Optical Fiber and the Fiber Channel

2

Cristian Antonelli

Contents

Abstract

The enormous potential of the fiber-optic channel to transmit data over long distances at high rates has been gradually unlocked by means of a number of key technological innovations underpinned by the mature understanding of lightwave propagation in optical fibers. This chapter reviews the main properties of the fiber-optic channel, starting from the structure of ideal linear optical fibers and proceeding to the derivation of the equations governing signal propagation in

C. Antonelli (✉)
Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila, Italy

National Laboratory of Advanced Optical Fibers for Photonics (FIBERS), CNIT, L'Aquila, Italy
e-mail: cristian.antonelli@univaq.it

T. Kawanishi (ed.), *Handbook of Radio and Optical Networks Convergence*,
https://doi.org/10.1007/978-981-97-2282-2_17

single-mode fibers in the presence of various physical phenomena. These are chromatic dispersion, attenuation and amplification, polarization-mode dispersion, polarization-dependent loss, and nonlinear distortions. In the context of nonlinear propagation, the chapter also provides an introduction to the modeling of the nonlinear interference noise in coherent systems that are used today. The chapter targets readers who are familiar with the fundamental concepts involved in the study of electromagnetic fields and signal theory but have no specific background on fiber-optic transmission. It consists of five sections that are largely self-contained but with references used to avoid long derivations and maintain brevity.

Keywords

Optical fiber · Fiber modes · Propagation effects · Polarization-mode dispersion · Polarization-dependent loss · Nonlinear interference noise

Introduction

The enormous potential of the fiber-optic channel to transmit data over long distances at high rates has been gradually unlocked by means of a number of key technological innovations underpinned by the mature understanding of lightwave propagation in optical fibers. This chapter reviews the main properties of the fiber-optic channel, starting from the structure of ideal linear optical fibers and proceeding to the derivation of the equations governing signal propagation in single-mode fibers in the presence of various physical phenomena.

The chapter is organized into five sections. The first section "Fiber Modes" introduces the formalism used throughout the chapter and sets up the fiber-mode problem. Its solution is discussed in the case of step-index fibers and in particular in the relevant regime of weak guidance. The subsequent sections are devoted to deriving the evolution equation for information carrying signals within the perturbation approach reviewed in the second section "Transmission of Information-Carrying Signals." The third section "Chromatic Dispersion" discusses the modeling of chromatic dispersion, the signal spreading that it implies, and its compensation. The fundamental concepts related to in-line amplification are introduced in the section "Loss, Gain, and Noise." The section "Polarization Effects" is devoted to the review of random polarization-mode coupling, polarization-mode dispersion (PMD), and polarization-dependent loss (PDL), while introducing the Jones and Stokes representation of lightwave signals. Finally, the section "Nonlinear Propagation" reviews the main nonlinear phenomena that characterize signal propagation in the fiber-optic channel, while the section "Nonlinear Interference Noise in the Context of Digital Coherent Systems" introduces the modeling of the nonlinear interference noise in uncompensated coherent systems of the kind that are used today.

The chapter is largely self-contained, and pointers to important references are provided to avoid detailed derivations and focus on the key concepts.

Fiber Modes

The optical fiber is a circularly symmetric dielectric waveguide. The simplest fiber design is the one known as *step index* and is illustrated in Fig. 1. This is characterized by a stepwise dependence of the refractive index on the radial coordinate, where the fiber core has a slightly larger refractive index compared to the surrounding cladding. The fiber's ability to guide lightwaves can be understood intuitively with reference to the effect of total internal reflection in the framework of ray optics. In the case of an ideally straight fiber illustrated in the same figure, this is achieved if Snell's refraction law at the core-cladding interface, $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, admits no real solution for θ_2, that is, if $\theta_1 \geq \sin^{-1}(n_2/n_1)$. Therefore, light rays impinging upon the fiber input with an angle θ, for which $n_0 \sin(\theta) = n_1 \sin(\pi/2 - \theta_1)$, with $n_0 = 1$, are guided if $\theta \leq \theta_0 = \sqrt{n_1^2 - n_2^2}$. For small values of $n_1 - n_2$, this quantity is a good approximation of the fiber numerical aperture, which like for any other optical system, is defined as the sine of the acceptance-cone angle, $\mathrm{NA} = \sin(\theta_0)$. A rigorous study of propagation in optical fibers, however, involves the concept of fiber modes, and in what follows we summarize the main aspects of this approach.

Our starting point is Maxwell's equations in a dielectric medium,

$$\nabla \times \vec{e} = -\mu_0 \frac{\partial \vec{h}}{\partial t}, \tag{1}$$

$$\nabla \times \vec{h} = \frac{\partial \vec{d}}{\partial t}, \tag{2}$$

$$\nabla \cdot \vec{d} = 0, \tag{3}$$

$$\nabla \cdot \vec{h} = 0, \tag{4}$$

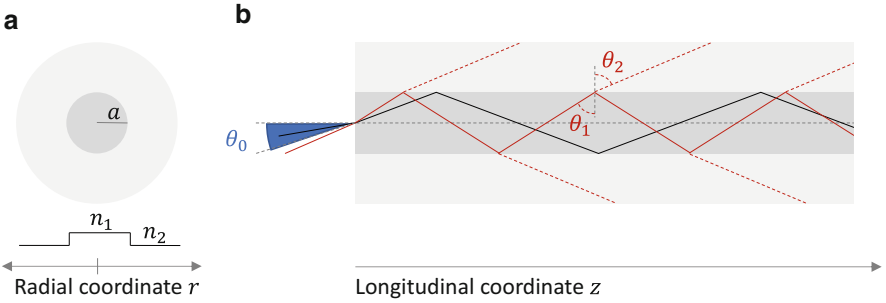

Fig. 1 (**a**) Step-index refractive index profile, where the core index n_1 is slightly larger than the cladding index n_2. (**b**) Illustration of total internal reflection. Incident light rays within the acceptance angle θ_0 undergo total reflection at the core-cladding interface (black lines). The others are partially dispersed in the cladding at each reflection (red lines)

where $\vec{e}$ and $\vec{h}$ are the electric and magnetic fields, respectively, and where $\vec{d} = \varepsilon_0 \vec{e} + \vec{p}$ is the electric displacement vector, with $\vec{p}$ denoting the polarization density vector. In the reference case of an ideally straight fiber with instantaneous material response, $\vec{p} = \varepsilon_0 \chi \vec{e}$, where χ is the dielectric susceptibility, and $\vec{d} = \varepsilon_0 \varepsilon_r \vec{e}$, with $\varepsilon_r = \chi + 1$ being the relative dielectric constant. The refractive index n is related to ε_r through $\varepsilon_r = n^2$. Defining the Fourier transform pair as

$$\widetilde{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{i\omega t} \mathrm{d}t, \quad f(t) = \int_{-\infty}^{+\infty} \widetilde{f}(\omega) e^{-i\omega t} \frac{\mathrm{d}\omega}{2\pi}, \tag{5}$$

in the frequency domain, allows Maxwell's equations to take the form

$$\nabla \times \widetilde{\vec{e}} = i\omega\mu_0 \widetilde{\vec{h}}, \tag{6}$$

$$\nabla \times \widetilde{\vec{h}} = -i\omega \widetilde{\vec{d}}, \tag{7}$$

$$\nabla \cdot \widetilde{\vec{d}} = 0, \tag{8}$$

$$\nabla \cdot \widetilde{\vec{h}} = 0, \tag{9}$$

with $\widetilde{\vec{d}} = \varepsilon_0 \varepsilon_r \widetilde{\vec{e}}$ in the case of the ideal straight fiber.

The study of the fiber modes is conveniently carried out by considering a monochromatic excitation of a single fiber mode, in which case the propagating field is a solution of Eqs. (1)–(4) of the following form

$$\vec{e}(x, y, z, t) = \mathrm{Re}\left\{ A \vec{F}(x, y) e^{-i(\omega t - \beta z)} \right\}, \tag{10}$$

$$\vec{h}(x, y, z, t) = \mathrm{Re}\left\{ A \vec{G}(x, y) e^{-i(\omega t - \beta z)} \right\}, \tag{11}$$

where z is the fiber longitudinal coordinate, as illustrated in Fig. 1. The complex-valued three-dimensional vectors $\vec{F}(x, y)$ and $\vec{G}(x, y)$ describe the lateral profiles of the electric and magnetic modes, respectively, which must be invariant during propagation, whereas the scalar coefficient A represents the mode excitation at the fiber input. The term $\omega t - \beta z$ describes the phase of the field as it evolves along the fiber, where β is the mode propagation constant. The z-components of the electric and magnetic mode's lateral profiles play a special role in the mode analysis, as all the other components can be expressed in terms of them. To see this, we first express $\vec{F} = \vec{F}_t + F_z \hat{z}$ and $\vec{G} = \vec{G}_t + G_z \hat{z}$, where the subscript t refers to the vector content across the fiber cross section (that is, in the (x, y) plane), and similarly $\nabla = \nabla_t + \hat{z} \frac{\partial}{\partial z}$ (the hat denotes unit-length vectors). We then insert $\widetilde{\vec{e}} = \vec{F} \exp(i\beta z)$ and $\widetilde{\vec{h}} = \vec{G} \exp(i\beta z)$ into Eqs. (6) and (7), and multiply both sides of the two equations by

$\hat{z}\times$, with the result $\nabla_t F_z - i\beta\vec{F}_t = i\omega\mu_0\hat{z}\times\vec{G}_t$ and $\nabla_t G_z - i\beta\vec{G}_t = -i\omega\varepsilon_0\varepsilon_r\hat{z}\times\vec{F}_t$. From these equations, one can readily obtain

$$\vec{F}_t = \frac{i\beta}{\varepsilon_r k^2 - \beta^2}\left(\nabla_t F_z - \frac{\omega\mu_0}{\beta}\hat{z}\times\nabla_t G_z\right) \tag{12}$$

$$\vec{G}_t = \frac{i\beta}{\varepsilon_r k^2 - \beta^2}\left(\nabla_t G_z - \frac{\omega\varepsilon_0\varepsilon_r}{\beta}\hat{z}\times\nabla_t F_z\right), \tag{13}$$

where $k = \omega/c$ ($c = 1/\sqrt{\varepsilon_0\mu_0}$ is the speed of light in vacuum) is the propagation constant of a plane wave in vacuum.

