pp. 875–877, July 1999.
50. D. L. Peterson, B. C. Ward, K. B. Rochford, P. J. Leo, and G. Simer, “Polarization mode dispersion compensator field trial and field fiber characterization,” *Optics Express*, vol. 10, no. 14, pp. 614–621, July 2002. [Online]. Available: http://www.opticsexpress.org/
51. C. D. Poole and C. R. Giles, “Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber,” *Optics Letters*, vol. 13, no. 2, pp. 155–157, 1988.
52. C. D. Poole and R. E. Wagner, “Phenomenological approach to polarization mode dispersion in long single-mode fibers,” *Electronics Letters*, vol. 22, no. 19, pp. 1029–1030, 1986.
53. C. D. Poole, J. H. Winters, and J. A. Nagel, “Dynamical equation for polarization dispersion,” *Optics Letters*, vol. 16, no. 6, pp. 372–374, 1991.
54. C. D. Poole and D. L. Favin, “Polarization-mode dispersion measurements based on transmission spectra through a polarizer,” *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, 1994.
55. S. C. Rashleigh, “Origins and control and polarization effects in single-mode fiber,” *Journal of Lightwave Technology*, vol. LT-1, no. 2, pp. 312–331, 1983.
56. S. C. Rashleigh and R. Ulrich, “Polarization mode dispersion in single mode fibers,” *Optics Letters*, vol. 3, no. 2, pp. 60–62, 1978.
57. W. Shieh, “Principal states of polarization for an optical pulse,” *IEEE Photonics Technology Letters*, vol. 11, no. 6, pp. 677–679, June 1999.
58. W. Shieh and H. Kogelnik, “Dynamic eigenstates of polarization,” *IEEE Photonics Technology Letters*, vol. 13, pp. 40–42, 2001.
59. M. Shtaif and A. Mecozzi, “Polarization-dependent loss and its effect on the signal-to-noise ratio in fiber-optic systems,” *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 671–673, Feb. 2004.
60. *Measurement of Polarization Depedent Loss (PDL) of Single-Mode Fiber Optic Components*, Telecommunications Industry Association Std. TIA/EIA-455-157, 2000. [Online]. Available: http://www.tiaonline.org/standards/
61. S.-C. Wang, private communication, 2002, Insensitivity of Fourier phase to first and last section birefringent phase was first identified by Dr. Wang.
62. C. Xie and L. F. Mollenauer, “Performance degradation induced by polarization-dependent loss in optical fiber transmission systems with and without polarization-mode dispersion,” *Journal of Lightwave Technology*, vol. 21, no. 9, pp. 1953–1957, Sept. 2003.
63. C. Xie, L. F. Mollenauer, and L. Moller, “Pulse distortion induced by polarization-mode dispersion and polarization-dependent loss in lightwave transmission systems,” *IEEE Photonics Technology Letters*, vol. 15, no. 8, pp. 1073–1075, Aug. 2003.
64. M. Yoshida-Dierolf and V. Dierolf, “Analytical form of frequency dependence of DGD in concatenated single-mode fiber systems,” *Journal of Lightwave Technology*, vol. 21, no. 10, pp. 2217–2223, Oct. 2003.