Journal of the Optical Society of Korea
Vol. 12, No. 4, December 2008, pp. 240-243

DOI: 10.3807/JOSK.2008.12.4.240

Derivation of the Foschini and Shepp's Joint-Characteristic Function for the First-and Second-Order Polarization-Mode-Dispersion Vectors Using the Fokker-Planck Method

Jae-Seung Lee*

Department of Electronics, Kwangwoon University, Kwangwoon-Gil 26, Nowon-Gu, Seoul, 139-701, Korea

(Received October 14, 2008 : revised November 18, 2008 : accepted December 1, 2008)

Using the well-known Fokker-Planck method, this paper presents a standard way to find the joint-characteristic function for the first- and second-order polarization-mode-dispersion vectors originally derived by Foschini and Shepp. Compared with the Foschini and Shepp's approach, the Fokker-Planck approach gives a more accurate and straightforward way to find the joint-characteristic function.

Keywords: Polarization Mode Dispersion, Fokker-Planck Equation

OCIS codes : (060.2330) Fiber optics communications; (060.2270) Fiber characterization; (060.2300) Fiber measurements

I. INTRODUCTION

Polarization-mode dispersion (PMD) is caused by small random birefringences within optical fibers and is an important stochastic parameter especially for high bit-rate channels. For a fixed optical frequency component, its output polarization state after an optical fiber is described by a PMD vector, also called the first-order PMD vector. The magnitude of the first-order PMD vector is equal to the differential group delay between two principal states [1]. When there are many optical channels or the channel bit rate is high, the angular frequency derivative of the PMD vector, called the second-order PMD vector, becomes also important.

In 1991, Foschini and Shepp derived a joint- characteristic function using a sine-cosine Fourier Transform representation for white Gaussian processes [2]. In the companion paper [3], it was explained in detail that the joint-characteristic function can be applied to the first- and second-order PMD vectors for highly- birefringent fibers in long distance limit. This function has played a key role in subsequent papers to describe various probability distribution characteristics of PMD vectors [4-6]. However, the original procedure to find the joint-characteristic function is lengthy and complicated. An alternative way to find the joint-characteristic function was reported by Gordon [7]. He used the Lax procedure [8] and assumed a spherically-symmetric birefringence vector distribution in Stokes space. In this paper, we use a more familiar Fokker-Plank method [9] to find the joint-characteristic function more correctly. The birefringence vector distribution is chosen to be the same as [7].

*Corresponding author: jslee@kw.ac.kr

II. FIRST-ORDER PMD VECTOR

As a demonstration of the Fokker-Planck method, we first find the probability-density function (pdf) of the first-order PMD vector. The dynamical PMD equation,

$$\frac{\partial \boldsymbol{\tau}}{\partial L} = \boldsymbol{\beta} \times \boldsymbol{\tau} + \boldsymbol{\beta}_\omega \tag{1}$$

describes the evolution of the first-order PMD vector, $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$, along the fiber [1]. L is the transmission distance. $\boldsymbol{\beta}$ is the birefringence vector given as $\boldsymbol{\beta} = a(\Gamma_1,$

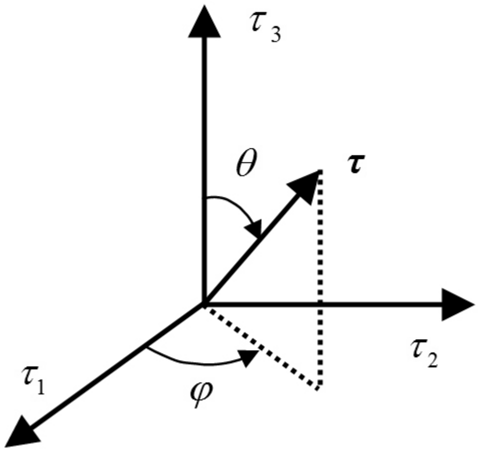

FIG. 1. Spherical coordinate for τ space.