The equation for F_z is derived by applying the operator $\nabla\times$ to both sides of Eq. (6) and by taking $\nabla\times\vec{\tilde{h}}$ from Eq. (7), with the result

$$\nabla^2\vec{\tilde{e}} + n^2k^2\vec{\tilde{e}} = \nabla\left(\nabla\cdot\vec{\tilde{e}}\right) \simeq 0, \tag{14}$$

in whose derivation the identity $\nabla\times\nabla\times\vec{f} = \nabla\left(\nabla\cdot\vec{f}\right) - \nabla^2\vec{f}$ is used. Typically $\nabla\cdot\vec{\tilde{e}}$ is set to zero, although from $\nabla\cdot\vec{\tilde{d}} = 0$ one can only conclude that $\nabla\cdot\vec{\tilde{e}} = -2\vec{\tilde{e}}\cdot(\nabla n)/n$. Therefore setting $\nabla\cdot\vec{\tilde{e}}$ to zero is rigorously correct in the case of a step-index fiber (where n is uniform in the core and in the cladding), while in general, this is a valid approximation for fibers with a smoothly varying refractive index profile, such that $|\nabla n| \ll n$. For the magnetic-field mode, Eq. (14) becomes

$$\nabla^2\vec{\tilde{h}} + k^2\vec{\tilde{h}} = 0, \tag{15}$$

regardless of the refractive index profile (Eqs. (14) and (15) are in the form of the Helmholtz equation). By entering the modes' ansatz in Eqs. (14) and (15), one finds

$$\nabla_t^2\vec{F} + \left(n^2k^2 - \beta^2\right)\vec{F} = 0, \tag{16}$$

$$\nabla_t^2\vec{G} + \left(n^2k^2 - \beta^2\right)\vec{G} = 0. \tag{17}$$

The component F_z is the solution of Eq. (16) with $\vec{F}$ replaced with F_z, and, given the circular symmetry of the optical fiber, it is natural to search for it in cylindrical coordinates, with $F_z = F_z(r,\ \phi)$. Separation of variables $F_z(r,\ \phi) = \psi(r)\,f(\phi)$ yields

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}\psi}{\mathrm{d}r} + \left(n^2k^2 - \beta^2 - \frac{\ell^2}{r^2}\right)\psi = 0, \tag{18}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}\phi^2} + \ell^2 f = 0, \tag{19}$$

where ℓ is a constant introduced in the process of separating the variables. Since the solution of Eq. (19) is in the form of $f = \exp(\pm i\ell\phi)$ (or, equivalently, in the form of $f = \cos(\ell\phi)$ or $f = \sin(\ell\phi)$), and obviously must be 2π-periodic, then ℓ must be an integer. At this point, it becomes necessary to specify the refractive index profile in order to extract ψ. In what follows, we discuss the case of the step-index fiber and introduce the approximation of linearly polarized modes in the relevant case of weakly guiding fibers [1].

Step-Index Fiber

As illustrated in Fig. 1, for a step-index profile, n is equal to n_1 in the core and to n_2 in the cladding ($n_2 < n_1$), which for the purpose of the mode analysis is assumed to extend to indefinitely large radial distances. In this case, Eq. (18) becomes a Bessel equation and its solution depends on the sign of $n^2k^2 - \beta^2$. In particular, the Bessel function of the first kind J_ℓ is a suitable solution for the core, as is does not have singularities in $r = 0$, while the modified Bessel function of the second kind K_ℓ is a suitable solution for the cladding, as it does not diverge for $r \to \infty$. This is the case if $\kappa_1^2 = n_1^2k^2 - \beta^2 > 0$ and $\kappa_2^2 = \beta^2 - n_2^2k^2 > 0$, in which case

$$\psi(r) = J_\ell(\kappa_1 r), r < a \tag{20}$$

$$\psi(r) = K_\ell(\kappa_2 r), r > a. \tag{21}$$

The requirement $\kappa_{1,2}^2 \geq 0$ implies the condition $n_1k \leq \beta \leq n_2k$, showing that the propagation constant of the mode is bounded between those of plane waves propagating in homogeneous media with refractive indexes n_2 and n_1. It is therefore customary to express the propagation constant as $\beta = n_{\mathrm{eff}}k$, where n_{eff} is referred to as the effective refractive index, and $n_2 \leq n_{\mathrm{eff}} \leq n_1$. The extraction of the propagation constant follows from imposing the boundary conditions. To this end, one can express F_z and G_z as $F_z = f_a \frac{J_\ell(\kappa_1 r)}{J_\ell(\kappa_1 a)} \cos(\ell\phi)$ and $G_z = g_a \frac{J_\ell(\kappa_1 r)}{J_\ell(\kappa_1 a)} \sin(\ell\phi)$ in the core, and $F_z = f_a \frac{K_\ell(\kappa_2 r)}{K_\ell(\kappa_2 a)} \cos(\ell\phi)$ and $G_z = g_a \frac{K_\ell(\kappa_2 r)}{K_\ell(\kappa_2 a)} \sin(\ell\phi)$ in the cladding, and evaluate $\vec{F}_t$ and $\vec{G}_t$ using Eqs. (16) and (17). It can be readily seen that by doing so, $\vec{F}_t$ and $\vec{G}_t$ assume the form of linear combinations of the two constants f_a and g_a. Imposing the continuity of the tangential components $\vec{F}_t \cdot \hat{\phi}$ and $\vec{G}_t \cdot \hat{\phi}$ at the core-cladding interface (the continuity of F_z and G_z is guaranteed by construction) yields a 2×2 homogeneous system of linear equations for f_a and g_a. Setting the determinant of the coefficients matrix to zero yields, after some tedious yet straightforward algebra, the famous characteristic equation [2].

$$\ell^2\left(\frac{\beta}{k}\right)^2\left[\left(\frac{1}{\kappa_1 a}\right)^2+\left(\frac{1}{\kappa_2 a}\right)^2\right]^2=\left[\frac{n_1^2 J'_\ell(\kappa_1 a)}{\kappa_1 a J_\ell(\kappa_1 a)}+\frac{n_2^2 K'_\ell(\kappa_2 a)}{\kappa_2 a K_\ell(\kappa_2 a)}\right]\times\left[\frac{J'_\ell(\kappa_1 a)}{\kappa_1 a J_\ell(\kappa_1 a)}+\frac{K'_\ell(\kappa_2 a)}{\kappa_2 a K_\ell(\kappa_2 a)}\right]. \tag{22}$$

This equation is to be solved in favor of β (the prime here is used to denote the derivatives of the Bessel functions with respect to their argument). In the search of solutions to Eq. (22), it is convenient to define the normalized frequency V as

$$V=\sqrt{(\kappa_1 a)^2+(\kappa_2 a)^2}=ak\sqrt{n_1^2-n_2^2}. \tag{23}$$

This quantity only depends on the fiber parameters, and its importance stems from the fact that since $\kappa_1 a \leq V$ and $\kappa_2 a \leq V$, more and more solutions – each corresponding to a different fiber mode – can be found by increasing its value (A graphical approach to solving the characteristic equation is reviewed in Ref. [3]). For $\ell = 1$, Eq. (22) is known to admit solutions for arbitrary values of V, and it admits only a single solution for $V \leq V_{0,1}$, where $V_{0,\,1} \simeq 2.405$ is the first zero of J_0 [3]. In this situation, only one mode is guided, and the fiber is referred to as a single-mode fiber. At a given wavelength, the single-mode operation can be attained either by reducing the core radius a or by reducing the numerical aperture $\sqrt{n_1^2-n_2^2}$. The one mode guided in a single-mode fiber is referred to as the fundamental mode and is labeled as HE_{11} [3].

The derivation reviewed in this section suggests that the analysis of the fiber modes, which is straightforward in principle, is tedious from the algebraic standpoint, and extracting the modes' characteristics on the basis of their analytical expressions is rather complicated, even in the case of the fundamental mode. Fortunately, this mathematical hardship reduces considerably in the relevant regime of weak guidance, which is described in the following section.

Weakly Guiding Fibers: Linearly Polarized Modes

The condition of weak guidance is characterized by a small refractive-index difference $\Delta n = n_1 - n_2 \ll n_{1,2}$. Typical values of $\Delta n/n_{1,2}$ are of the order of 0.1%. The key observations enabling a simplified study of the fiber modes under this assumption are (i) that the magnitude of the longitudinal components of the modes is small compared to that of transverse components and (ii) that the modes polarization is linear and fairly constant across the entire fiber section. These assumptions will be justified later. Therefore, the starting point is assuming $F_z = 0$, and

$$\vec{F}_t=\frac{J_\ell(\kappa_1 r)}{J_\ell(\kappa_1 a)}\cos(\ell\phi)\widehat{y}, r \leq a, \tag{24}$$

$$\vec{F}_t = \frac{K_\ell(\kappa_2 r)}{K_\ell(\kappa_2 a)} \cos(\ell\phi)\hat{y}, r \geq a, \tag{25}$$

so that $\vec{F}$ is a solution of Helmholtz equation (14) fulfilling the boundary conditions for the tangential component of the electric field. Note that this solution is approximate, as also the radial component of $\vec{F}$ is continuous at the core-cladding interface, whereas it should be larger in the cladding by the factor n_1^2/n_2^2 (which is approximately one for a small index difference). From $\nabla \times \vec{\tilde{e}} = i\omega\mu_0 \vec{\tilde{h}}$, one can readily find

$$\vec{G} = \frac{\beta}{\omega\mu_0}\hat{z} \times \vec{F}_t - \frac{i}{\omega\mu_0}\frac{\partial F_y}{\partial x}\hat{z}, \tag{26}$$

and, after some algebra involving some of the Bessel functions recurrence relations,

$$G_z = \frac{i\kappa_1}{2\omega\mu_0}\left\{\frac{J_{\ell-1}(\kappa_1 r)}{J_\ell(\kappa_1 r)}\cos[(\ell-1)\phi] - \frac{J_{\ell+1}(\kappa_1 r)}{J_\ell(\kappa_1 r)}\cos[(\ell+1)\phi]\right\}, r \leq a \tag{27}$$

$$G_z = \frac{i\kappa_2}{2\omega\mu_0}\left\{-\frac{K_{\ell-1}(\kappa_2 r)}{K_\ell(\kappa_2 r)}\cos[(\ell-1)\phi] - \frac{K_{\ell+1}(\kappa_2 r)}{K_\ell(\kappa_2 r)}\cos[(\ell+1)\phi]\right\}, r \geq a. \tag{28}$$

Note that κ_1 and κ_2 are of the order of $\sqrt{\Delta n}$, and therefore the longitudinal component of $\vec{G}$ is consistently small within the weakly guiding approximation, and therefore $\vec{G} \simeq \frac{\beta}{\omega\mu_0}\hat{z} \times \vec{F}$. On the other hand, the transverse component of $\vec{G}$ allows derivation of the longitudinal component of $\vec{F}$ using $\nabla \times \vec{\tilde{h}} = -i\omega\varepsilon_0\varepsilon_r \vec{\tilde{e}}$, with the result

$$F_z = \frac{i\beta\kappa_1}{2\omega^2\varepsilon_0\mu_0}\left\{\frac{J_{\ell-1}(\kappa_1 r)}{n_1^2 J_\ell(\kappa_1 r)}\cos[(\ell-1)\phi] - \frac{J_{\ell+1}(\kappa_1 r)}{n_1^2 J_\ell(\kappa_1 r)}\cos[(\ell+1)\phi]\right\}, r \leq a \tag{29}$$

$$F_z = \frac{i\beta\kappa_2}{2\omega^2\varepsilon_0\mu_0}\left\{-\frac{K_{\ell-1}(\kappa_2 r)}{n_2^2 K_\ell(\kappa_2 r)}\cos[(\ell-1)\phi] - \frac{K_{\ell+1}(\kappa_2 r)}{n_2^2 K_\ell(\kappa_2 r)}\cos[(\ell+1)\phi]\right\}, r \geq a, \tag{30}$$

showing that F_z is also small. The continuity of F_z and G_z at the core-cladding interface follows from imposing the equality between the coefficients of $\cos[(\ell+1)\phi]$ (or, equivalently, of $\cos[(\ell-1)\phi]$), which yields the following characteristic equation

$$\kappa_1 a \frac{J_{\ell+1}(\kappa_1 a)}{J_\ell(\kappa_1 r)} = \sqrt{V^2 - (\kappa_1 a)^2}\frac{K_{\ell+1}\left(\sqrt{V^2 - (\kappa_1 a)^2}\right)}{K_\ell\left(\sqrt{V^2 - (\kappa_1 a)^2}\right)}. \tag{31}$$

This equation always admits a solution for $\ell = 0$, and similarly to the case of the exact fiber modes, this is unique for $V < 2.405$. The modes of a weakly guiding fiber are denoted as $\mathrm{LP}_{\ell n}$, where LP stands for *linearly polarized*, and n is the solution number for a given value of ℓ. The fundamental mode is therefore denoted as LP_{01}.

Note that by starting with an $\widehat{x}$-polarized electric field mode or by replacing $\cos(\ell\phi)$ with $\sin(\ell\phi)$ would yield the same characteristic equation. Therefore, LP modes are fourfold degenerate for $\ell \neq 0$ (with respect to their dependence on ϕ and to polarization), while they are only twofold degenerate for $\ell = 0$ (with respect to polarization only). The fiber is single-mode if $V < 2.405$, consistent with the exact mode analysis, while additional modes are guided as V is increased. This is illustrated in the left panel of Fig. 2, where a graphical approach to solving the characteristic equation is implemented for three values of the normalized frequency V. The right panel of the same figure shows, by solid lines, the corresponding fundamental-mode amplitudes versus the normalized radial coordinate r/a and, by dashed lines, their Gaussian approximation proportional to $\exp\left(-r^2/r_0^2\right)$, where the parameter r_0 is taken from Marcuse's formula $r_0/a \simeq \sqrt{2/5} + 1.478V^{-3/2} + 4.76V^{-6}$ [4] (The mode amplitudes are normalized so as to yield the same value of $\int \left|\vec{F}\right|^2 \mathrm{d}x\mathrm{d}y$). The inset is a plot of Marcuse's formula.

We conclude this section by recalling that both true fiber modes and LP modes form orthogonal bases according to the orthogonality definition $\int \widehat{z} \cdot \left(\vec{F}_n \times \vec{G}_m^*\right) \mathrm{d}x\mathrm{d}y \propto \delta_{n,m}$, where by $\delta_{n,m}$, we denote the Kroneker delta [5, 6].

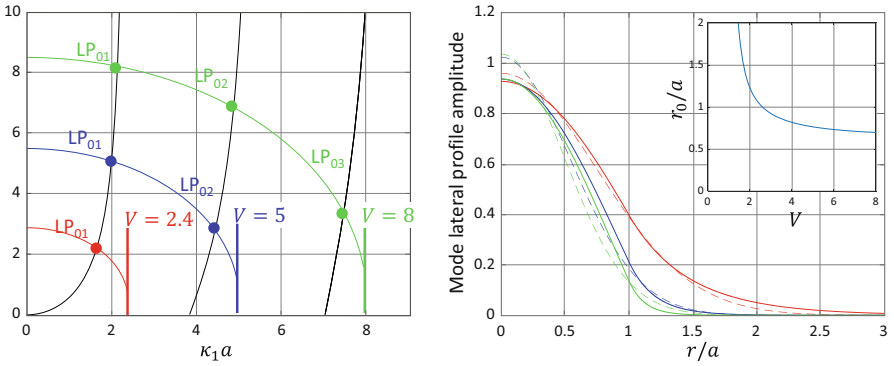

Fig. 2 The left panel illustrates a graphical approach to solving Eq. (31). The black curves are a plot of the left-hand side of the equality. The red, blue, and green curves are a plot of the right-hand side of the equality for the displayed values of the normalized frequency V. The dots correspond to the solutions, whose number increases as the value of V is increased. The right panel shows a plot of the amplitudes of the corresponding fundamental modes versus the radial coordinate r normalized to the core radius a. The solid curves refer to the exact expression of the LP_{01} mode amplitude, while the dashed curves refer to the Gaussian approximation proportional to $\exp\left(-r^2/r_0^2\right)$, where the parameter r_0 is taken from Marcuse's formula $r_0/a \simeq \sqrt{2/5} + 1.478V^{-3/2} + 4.76V^{-6}$ [4]. The inset is a plot of Marcuse's formula

Normalization and Power

The mean power carried by the electromagnetic field in a given mode is the flux of the Poynting vector throughout the fiber cross section, which according to Eqs. (10) and (11), is given by $P = \frac{|A|^2}{2} \int \mathrm{Re}\left\{ \left(\vec{F} \times \vec{G}^* \right) \cdot \hat{z} \right\} \mathrm{d}x\mathrm{d}y$. This expression suggests that by defining $A = E/\mathcal{N}$, with

$$\mathcal{N}^2 = \frac{1}{2} \int \mathrm{Re}\left\{ \left(\vec{F} \times \vec{G}^* \right) \cdot \hat{z} \right\} \mathrm{d}x\mathrm{d}y, \tag{32}$$

the complex amplitude E describing the mode excitation is such that $|E|^2$ is the power in Watts carried by the electromagnetic field in that mode. In the weakly guiding approximation $\mathcal{N}^2 \simeq \frac{n_{\mathrm{eff}}}{2Z_0} \int \left| \vec{F}_t \right|^2 \mathrm{d}x\mathrm{d}y$, where $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ is the vacuum impedance, as follows from $\vec{G} \simeq \frac{\beta}{\omega\mu_0} \hat{z} \times \vec{F}$ (Analytical approximations for $\mathcal{N}$ can be found in Ref. [3]). In the remainder of this chapter, we assume such normalization. Note that the orthogonality condition can hence be expressed also in the following form

$$\int \left(\vec{F}_n \times \vec{G}_m^* + \vec{F}_m^* \times \vec{G}_n \right) \cdot \hat{z} \mathrm{d}x\mathrm{d}y = 4\mathcal{N}^2 \delta_{n,m}, \tag{33}$$

which will be useful in the next section.