$\Gamma_2,\Gamma_3)$. $\Gamma_i=\Gamma_i(L)(i=1,2,3)$ is a zero-mean white Gaussian process having a spherically-symmetric correlation property, $E\{\Gamma_i(L)\Gamma_j(L')\}=2\delta_{ij}\delta(L-L')$, where E{ · } is the ensemble average and δ_{ij} is the Kronecker delta. The proportional constant, a, is dependent on the angular frequency, ω. The fiber correlation length is neglected because it is much smaller than the scale of L[7]. We express $\partial\beta/\partial\omega$ and $\partial\beta^2/\partial\omega^2$ as $\boldsymbol{\beta}_\omega=a_\omega(\Gamma_1,\ \Gamma_2,\ \Gamma_3)$ and $\boldsymbol{\beta}_{\omega\omega}=a_{\omega\omega}(\Gamma_1,\ \Gamma_2,\ \Gamma_3)$, respectively.

We note that the vector equation (1) is composed of three Langevin equations, $\partial x_i/\partial L=\sum_{j=1}^{N} g_{ij}\Gamma_j$ (i=1, 2, 3), where $x_i=\tau_i$ and N=3. Thus we find the Fokker-Planck equation for (1) according to [9] as

$$\frac{\partial P}{\partial L}=-\sum_{i=1}^{N}\frac{\partial}{\partial x_i}D_iP+\sum_{i,j=1}^{N}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}D_{ij}P \tag{2}$$

where D_i and D_{ij} are drift and diffusion coefficients, respectively, defined as $D_i=\sum_{k,j=1}^{N} g_{kj}\partial g_{ij}/\partial x_k$ and $D_{ij}=\sum_{k=1}^{N} g_{ik}g_{jk}$. $P=P(\tau,L)$ is the pdf for $\boldsymbol{\tau}$. Instead of the rectangular coordinate, we use a spherical coordinate to find the following Fokker-Planck equation:

As is shown in Fig. 1, we use a spherical coordinate using following transform relations:

$$\frac{\partial}{\partial\tau_1}=\sin\theta\cos\varphi\frac{\partial}{\partial\tau}+\cos\theta\cos\varphi\frac{\partial}{\tau\partial\theta}-\frac{\sin\varphi}{\tau\sin\theta}\frac{\partial}{\partial\varphi} \tag{3}$$

$$\frac{\partial}{\partial\tau_2}=\sin\theta\sin\varphi\frac{\partial}{\partial\tau}+\cos\theta\sin\varphi\frac{\partial}{\tau\partial\theta}+\frac{\cos\varphi}{\tau\sin\theta}\frac{\partial}{\partial\varphi} \tag{4}$$

$$\frac{\partial}{\partial\tau_3}=\cos\theta\frac{\partial}{\partial\tau}-\sin\theta\frac{\partial}{\tau\partial\theta} \tag{5}$$

The positive τ_3 -axis is the polar axis with polar and azimuth angles θ and φ, respectively. Since our problem is invariant for the rotation about the τ_3-axis, we may set $\partial/\partial\varphi=0$ and find

$$\frac{\partial P(\boldsymbol{\tau},L)}{\partial L}=\left(a_\omega^2\nabla^2+a^2\tau^2\nabla_{angle}^2\right)P(\boldsymbol{\tau},L) \tag{6}$$

$$\tau^2\nabla_{angle}^2=\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{7}$$

The subscript, *angle*, implies the angular part of the Laplacian. Since the input signal has experienced no PMD degradations, the initial condition is $P(\tau,0)=\delta(\tau)$ which is spherically symmetric. Thus (6) becomes a 3-dimensional diffusion equation, $\partial P(\tau,L)/\partial L=a_\omega^2\nabla^2P(\tau,L)$, and we obtain a Gaussian pdf, $P(\tau,L)=\exp(-\tau^2/2\sigma^2)/(\sqrt{2\pi}\,\sigma)^3$, where $\sigma^2(=2a_\omega^2L)$ is the variance of each rectangular component of τ. Accordingly, the magnitude of τ has the well-known Maxwellian pdf, $\sqrt{2/\pi}\,(\tau^2/2\sigma^3)\exp(-\tau^2/2\sigma^2)$.