Transmission of Information-Carrying Signals

In what follows, we focus on the transmission of information-carrying signals in single-mode fibers. In this context, two independent waveforms are used to excite the x- and y-polarized replicas of the fundamental mode, which we denote as $\vec{F}_{1x} = F\hat{x}$ and $\vec{F}_{1y} = F\hat{y}$, respectively, where $F = F(x, y, \omega)$ is the amplitude of LP_{01}. We express the positive-frequency content of the propagating field as

$$\tilde{\vec{e}}\,(x, y, z, \omega) = \tilde{E}_x(z, \omega - \omega_0) \frac{\vec{F}_{1x}(x, y, \omega)}{2\mathcal{N}(\omega)} + \tilde{E}_y(z, \omega - \omega_0) \frac{\vec{F}_{1y}(x, y, \omega)}{2\mathcal{N}(\omega)}, \omega > 0, \tag{34}$$

where $\tilde{E}_x(z, \omega)$ and $\tilde{E}_y(z, \omega)$ describe the baseband excitations of the two polarization modes. At $z = 0$, they describe the input waveforms $\tilde{E}_{\mathrm{in},x}(\omega) = \tilde{E}_x(0, \omega)$ and $\tilde{E}_{\mathrm{in},y}(\omega) = \tilde{E}_y(0, \omega)$ that are used to modulate the transmitter laser, whose frequency ω_0 is referred to as the carrier frequency. A similar expression describes the magnetic field,

$$\tilde{\vec{h}}\,(x, y, z, \omega) = \tilde{E}_x(z, \omega - \omega_0) \frac{\vec{G}_{1x}(x, y, \omega)}{2\mathcal{N}(\omega)} + \tilde{E}_y(z, \omega - \omega_0) \frac{\vec{G}_{1y}(x, y, \omega)}{2\mathcal{N}(\omega)}, \omega > 0. \tag{35}$$

An alternative more intuitive expression of Eq. (34) follows from the notion that the modes' lateral profiles depend only weakly on frequency, so that within the signal bandwidth $\vec{F}(x,y,\omega) \simeq \vec{F}(x,y,\omega_0)$ and $\mathcal{N}(\omega) \simeq \mathcal{N}(\omega_0)$. Therefore, within this approximation, Eq. (34) can be expressed in the time domain as

$$\vec{e}\,(x,y,z,t) \simeq \mathrm{Re}\left\{ \left[E_x(z,t)\hat{x} + E_y(z,t)\hat{y}\right] \frac{F(x,y,\omega_0)}{\mathcal{N}(\omega_0)} e^{-i\omega_0 t} \right\}. \tag{36}$$

The waveforms E_x and E_y are the inverse Fourier transforms of $\widetilde{E}_x$ and $\widetilde{E}_y$, respectively, and they form the field complex envelopes in the two polarizations. Their evolution along a fiber is a consequence of the fiber's nonideal nature, in keeping with what we have seen in previous sections, that in ideal fibers, these quantities are independent of z. Describing the character of these nonideal characteristics is at the core of this chapter. In what follows, we review the derivation of the propagation equations for E_x and E_y, where the deviation of the constitutive relation between $\vec{p}$ and $\vec{e}$ from the simplest form $\vec{p} = \varepsilon_0(\varepsilon_r - 1)\vec{e}$ assumed in the section "Fiber Modes" is taken into account within a perturbation approach.

The derivation of the perturbed equations goes as follows. We start from Maxwell's equations in the frequency domain,

$$\nabla \times \widetilde{\vec{e}} = i\omega\mu_0 \widetilde{\vec{h}}, \tag{37}$$

$$\nabla \times \widetilde{\vec{h}} = -i\omega\varepsilon_0\varepsilon_r \widetilde{\vec{e}} - i\omega\varepsilon_0 \widetilde{\Delta\vec{p}}, \tag{38}$$

where $\widetilde{\Delta\vec{p}}$ is the Fourier transform of $\Delta\vec{p} = \vec{p} - \varepsilon_0(\varepsilon_r - 1)\vec{e}$. For fiber modes $\vec{E}_n = \vec{F}_n \exp(i\beta z)$ and $\vec{H}_n = \vec{G}_n \exp(i\beta z)$, with $n = x,y$, they simplify to

$$\nabla \times \vec{E}_n = i\omega\mu_0 \vec{H}_n, \tag{39}$$

$$\nabla \times \vec{H}_n = -i\omega\varepsilon_0\varepsilon_r \vec{E}_n. \tag{40}$$

We then perform the scalar product between Eq. (37) and $\vec{H}_n^*$, and between Eq. (38) and $\vec{E}_n^*$, and use the identity $\nabla \cdot \left(\vec{A} \times \vec{B}\right) = \vec{B} \cdot \nabla \times \vec{A} - \vec{A} \cdot \nabla \times \vec{B}$, with the result

$$\nabla \cdot \left(\widetilde{\vec{e}} \times \vec{H}_n^*\right) + \widetilde{\vec{e}} \cdot \nabla \times \vec{H}_n^* = i\omega\mu_0 \widetilde{\vec{h}} \cdot \vec{H}_n^*, \tag{41}$$

$$\nabla \cdot \left(\widetilde{\vec{h}} \times \vec{E}_n^*\right) + \widetilde{\vec{h}} \cdot \nabla \times \vec{E}_n^* = -i\omega\varepsilon_0\varepsilon_r \widetilde{\vec{e}} \cdot \vec{E}_n^* - i\omega \widetilde{\Delta\vec{p}} \cdot \vec{E}_n^*. \tag{42}$$

Substituting $\nabla \times \vec{E}_n$ and $\nabla \times \vec{H}_n$ from Eqs. (39) and (40), respectively, and subtracting the resulting equations from each other, we obtain the following equality,

$$\nabla \cdot \left(\widetilde{\vec{e}} \times \vec{H}_n^* - \widetilde{\vec{h}} \times \vec{E}_n^* \right) = i\omega \Delta \widetilde{\vec{p}} \cdot \vec{E}_n^*. \tag{43}$$

To conclude, we integrate both sides of the above over the (x, y) plane, to which end we use $\nabla = \nabla_t + \hat{z}\frac{\partial}{\partial z}$, and recall that owing to Green's theorem, the term ∇_t does not contribute to the integral. Therefore, using the expansions (34) and (35), we are left with the evolution equation

$$\left(\frac{\partial}{\partial z} - i\beta \right) \sum_{u=x,y} \widetilde{E}_u(z, \omega - \omega_0) \int \hat{z} \cdot \frac{\vec{F}_u \times \vec{G}_n^* - \vec{G}_u \times \vec{F}_n^*}{2\mathcal{N}} \mathrm{d}x\mathrm{d}y$$
$$= i\omega \int \Delta \widetilde{\vec{p}} \cdot \vec{F}_n^* \mathrm{d}x\mathrm{d}y, \tag{44}$$

which, using the orthogonality condition (33), yields

$$\left(\frac{\partial}{\partial z} - i\beta \right) \widetilde{E}_n(z, \omega - \omega_0) = \frac{i\omega}{2\mathcal{N}} \int \Delta \widetilde{\vec{p}} \cdot \vec{F}_n^* \mathrm{d}x\mathrm{d}y. \tag{45}$$

This equation can be further simplified within the narrow-band signal approximation discussed in the derivation of (36), which allows replacement of the ratio $\omega/\mathcal{N}(\omega)$ with $\omega_0/\mathcal{N}(\omega_0)$.

In what follows, we adopt the bra-ket notation [7] for the description of the field vector and its Fourier transform, which we define as

$$|E(z,t)\rangle = \begin{bmatrix} E_x(z,t) \\ E_y(z,t) \end{bmatrix}, \quad |\widetilde{E}(z,\omega)\rangle = \begin{bmatrix} \widetilde{E}_x(z,\omega) \\ \widetilde{E}_y(z,\omega) \end{bmatrix}. \tag{46}$$

With this notation, the derived evolution equation can be cast in the following compact form,

$$\frac{\partial |\widetilde{E}(z, \omega - \omega_0)\rangle}{\partial z} = i\beta(\omega) |\widetilde{E}(z, \omega - \omega_0)\rangle + |\widetilde{L}\rangle \tag{47}$$

where the two components of $|\widetilde{L}\rangle$ are given by

$$\widetilde{L}_u = \frac{i\omega_0}{2\mathcal{N}(\omega_0)} \int \vec{F}_{1u}^*(x,y) \cdot \Delta \widetilde{\vec{p}}\ (x,y,z,\omega)\, \mathrm{d}x\mathrm{d}y, \quad u = x, y. \tag{48}$$

In the time domain, the perturbation term can be expressed as

$$L_u = \frac{i\omega_0}{2\mathcal{N}(\omega_0)} \int \vec{F}_{1u}^*(x,y) \cdot \Delta \vec{P}_{\omega_0}(x,y,z,t)\, \mathrm{d}x\mathrm{d}y, \quad u = x, y, \tag{49}$$

where $\Delta \vec{P}_{\omega_0}(x, y, z, t)$ corresponds to the frequency content of $\Delta \vec{p}$ around ω_0. The time-domain representation will be seen to be more convenient for the study of the nonlinear effects considered in this chapter.

Chromatic Dispersion

In the absence of perturbations to the ideal fiber structure, the only cause of distortion of the propagating waveform is represented by the frequency dependence of the propagation constant β. This effect is referred to as waveguide dispersion. In practice, the waveguide dispersion is accompanied by additional dispersion due to the frequency dependence of the fiber refractive index, and this contribution is referred to as the material dispersion. The simplest constitutive relation that can describe this effect is of the form $\vec{p} = \varepsilon_0 \chi(t) \otimes \vec{e}$, where $\chi(t) = \varepsilon_r(t) - 1$ includes a small deviation from the instantaneous contribution considered in the mode analysis. Within this representation, the frequency-dependent refractive index is defined trough the relation $\widetilde{\chi}_R(\omega) = n^2(\omega) - 1$, where $\widetilde{\chi}_R(\omega)$ is the real part of $\widetilde{\chi}(\omega)$, and the functional dependence of the refractive index on frequency is typically modeled with the Sellmeier expansion [8]. Therefore, material dispersion is readily taken into account through the frequency dependence of β. On the other hand, the presence of an imaginary part of $\widetilde{\chi}$ introduces a perturbation, and its implications are discussed in the next section.