III. JOINT CHARACTERISTIC FUNCTION FOR THE FIRST- AND SECOND-ORDER PMD VECTORS

From (1), we find a differential equation for the second-order PMD vector, $\boldsymbol{\tau}_\omega=\partial\tau/\partial\omega$, as follows:

$$\frac{\partial\boldsymbol{\tau}_\omega}{\partial L}=\boldsymbol{\beta}_\omega\times\boldsymbol{\tau}+\boldsymbol{\beta}\times\boldsymbol{\tau}_\omega+\boldsymbol{\beta}_{\omega\omega} \tag{8}$$

We note that (1) and (8) can be regarded as Langevin equations with N=6, $\partial x_i/\partial L=\sum_{j=1}^{N} g_{ij}\Gamma_j$, where $x_i=\tau_i$ for $i=1,\ 2,\ 3$ and $x_i=\tau_{\omega i}$ for $i=4,\ 5,\ 6$. According to (2), we find a Fokker-Planck equation as follows:

$$\begin{aligned}\frac{\partial P}{\partial L}=&\Big[a_\omega^2\nabla_1^2+a_{\omega\omega}^2\nabla_2^2+2a_\omega a_{\omega\omega}\nabla_1\cdot\nabla_2+a_\omega^2\left\{\tau^2\nabla_2^2-(\boldsymbol{\tau}\cdot\nabla_2)^2\right\}\\&+a^2\left\{\tau^2\nabla_{1,angle}^2+\tau_\omega^2\nabla_{2,angle}^2+2(\boldsymbol{\tau}\cdot\boldsymbol{\tau}_\omega)\nabla_1\cdot\nabla_2-2(\boldsymbol{\tau}\cdot\nabla_2)(\boldsymbol{\tau}_\omega\cdot\nabla_1)+2\boldsymbol{\tau}\cdot\nabla_1\right\}\\&+2aa_\omega\left\{(\boldsymbol{\tau}\cdot\boldsymbol{\tau}_\omega)\nabla_2^2-(\boldsymbol{\tau}_\omega\cdot\nabla_2)(\boldsymbol{\tau}\cdot\nabla_2)+\boldsymbol{\tau}^2(\nabla_1\cdot\nabla_2)-(\boldsymbol{\tau}\cdot\nabla_2)(\boldsymbol{\tau}\cdot\nabla_1)\right.\\&\left.-2\boldsymbol{\tau}\cdot\nabla_2-\boldsymbol{\tau}_\omega\cdot(\nabla_1\times\nabla_2)\right\}+2\left(aa_{\omega\omega}-a_\omega^2\right)\boldsymbol{\tau}\cdot(\nabla_1\times\nabla_2)\Big]P\end{aligned} \tag{9}$$

where ∇_1 and ∇_2 are del operators in τ and τ_ω spaces, respectively. $P=P(\tau,\tau_\omega,L)$ is the joint pdf of τ and τ_ω. The initial condition is $P(\tau,\tau_w,0)=\delta(\tau)\delta(\tau_w)$. With this initial condition, $P(\tau,\tau_\omega,L)$ should be dependent only on the magnitudes of τ and τ_ω and the angle between them. Thus the differential operators in (9) become

$$\boldsymbol{\tau}\cdot\nabla_1=\tau\frac{\partial}{\partial\tau} \tag{10}$$

$$\boldsymbol{\tau}_\omega\cdot\nabla_2=\tau_\omega\frac{\partial}{\partial\tau_\omega} \tag{11}$$

$$\boldsymbol{\tau}\cdot\nabla_2 = \tau\left(\cos\gamma\frac{\partial}{\partial\tau_\omega} - \frac{\sin\gamma}{\tau_\omega}\frac{\partial}{\partial\gamma}\right) \tag{12}$$

$$\boldsymbol{\tau}_\omega\cdot\nabla_1 = \tau_\omega\left(\cos\gamma\frac{\partial}{\partial\tau} - \frac{\sin\gamma}{\tau}\frac{\partial}{\partial\gamma}\right) \tag{13}$$

$$\nabla_1\cdot\nabla_2 = \cos\gamma\frac{\partial^2}{\partial\tau_\omega\partial\tau} - \sin\gamma\left(\frac{\partial}{\tau_\omega\partial\tau} + \frac{\partial}{\tau\partial\tau_\omega}\right)\frac{\partial}{\partial\gamma} - \frac{1}{\tau_\omega\tau}\left(\frac{\partial}{\sin\gamma\partial\gamma} + \cos\gamma\frac{\partial^2}{\partial\gamma^2}\right) \tag{14}$$

$$\nabla_2^2 = \frac{\partial^2}{\partial\tau_\omega^2} + \frac{2}{\tau_\omega}\frac{\partial}{\partial\tau_\omega} + \frac{1}{\tau_\omega^2\sin\gamma}\frac{\partial}{\partial\gamma}\sin\gamma\frac{\partial}{\partial\gamma} \tag{15}$$

$$\tau^2\nabla^2_{1,angle} = \tau_\omega^2\nabla^2_{2,angle} = \frac{1}{\sin\gamma}\frac{\partial}{\partial\gamma}\sin\gamma\frac{\partial}{\partial\gamma} \tag{16}$$

$$\boldsymbol{\tau}_\omega\cdot(\nabla_1\times\nabla_2) = \boldsymbol{\tau}\cdot(\nabla_1\times\nabla_2) = 0 \tag{17}$$

where γ is the angle between $\boldsymbol{\tau}$ and $\boldsymbol{\tau}_\omega$. Then, we find that a^2, aa_ω, and $(aa_\omega - a_\omega^2)$ terms in (9) disappear. The Fokker-Planck equation becomes

$$\frac{\partial P}{\partial L} = \left[a_\omega^2\nabla_1^2 + a_{\omega\omega}^2\nabla_2^2 + 2a_\omega a_{\omega\omega}\nabla_1\cdot\nabla_2 + a_\omega^2\left\{\tau^2\nabla_2^2 - (\boldsymbol{\tau}\cdot\nabla_2)^2\right\}\right]P. \tag{18}$$

Introducing the joint-characteristic function for the joint pdf as a Fourier integral

$$Q(\boldsymbol{k},\boldsymbol{k}_\omega,L) = \iint d\boldsymbol{\tau} d\boldsymbol{\tau}_\omega \exp(-j\boldsymbol{k}\cdot\boldsymbol{\tau} - j\boldsymbol{k}_\omega\cdot\boldsymbol{\tau}_\omega)P(\boldsymbol{\tau},\boldsymbol{\tau}_\omega,L) \tag{19}$$

we find

$$\frac{\partial Q}{\partial L} = \left[-a_\omega^2k^2 - a_{\omega\omega}^2k_\omega^2 - 2a_\omega a_{\omega\omega}\boldsymbol{k}\cdot\boldsymbol{k}_\omega + a_{\omega\omega}^2\left\{k_\omega^2\nabla_k^2 - (\boldsymbol{k}_\omega\cdot\nabla_k)^2\right\}\right]Q \tag{20}$$

where ∇_k is the del operator in $\boldsymbol{k}$ vector space and $Q(\boldsymbol{k},\boldsymbol{k}_\omega,0) = 1$. Note that our definition of the joint-characteristic function (19) is a complex conjugate of the conventional expression. The $\boldsymbol{k}_\omega$ vector is just a parameter in (20). Thus we set $\boldsymbol{k}_\omega$ to $(0,0,k_\omega)$ for a moment and decompose $Q(\boldsymbol{k},\boldsymbol{k}_\omega,0)$ as

$$Q = \exp\left\{-(a_\omega k_3 + a_{\omega\omega}k_\omega)^2L\right\}q(k_1,k_2,L) \tag{21}$$

where $\boldsymbol{k}=(k_1,k_2,k_3)$ and q satisfies

$$\frac{\partial q}{\partial L} = \left\{-a_\omega^2(k_1^2+k_2^2) + a_\omega^2k_\omega^2\left(\frac{\partial^2}{\partial k_1^2} + \frac{\partial^2}{\partial k_2^2}\right)\right\} \tag{22}$$