Dispersion is conveniently studied by using a Taylor expansion of $\beta(\omega)$ within the signal bandwidth. Typically, a second-order expansion suffices,

$$\beta(\omega) \simeq \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega - \omega_0)^2, \tag{50}$$

and substituting this expression into Eq. (47) and performing the inverse Fourier transform yields the following propagation equation in the time domain

$$\frac{\partial |E\rangle}{\partial z} = i\beta_0 |E\rangle - \beta_1 \frac{\partial |E\rangle}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2 |E\rangle}{\partial t^2}. \tag{51}$$

The first and second terms do not produce any waveform distortions, and they can be removed by means of a simple transformation. In fact, for $\beta_2 = 0$, the solution of Eq. (51) is $|E(z, t)\rangle = \exp(i\beta_0 z)|E(0, t - \beta_1 z)\rangle$, which shows that the effect of propagation is to introduce the phase $\beta_0 z$ and the group delay $\beta_1 z$. A phase velocity and a group velocity are defined accordingly as $v_p = \beta_0^{-1}$ and $v_g = \beta_1^{-1}$. Therefore, by defining $|E(z, t)\rangle = \exp(i\beta_0 z)|E'(z, \tau)\rangle$, with $\tau = t - \beta_1 z$, the propagation equation for $|E'(z, \tau)\rangle$ is obtained from (51) in the following form

$$\frac{\partial |E'\rangle}{\partial z} = -i\frac{\beta_2}{2}\frac{\partial^2 |E'\rangle}{\partial \tau^2}. \tag{52}$$

In what follows, we drop the primes and denote the time in the moving reference frame simply by t, rather than τ. The formal solution of Eq. (52) is therefore $|E\ (z,\ t)\rangle\ =\ \exp(-i\ z\ \partial^2/\partial\ t^2)|E\ (0,\ t)\rangle$ [9], while in the frequency domain, it is obviously given by

$$|\widetilde{E}\,(z,\ \ \omega)\rangle = e^{\frac{i}{2}\beta_2\,\omega^2 z}\ |\widetilde{E}\,(z,\ \ \omega)\rangle. \tag{53}$$

The coefficient β_2 describes group velocity dispersion, as is implied by $\mathrm{d}v_g/\mathrm{d}\omega = -\beta_2/\beta_1^2$. Some intuition for the effect of this phenomenon, most commonly referred to as chromatic dispersion, can be developed by considering a signal of bandwidth B, so that different narrow sub-bands of the signal propagate with different group velocities. As a result, the propagating waveform undergoes temporal spreading, and the time duration of the dispersed waveform is given by the delay between the extreme spectral components at the frequencies $\omega_{\min}$ and $\omega_{\max}$,

$$T = |\beta_1(\omega_{\max})L - \beta_1(\omega_{\min})L| \simeq 2\pi|\ \beta_2|BL, \tag{54}$$

where L is the fiber length. An alternative parameter is commonly used to quantify chromatic dispersion, which is referred to as the dispersion parameter and denoted by D. This parameter is defined through the equality $T = D\Delta\lambda$, where $\Delta\lambda$ is the signal bandwidth in the wavelength domain, namely $D = -2\pi c\beta_2/\lambda^2$. At 1550 nm, typical values of D are 17 ps/km-nm in standard single-mode fiber (SSMF, G.652), and less than 4 ps/km-nm in dispersion-shifted fiber (DSF, G.653). The situation in which the group velocity grows with frequency is referred to as anomalous dispersion ($D > 0$) and is typical in communications fibers around 1550 nm, while the opposite case ($D < 0$) is referred to as normal dispersion, and in the 1550 nm region, it characterizes special fibers, like dispersion compensating fibers that are discussed subsequently.

While Eq. (54) provides a qualitative description of dispersive pulse broadening, a quantitative analysis may be carried out by looking at the broadening factor $b(z) = \sigma(z)/\sigma(0)$, where $\sigma(z)$ is the root-mean-square duration of the propagating waveform, which is defined as

$$\sigma^2\,(z) = \int t^2 p\,(z,t)\,\mathrm{d}t - \left[\int t p\,(z,t)\,\mathrm{d}t\right]^2 \tag{55}$$

with $p\ (z,\ t)\ =\ |E\ (z,\ t)\ |^2/\int\ |E\ (z,\ \ t)\ |^2\ \mathrm{d}t$ having the properties of a probability density function. A closed-form expression for the broadening factor can be obtained for arbitrary input waveforms by taking advantage of the similarity between Eq. (52) and the Schrödinger equation studied in quantum mechanics. The resulting form of the broadening factor is $b^2\ (z) = 1 + b_1 z + b_2 z^2$ [10], and it shows that asymptotically, for large propagation distances, $b\ (z) \propto z$.

It is instructive to consider the case of an input waveform with a Gaussian profile, that is $E\,(0,t) = E_0 \exp\ \left[-\frac{t^2}{4\sigma_0^2}\ \ (1 + iC)\right]$, where $\sigma_0 = \sigma(0)$ is the root-mean-square

duration of the intensity waveform and C is the chirp parameter. Using the known relation between a Gaussian waveform and its Fourier transform, one can readily evaluate $\widetilde{E}(z,\omega) = E_0\sqrt{\frac{4\pi\sigma_0^2}{1+iC}}\exp\left\{-\frac{1}{2}\left[\frac{2\sigma_0^2}{1+iC} - i\beta_2 z\right]\omega^2\right\}$, and hence $E(z,t) = E_0\sqrt{\frac{2\sigma_0^2}{1+iC}\left[\frac{2\sigma_0^2}{1+iC} - i\beta_2 z\right]^{-1}}\exp\left\{-\frac{1}{2}\left[\frac{2\sigma_0^2}{1+iC} - i\beta_2 z\right]^{-1}t^2\right\}$. This yields the following expression for the broadening factor

$$b(z) = \sqrt{\left(1+\frac{C\beta_2 z}{2\sigma_0^2}\right)^2 + \left(\frac{\beta_2 z}{2\sigma_0^2}\right)^2} \simeq \sqrt{1+C^2}\frac{z}{2\sigma_0^2/|\beta_2|}, \tag{56}$$

where the second equality holds asymptotically for large values of z. The quantity $L_D = T_0^2/|\beta_2|$, with $T_0^2 = 2\sigma_0^2$, is referred to as the *dispersion length*, and it represents the propagation length over which the effect of chromatic dispersion becomes appreciable. The broadening factor has a minimum for $z_{\min} = -\text{sign}(\beta_2)\frac{C}{1+C^2}L_D$, which is physically meaningful when β_2 and C have opposite signs, and hence $z_{\min} > 0$. At $z = z_{\min}$, the pulse loses its chirp and becomes transform-limited, namely it achieves a minimum time duration for its bandwidth (note that the signal bandwidth is not affected by chromatic dispersion, which is an all-pass filter).

This approach to characterizing CD is particularly useful in systems with intensity modulation and direct detection, where it is important to avoid pulse spreading outside one symbol time slot. In these systems, chromatic dispersion can be compensated optically by means of additional dispersion of opposite sign, so that $\beta_2 L + \beta_{2,\text{ comp}} L_{\text{comp}} = 0$, where $\beta_{2,\text{ comp}}$ and L_{comp} are the chromatic dispersion coefficient and length of the compensating fiber, respectively. The dispersion of compensating fibers can be fairly large (of the order of 300 ps/km-nm at 1550 nm), so that short compensating fiber spans can be used to compensate large amounts of accumulated chromatic dispersion. Optical chromatic dispersion compensation can be implemented at the end of the link, or in a distributed fashion, typically by deploying compensating modules within the in-line amplifiers, and the resulting system is referred to as dispersion-managed [11]. In digital coherent systems that are deployed today, chromatic dispersion is compensated electronically by digital signal processing after coherent detection, thereby considerably simplifying the link design. As is discussed in the section "Nonlinear Interference Noise in the Context of Digital Coherent Systems," chromatic dispersion is the primary propagation effect in coherent systems. It is instructive to note that in these systems, chromatic dispersion accumulates to the extent that it transforms the launched waveform into its Fourier transform [12],

$$E(z,t) \simeq \sqrt{\frac{i}{2\pi\beta_2 z}}e^{-\frac{it^2}{2\beta_2 z}}\widetilde{E}\left(\frac{t}{\beta_2 z}\right), \tag{57}$$

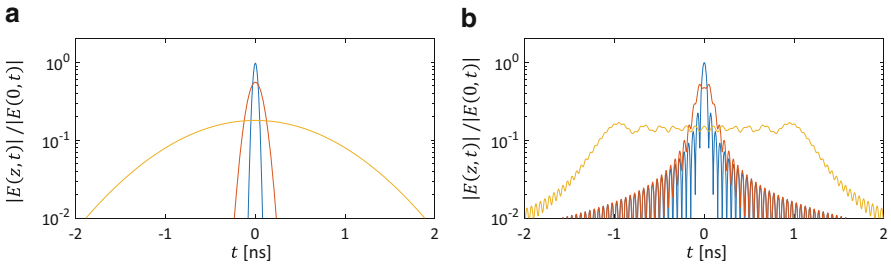

Fig. 3 (**a**) Normalized amplitude of an input Gaussian waveform with a full width at half-maximum of 60 ps for propagation distances of $z = 10$, 100, and 1000 km, in a fiber with $\beta_2 = -21$ ps^2/km. (**b**) Same as (**a**) for an input sinc waveform with the same full width at half-maximum

which is valid for $2\pi|\beta_2|zB \gg 1$. Figure 3a and b show the evolution of an input Gaussian waveform and a sinc waveform with the same full width at half-maximum amplitude of 60 ps, respectively, for $z = 10$, 100, and 1000 km. It can be seen that while the Gaussian waveform preserves its shape, the sinc pulse approaches the squared shape of its Fourier transform.

We conclude this section by recalling that higher-order dispersion may be included in the analysis by using a higher-order Taylor expansion of $\beta(\omega)$, thereby introducing additional time-derivative terms in the time-domain propagation equation. This extension becomes relevant in the context of ultrashort pulse transmission, where the signal bandwidth is considerably broader than signals typically used in optical communications.

Loss, Gain, and Noise

The linear constitutive relation $\vec{p} = \varepsilon_0 \chi(t) \otimes \vec{e}$ introduced in the previous section shows that the dielectric susceptibility is the impulse response of the fiber atoms and molecules to the electric field. As such, $\chi(t)$ must fulfill the causality requirement, which means that the real and imaginary parts of its Fourier transform, $\widetilde{\chi}_R(\omega)$ and $\widetilde{\chi}_I(\omega)$, are related to each other through the Kramers-Kronig relations, and one cannot vanish without the other vanishing too [13] (unless $\chi(t) = \chi\delta(t)$ where χ is a real-valued constant, as assumed in the mode analysis). Here we are interested in the non-vanishing $\widetilde{\chi}_I(\omega)$, which is known to introduce loss or gain in the propagation equation. This can be readily seen by substituting $\Delta\widetilde{\vec{p}} = i\varepsilon_0\chi_I(\omega)\widetilde{\vec{e}}$ into Eq. (48), which yields

$$\langle \widetilde{L} \rangle = -\frac{\varepsilon_0 \chi_I(\omega)\omega_0}{4\mathcal{N}^2(\omega_0)} \int \left|\vec{F}(x,y)\right|^2 \langle \widetilde{E}(z,\ \omega-\omega_0)\rangle \,\mathrm{d}x\mathrm{d}y = -\frac{\omega_0 \chi_I(\omega)}{2cn_{\mathrm{eff}}} \langle \widetilde{E}(z,\ \omega-\omega_0)\rangle \tag{58}$$

where the second equality holds within the LP approximation. By defining $\widetilde{\alpha}(\omega) = \frac{\omega_0}{cn_{\text{eff}}}\chi_I(\omega)$, one obtains the familiar term accounting for loss in the propagation equation,

$$\frac{\partial|\widetilde{E}(z,\ \omega-\omega_0)\rangle}{\partial z} = -\frac{\widetilde{\alpha}(\omega)}{2}|\widetilde{E}(z,\ \omega-\omega_0)\rangle + i\frac{\beta_2}{2}(\omega-\omega_0)^2 |\widetilde{E}(z,\ \omega-\omega_0)\rangle. \tag{59}$$

The physical mechanisms responsible for loss in optical fibers are scattering and absorption, and the fundamental lower bound to the fiber loss is set by the Rayleigh scattering process, which is caused by scatterers of sub-wavelength dimensions. Clearly, loss is frequency dependent, and this dependence is amply discussed in the literature; however, for modeling purposes, the loss coefficient can be approximated as constant within the C-band, $\widetilde{\alpha}(\omega) \simeq \alpha(\omega_0)$, so that in the time domain, the propagation equation becomes

$$\frac{\partial|E\rangle}{\partial z} = -\frac{\alpha}{2}|E\rangle - i\frac{\beta_2}{2}\frac{\partial^2|E\rangle}{\partial t^2}, \tag{60}$$

which implies that the effect of loss is to introduce an exponential decay of the complex amplitude of the propagating signal,

$$|E(z,t)\rangle = e^{-\frac{\alpha z}{2}}e^{-i\frac{\beta_2}{2}z\frac{\partial^2}{\partial t^2}}|E(0,\ t)\rangle, \tag{61}$$

$$|\widetilde{E}(z,\ \omega)\rangle = e^{-\frac{\alpha z}{2}}e^{i\frac{\beta_2}{2}z\omega^2}|\widetilde{E}(0,\ \omega)\rangle. \tag{62}$$

Attenuation is typically quantified with reference to the optical power $P(z, t) = |E_x(z, t)|^2 + |E_y(z, t)|^2 = \langle E(z,\ t)|E(z,\ t)\rangle$, which decays with propagation distance as $\exp(-\alpha z)$. Typical values for the loss coefficient are of the order of 0.2 dB/km, corresponding to an attenuation by a factor of approximately 100 after propagation over 100 km of fiber.

Attenuation is compensated by means of optical amplifiers deployed along the link (see Fig. 4), so that transmission distances of thousands of kilometers can be attained by using fibers spans whose length is in the range of several tens of kilometers. However, while restoring the optical power to the desired level, each amplifier introduces additive noise. We stress that this is not a technological issue but rather a fundamental effect dictated by quantum mechanics [14]. In particular, an optical amplifier providing a power gain of G produces complex-circular Gaussian noise, which is referred to as *amplified spontaneous emission* (ASE) noise, and whose spectral density in each polarization is given by $p_{\text{ASE}} = \hbar\omega n_{\text{sp}}(G-1)$, where $\hbar$ is the modified Plank's constant and n_{sp} is the so-called inversion factor, which is equal to 1 for ideal amplifiers and is greater than 1 in all practical cases (in typical amplifiers n_{sp} ranges between 1.6 and 2). It should be noted that the direct dependence of p_{ASE} on frequency is negligible in practice because the range in which ω

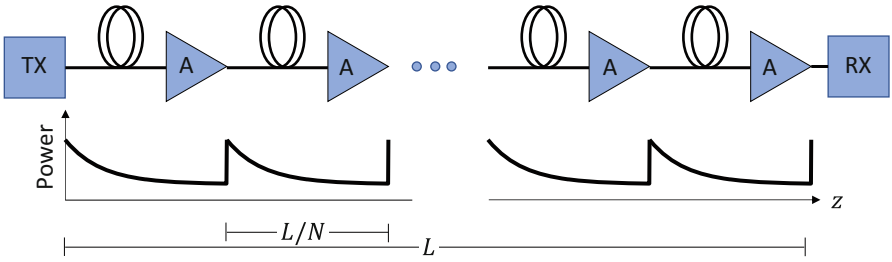

Fig. 4 A typical link of length L, where a transmitter and a receiver are connected by means of N fiber spans, each terminated by an optical amplifier. The power of the propagating signal decays exponentially along the fiber spans and is restored at each amplification stage

varies within the amplification bandwidth is very small relative to the central frequency of the amplification band. Therefore, ω can be safely replaced with the signal center frequency, and the noise approximated as white within regions where the amplification spectrum is reasonably flat (as it should be in communication systems). The amount of noise accumulated in a link of the kind sketched in Fig. 4 depends also on the number of amplifiers and the distance between them. In particular, the power of the noise impinging upon the receiver within the signal bandwidth B is given by $P_{\mathrm{ASE}} = B \times \hbar\omega_0 n_{\mathrm{sp}} (G-1) \times N = B\hbar\omega_0 n_{\mathrm{sp}} [\exp(\alpha L/N) - 1]N$, where the second equality follows from the fact that each amplifier compensates for the loss of a fiber span of length L/N, with L denoting the link length. This simple expression shows that the power of the noise impairing the received signal is a monotonically reducing function of N, thereby implying that continuously distributed amplification is optimal, and the corresponding noise power is $\lim_{N \to \infty} P_{\mathrm{ASE}} = B\hbar\omega_0 n_{\mathrm{sp}} \alpha L$. Distributed amplification can be approached in practice by implementing Raman amplification [15] (Raman amplification is discussed in ▶ Chap. 4, "Optical Amplification," of this handbook). Nonetheless, practical considerations dictate that the vast majority of systems that are used today are based on erbium-doped fiber amplifiers (EDFAs), spaced by 40–100 km of transmission fiber, so that at 0.2 dB/km fiber loss, they are required to provide a power gain of approximately 8–20 dB. Obviously, deploying more closely spaced amplifiers in order to reduce the ASE power implies higher costs.

It is instructive to notice that by pre-amplifying each span (as opposed to terminating each span with an amplifier), the ASE noise power is minimized by deploying a single amplifier at the input of the link. However, in this case, unrealistically large gains and optical power levels at the fiber input would be required (For instance, an optical power of 100 W should be injected in a 200-km link to receive 0.01 mW). If such amplifiers were available, this would also be detrimental in terms of nonlinear signal distortions, which are discussed in the section "Nonlinear Propagation." However, one may check that distributed amplification is optimal also in this case, if the optical signal power at the fiber input is constrained to some prescribed level.