Eq. (22) can be solved using the separation-of-variables method as

$$q(k_1,k_2,L) = \frac{\exp\left\{-\frac{k_1^2+k_2^2}{2k_\omega}\tanh(2a_\omega^2k_\omega L)\right\}}{\cosh(2a_\omega^2k_\omega L)} \tag{23}$$

Consequently, for an arbitrary direction of k_ω, the joint-characteristic function becomes

$$Q(\boldsymbol{k},\boldsymbol{k}_\omega,L) = \frac{\exp\left\{-\frac{k_\perp^2}{2k_\omega}\tanh(2a_\omega^2k_\omega L) - (a_\omega k_{//} + a_{\omega\omega}k_\omega)^2L\right\}}{\cosh(2a_\omega^2k_\omega L)} \tag{24}$$

where $k_{//}$ is the component of $\boldsymbol{k}$ parallel to $\boldsymbol{k}_\omega$ and $k_\perp^2 = k^2 - k_{//}^2$. Our result is more exact than the previous joint-characteristic function. When $a_{\omega\omega}$ is neglected, we find that (24) is the same as the joint-characteristic function deived in [2] and [7] using the relation $2a_\omega^2L = \sigma^2 = E\{\tau^2\}/3$.

The $a_{\omega\omega}$ term is not negligible when the transmission distance is not so large. To be more specific, we note that

$$Q(0,\boldsymbol{k}_\omega,L) = \frac{\exp(-a_{\omega\omega}^2k_w^2L)}{\cosh(2a_\omega^2k_\omega L)} \tag{25}$$

$$= \frac{\exp\left(-\frac{a_{\omega\omega}^2}{4a_\omega^4L}X^2\right)}{\cosh X} \tag{26}$$

where $X=2a_\omega^2k_\omega L$. The 3-dimensional inverse Fourier transform of $Q(0,\boldsymbol{k}_\omega,L)$ yields the pdf of $\boldsymbol{\tau}_\omega$. Although $Q(0,\boldsymbol{k}_\omega,L)$ becomes $1/\cosh(2a_\omega^2k_\omega L)$ in the limit of large L, the correction term $\exp(-a_{\omega\omega}^2k_w^2L)$ is important for reasonable values of L. This is illustrated in Fig. 2 where we have used some typical values such as a_ω=0.01735 $ps/\sqrt{km}$ and $a_{\omega\omega}$=0.002332 $ps^2/\sqrt{km}$. $Q(0,k_\omega,L)$ converges to $1/\cosh(2a_\omega^2k_\omega L)$ very slowly and our result is more accurate for moderate transmission distances.

Actually, (18) has no a terms. Thus we may set $\beta=0$ in (1) and (8). If we neglect $a_{\omega\omega}$ terms further as [7] does, we have $\partial\boldsymbol{\tau}/\partial L = \boldsymbol{\beta}_\omega$ and $\partial\boldsymbol{\tau}_\omega/\partial L = (\partial\boldsymbol{\tau}/\partial L)\times\boldsymbol{\tau}$. Thus we find

$$\boldsymbol{\tau} = \int_0^L \boldsymbol{\beta}_\omega(l)dl \tag{27}$$

$$\boldsymbol{\tau}_\omega = \int_0^L dl\boldsymbol{\beta}_\omega(l)\times\int_0^l \boldsymbol{\beta}_\omega(l')dl' \tag{28}$$