Polarization Effects

Polarization effects result from manufacturing imperfections and deployment-related issues that make the fiber lose its ideal circular symmetry. In this situation, the relation between $\vec{p}$ and $\vec{e}$ becomes anisotropic, and the perturbation can be expressed in the frequency domain as $\Delta \widetilde{\vec{p}} = \varepsilon_0 \widetilde{\chi}_P(\omega)\, \widetilde{\vec{e}}$, where $\widetilde{\chi}_P(z, \omega)$ is a traceless 2×2 tensor. In order to develop some intuition regarding the properties of $\widetilde{\chi}_P(\omega)$, we start by considering the case of a small refractive-index difference δn between the x- and y-polarized versions of the fundamental mode, corresponding to the relative dielectric constants $(n_{\text{eff}} \pm \delta n/2)^2 \simeq n_{\text{eff}}^2 \pm n_{\text{eff}}\, \delta n$. Therefore,

$$\Delta \widetilde{\vec{p}} = \varepsilon_0 n_{\text{eff}} \begin{bmatrix} \delta n & 0 \\ 0 & -\delta n \end{bmatrix} \widetilde{\vec{e}}, \tag{63}$$

and

$$\widetilde{\chi}_P = n_{\text{eff}} \begin{bmatrix} \delta n & 0 \\ 0 & -\delta n \end{bmatrix}. \tag{64}$$

This gives

$$\begin{aligned} \widetilde{L}_{x,y} &= \pm \frac{i\omega_0}{2\mathcal{N}(\omega_0)} \frac{\varepsilon_0 n_{\text{eff}}\, \delta n}{2} \int |F(x,y)|^2 \widetilde{E}_{x,y}(z, \omega - \omega_0) \mathrm{d}x\mathrm{d}y \\ &= \pm i \frac{\omega_0}{c} \frac{\delta n(z, \omega)}{2} \widetilde{E}_{x,y}(z, \omega - \omega_0), \end{aligned} \tag{65}$$

and by defining $\delta\beta = \frac{\omega_0}{c} \delta n$, the perturbation can be expressed as

$$|\widetilde{L}\rangle = \frac{i}{2} \begin{bmatrix} \delta\beta & 0 \\ 0 & -\delta\beta \end{bmatrix} |\widetilde{E}(z, \omega - \omega_0)\rangle. \tag{66}$$

In general, some difference in effective refractive index can be formed between any arbitrary pair of orthogonal polarization states. If we denote by $\mathbf{U}$ the unitary matrix expressing an arbitrary state vector in the basis formed by such two polarizations, then the perturbation term is obtained by simply replacing the diagonal matrix diag $\{\delta\beta, -\delta\beta\}$ with $\mathbf{B} = \mathbf{U}^{\dagger}\text{diag}\{\delta\beta, -\delta\beta\}\mathbf{U}$, where by the dagger we denote the Hermitian adjoint. The matrix $\mathbf{B}$ is Hermitian and traceless (the sum of the two eigenvalues vanishes), and therefore it implies that propagation is unitary, that is, the field power spectral density $\left\langle \widetilde{E} \middle| \widetilde{E} \right\rangle$ is preserved as the field vector evolves according to

$$\frac{\partial |\widetilde{E}(z, \omega - \omega_0)\rangle}{\partial z} = \frac{i}{2} \mathbf{B}(z, \omega)\, |\widetilde{E}(z, \omega - \omega_0)\rangle. \tag{67}$$

In fact, $\partial\ \langle\widetilde{E}\,|\,\widetilde{E}\rangle/\partial z = i\langle\widetilde{E}\,|\,\mathbf{B}\,|\,\widetilde{E}\rangle/2 - i\langle\widetilde{E}\,|\,\mathbf{B}^{\dagger}\,|\,\widetilde{E}\rangle/2 = 0$, as follows from $\mathbf{B}^{\dagger} = \mathbf{B}$. Therefore, the components of the propagated field vector $\widetilde{E}_x$ and $\widetilde{E}_y$ result from a unitary mixture of their input values, a phenomenon referred to as polarization-mode coupling. Note that this picture is not affected by the presence of loss of the kind described in the previous section, while it becomes considerably more involved with polarization-dependent loss. Finally, it is instructive to mention that by expanding the matrix $\mathbf{B}$ around the central frequency ω_0, the time-domain analogue of (67) is obtained in the form

$$\frac{\partial\,|\,E(z,t)\rangle}{\partial z} = \frac{i}{2}\mathbf{B}_0(z)\,|\,E(z,t)\rangle - \frac{1}{2}\mathbf{B}_1(z)\ \ \frac{\partial\,|\,E(z,t)\rangle}{\partial t} + \ldots, \tag{68}$$

where by $\mathbf{B}_0(z)$ and $\mathbf{B}_1(z)$, we denote the matrices $\mathbf{B}(z,\ \omega)$ and its frequency derivative evaluated at ω_0, respectively.

Jones and Stokes Representation

Jones calculus is a convenient framework to describe polarized light and its evolution in linear media. In this section, we focus on the case of unitary propagation, and define the Jones vector associated with the electric field as the field vector normalized to its modulus $|\,s\rangle = |\,\widetilde{E}\rangle/\langle\widetilde{E}\,|\,\widetilde{E}\rangle^{1/2}$, so that the unit-modulus condition can be expressed as $\langle s\,|\,s\rangle = |\,s_x\,|^2 + |\,s_y\,|^2 = 1$, and the scalar product between two Jones vectors is given by $\langle u\,|\,s\rangle = u_x^* s_x + u_x^* s_y$. The ratio between the magnitudes of s_x and s_y, and their relative phase determine the state of polarization of the field, which is defined by looking at the time trajectory drawn by the tip of the vector $\mathrm{Re}\,\{\exp(-i\omega\ t)\ |\,s\rangle\}$ in the x–y plane. This is a segment of unit length if s_x and s_y are in phase, a circle of unit radius if they have equal modulus and a phase difference of $\pm\pi/2$, whereas it is an ellipse in all of the other cases. The state of polarization is then referred to as linear, circular, or elliptical, respectively.

We now move to introducing the Stokes representation of the electric field. This is an alternative description based on the use of real-valued three-dimensional vectors, and it is isomorphic to the Jones representation [7]. If we denote by $\vec{s}$ the Stokes vector associated with the Jones vector $|s\rangle$, its three components are defined as

$$\begin{aligned} s_1 &= |\,s_x\,|^2 - |\,s_y\,|^2 = \langle s\,|\sigma_1\,|\,s\rangle = \mathrm{Tr}\{\sigma_1\,|\,s\rangle\,\langle s\,|\} \\ s_2 &= 2\,\mathrm{Re}\{s_x^* s_y\} = \langle s\,|\sigma_2\,|\,s\rangle = \mathrm{Tr}\{\sigma_2\,|\,s\rangle\,\langle s\,|\} \\ s_3 &= 2\mathrm{Im}\{s_x^* s_y\} = \langle s\,|\sigma_3\,|\,s\rangle = \mathrm{Tr}\{\sigma_3\,|\,s\rangle\,\langle s\,|\}, \end{aligned} \tag{69}$$

where σ_1, σ_2, and σ_3 are the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & -i \\ i & 0 \end{pmatrix}. \tag{70}$$

The Stokes vector can also be expressed more compactly as

$$\vec{s} = \langle s|\vec{\sigma}|s\rangle = \mathrm{Tr}\left\{\vec{\sigma} \mid s\rangle\langle s\mid\right\}, \tag{71}$$

where $\vec{\sigma}$ is a three-dimensional column vector whose components are the Pauli matrices. The Stokes vector length is $|\vec{s}| = \sqrt{s_1^2 + s_2^2 + s_3^2} = \langle s|s\rangle = 1$. As illustrated in Fig. 5, each polarization state corresponds to a point on the surface of what is known as the Poincaré sphere in the Stokes space: right- and left-circular polarizations are on the north- and south-poles, respectively, linear polarizations are on the equator (x and y on $\pm\hat{s}_1$, and $\pm 45^\circ$ on $\pm\hat{s}_2$), and elliptical polarizations span the rest of the sphere.

The relation between the Jones and Stokes vectors of Eq. (71) can be rearranged also as follows,

$$|s\rangle\langle s| = \frac{1}{2}\left(\mathbf{I} + \vec{s}\cdot\vec{\sigma}\right). \tag{72}$$

where the 2×2 matrix $|s\rangle\langle s|$ is the projection operator, which when applied to some given Jones vector, returns the component of that vector along $|s\rangle$. Equation (72) can be used to obtain the following relation between scalar products in the Jones and Stokes spaces

$$|\langle u|v\rangle|^2 = \frac{1}{2}\left(1 + \vec{u}\cdot\vec{v}\right), \tag{73}$$

which shows that orthogonal polarization states, for which $\langle u \mid v\rangle = 0$, are antiparallel in Stokes space, namely $\vec{u}\cdot\vec{v} = -1$.

The analysis of polarization mode coupling in the Jones and Stokes representations, which we review here, relies on two main properties of the Pauli matrices. One is trace-orthogonality, namely $\mathrm{Tr}\,(\sigma_n\sigma_m) = 2\delta_{n,m}$, and the other is cyclic permutation,

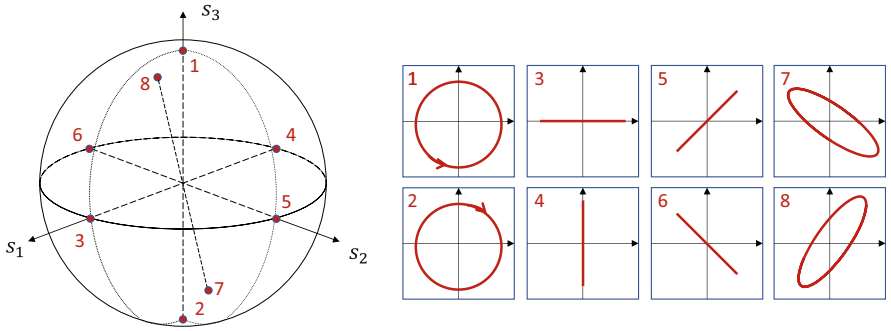

Fig. 5 The Poincaré sphere. The points on the poles represent circular polarizations, while those on the equator represent linear polarizations (x and y on the s_1 axis, $\pm 45^\circ$ on the s_2 axis). All the other points represent generically elliptical polarizations

namely $\sigma_n^2 = \mathbf{I}$ and $\sigma_n\sigma_m = -\sigma_m\sigma_n = i\sigma_k$, where n, m, and k are any cyclic permutation of 1, 2, and 3 [7].