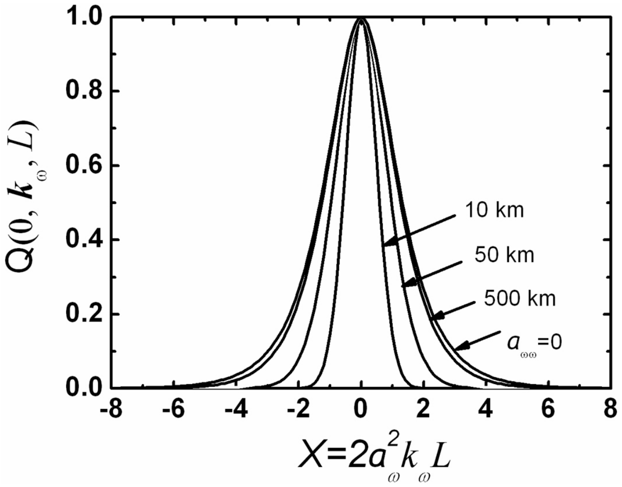

FIG. 2. Plot of $Q(0,\boldsymbol{k}_\omega,L)$ for transmission lengths, 10, 50, 500 km. a_ω=0.01735 $ps/\sqrt{km}$ and $a_{\omega\omega}$=0.002332 $ps^2/\sqrt{km}$. The a_ω=0 line corresponds to the result of Foschini and Shepp (infinite L).

These relations are equivalent to the starting point of Foschini and Shepp, (4) of [2]. They expand β_ω as a sine-cosine Fourier series with Gaussian random coefficients. The fact that these coefficients are mutually independent gives a closed form of the joint-characteristic function for τ and τ_ω.

Using the relation, $E\{\tau^2\}=-\nabla_k^2 Q(\boldsymbol{k},\boldsymbol{k}_\omega,L)|_{\boldsymbol{k}=\boldsymbol{k}_\omega=0}$, we find $E\{\tau^2\}=6a_\omega^2 L$. In a similar way, $E\{\tau_\omega^2\}$ is found as $12a_\omega^4 L^2+6a_{\omega\omega}^2 L$. The contributions from $a_{\omega\omega}^2$ term becomes negligible as L increases very large and the ratio $E\{\tau_\omega^2\}/(E\{\tau^2\})^2$ converges to 3 for large L as [3].

IV. CONCLUSION

We have found a more general expression of the joint-characteristic function for the first- and second-order PMD vectors using the standard Fokker-Planck method. Our Fokker-Planck method includes the second-order derivative term, $d^2a/d\omega^2=a_{\omega\omega}$, while the Foschini and Shepp's approach neglects it. We have also clarified that our method is equivalent to the Foschini and Shepp's approach in the limit of large transmission distance.

ACKNOWLEDGMENT

This work was partly supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST: R01-2007-000-21055-0) and also partly supported by the Kwangwoon University under a 2007 research grant.

REFERENCES

[1] C. D. Poole and J. Nagel, "Polarization effects in lightwave systems," *Optical Fiber Telecommunications III-A*, I. P. Kaminow and T. L. Koch, Ed. San Diego, CA: Academic, ch. 6, 1997.
[2] G. J. Foschini and L. A. Shepp, *Stochastic Analysis*, (New York, Academic Press, 1991), pp. 169–187.
[3] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," *J. Lightwave Tech.*, vol. 9, no. 11, pp. 1439-1456, Nov. 1991.
[4] G. J. Foschini, R. M. Jopson, L. E. Nelson, and H. Kogelnik, "The statistics of PMD-induced chromatic fiber dispersion," *J. Lightwave Tech.*, vol. 17, no. 9, pp. 1560–1565, Sept. 1999.
[5] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Probability densities of second-order polarization mode dispersion including polarization dependent chromatic fiber dispersion," *IEEE Photon. Technol. Lett.*, vol. 12, no. 3, pp. 293–295, Mar. 2000.
[6] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," *J. Lightwave Tech.*, vol. 19, no. 12, 1882–11886, Dec. 2001.
[7] J. P. Gordon, "Statistical properties of polarization mode dispersion," in *Polarization mode dispersion*, A. Galtarossa and C. R. Menyuk, eds., New York: Springer, pp. 52-59, 2005.
[8] M. Lax, "Classical Noise IV: Langevin Methods," *Rev. Mod. Phys.*, vol. 38, pp.541-566, 1966.
[9] H. Risken, *The Fokker-Planck Equation Methods of Solution and Applications*, 2nd ed. (New York: Springer-Verlag, 1996), ch. 3, pp. 54–56.