We start by noting that a traceless Hermitian matrix can be expressed as a linear combination of the Pauli matrices with real-valued coefficients. For matrix $\mathbf{B}$ of Eq. (67), this yields

$$\mathbf{B}(z,\omega) = -\beta_1(z,\omega)\sigma_1 - \beta_2(z,\omega)\sigma_2 - \beta_3(z,\omega)\sigma_3 = -\vec{\beta}(z,\omega)\cdot\vec{\sigma}, \tag{74}$$

where the vector $\vec{\beta}(z,\omega)$ is known as the birefringence vector and its z-dependence captures the longitudinal changes of the fiber perturbations. Its components are obtained from the simple relation $\beta_n = -\text{trace}\{\sigma_n\mathbf{B}\}/2$. The solution of the unitary evolution described by Eq. (67) can therefore be expressed as

$$|s(z,\omega)\rangle = \mathbf{U}(z,\omega)|s(0,\omega)\rangle, \tag{75}$$

where $\mathbf{U}$ is a unitary matrix that obeys the evolution equation

$$\frac{\partial\mathbf{U}}{\partial z} = -i\frac{\vec{\beta}(z,\omega)\cdot\vec{\sigma}}{2}\mathbf{U}. \tag{76}$$

If $\vec{\beta}$ is constant along the fiber, the solution of the above is

$$\mathbf{U}(z,\omega) = \exp\left\{-\frac{i}{2}\vec{\beta}(\omega)\cdot\vec{\sigma}z\right\}, \tag{77}$$

However, this situation is of no practical relevance, and Eq. (76) is normally solved numerically using the following relation for each step Δz,

$$\mathbf{U}(z+\Delta z,\omega) \simeq \exp\left\{-\frac{i}{2}\vec{\beta}(z,\omega)\cdot\vec{\sigma}\Delta z\right\}\mathbf{U}(z,\omega), \tag{78}$$

which corresponds to modeling the fiber as a concatenation of birefringent wave plates, each having a different birefringence vector. The formal solution of Eq. (76) can in general be expressed as

$$\mathbf{U}(z,\omega) = \exp\left\{-\frac{i}{2}\vec{r}(z,\omega)\cdot\vec{\sigma}\right\}, \tag{79}$$

where $\vec{r}$ is a real-valued vector, as follows from the fact that a unitary matrix can be written in the form $\mathbf{U} = \exp.\{-i\mathbf{H}/2\}$, where $\mathbf{H}$ is a traceless Hermitian matrix, so that $\mathbf{H} = \vec{r}\cdot\vec{\sigma}$. The significance of the Stokes vector $\vec{r}$ can be seen with reference to the expansion of $\mathbf{U}$ in terms of its eigenvectors [7],

$$\mathbf{U} = e^{-ir/2}\,|\,r\rangle\langle r\,| + e^{ir/2}\,|\,r_\perp\rangle\langle r_\perp\,|\,, \tag{80}$$

where $r = |\vec{r}|$ and $|r\rangle$ and $|r_\perp\rangle$ are orthogonal Jones vectors corresponding to the Stokes vectors $\pm\vec{r}/r$. The effect of propagation is therefore introducing a phase shift r between the components of the input state along the propagation eigenstates $|r\rangle$ and $|r_\perp\rangle$. Equation (80) also suggests that the direction of the birefringence vector $\vec{\beta}$ identifies the local propagation eigenstates, while its length is their differential phase shift per unit propagation distance.

Unitary propagation is most conveniently described in the Stokes space, where the tip of the field Stokes vector draws a trajectory on the Poincaré sphere as it evolves along the fiber. The equation of motion of $\vec{s}$ is derived from that of $|s\rangle$ by evaluating the z-derivative of $\vec{s} = \mathrm{Tr}\left(\vec{\sigma}\ |s\rangle\langle s|\right)$. Use of Eqs. (75) and (76) in this process yields

$$\frac{\partial \vec{s}}{\partial z} = \mathrm{Tr}\left\{\vec{\sigma}\, i\frac{-\left(\vec{\beta}\cdot\vec{\sigma}\right)\left(\vec{s}\cdot\vec{\sigma}\right)+\left(\vec{s}\cdot\vec{\sigma}\right)\left(\vec{\beta}\cdot\vec{\sigma}\right)}{4}\right\} = \vec{\beta}\times\vec{s}, \tag{81}$$

where the second equality is the result of some algebra that relies on the properties of Pauli matrices. This equation shows that the birefringence vector is the local rotation axis in the motion of the Stokes vector and its modulus is the local angular velocity. In the case of constant birefringence, the solution of Eq. (81) is readily obtained in the form $\vec{s}(z,\omega) = \exp\left(z\vec{\beta}\times\right)\vec{s}(0,\omega)$, where $\exp\left(z\vec{\beta}\times\right)$ is the Stokes-space equivalent of (77), and by $\vec{\beta}\times$, we denote the 3×3 matrix implementing the vector product between $\vec{\beta}$ and the vector to which it is applied. In this case, the trajectory of the Stokes vector tip is a circle on the Poincaré sphere. In general, the solution of Eq. (81) can be expressed as $\vec{s}(z,\omega) = R(z,\omega)\vec{s}(0,\omega)$, where R is a 3×3 unitary matrix referred to as the Müller matrix, and the trajectory of the Stokes vector tip loses its circular shape. Similarly to the case of the isomorphic matrix $\mathbf{U}$, a closed-form expression for R is not available in general. However, the relation between $\mathbf{U}$ and R discussed in the case of constant birefringence can be extended to Eq. (79), which in consideration of Eq. (80) yields $R = \exp\left(\vec{r}\times\right)$ (Fig. 6).

Polarization-Mode Dispersion

The frequency dependence of the birefringence vector is responsible for the phenomenon of polarization-mode dispersion (PMD), which takes the form of frequency-dependent random polarization-mode coupling. This phenomenon affects the system operation by introducing a delayed channel response that needs to be equalized. While this could in principle be done in the optical domain, as has been intensively researched in early 2000 [16, 17], when PMD was considered the bottleneck for the deployment of 40 Gb/s direct-detection systems [18], in coherent systems that are used today its equalization is performed digitally in the electrical

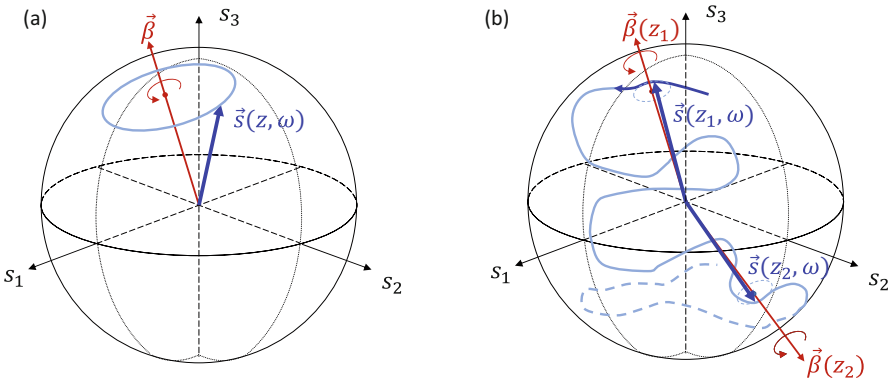

Fig. 6 (**a**) Trajectory of the field Stokes vector $\vec{s}$ on the Poincaré sphere in the presence of constant birefringence. (**b**) Trajectory of the field Stokes vector $\vec{s}$ on the Poincaré sphere in the presence of z-dependent birefringence

domain. In this section, we aim at providing the fundamental concepts and mathematical tools relating to the analysis of PMD in optical fiber links.

Our starting point is that the frequency derivative of $\mathbf{U}$ must be of the same form as Eq. (76). Indeed, by ω-differentiating the unitary condition $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$, one obtain that $\frac{\partial \mathbf{U}}{\partial \omega}\mathbf{U}^\dagger = -\mathbf{U}\frac{\partial \mathbf{U}^\dagger}{\partial \omega}$, which implies that $i\frac{\partial \mathbf{U}}{\partial \omega}\mathbf{U}^\dagger$ is Hermitian. Moreover, since the determinant of $\mathbf{U}$ is unity, as it follows from the fact that its eigenvalues are $\exp(\pm ir)$, the matrix $i\frac{\partial \mathbf{U}}{\partial \omega}\mathbf{U}^\dagger$ is also traceless [7], and therefore it can be expressed as

$$\frac{\partial \mathbf{U}(z,\omega)}{\partial \omega} = -\frac{i}{2}\vec{\tau}(z,\omega)\cdot\vec{\sigma}\,\mathbf{U}(z,\omega). \tag{82}$$

The real-valued vector $\vec{\tau}$ is the famous PMD vector and its physical meaning can be seen in the relevant regime where it can be approximated as frequency-independent within the signal bandwidth. Under this approximation, the solution of Eq. (82) is

$$\mathbf{U}(z,\omega) = \exp\left\{-\frac{i}{2}\vec{\tau}\cdot\vec{\sigma}\ (\omega-\omega_0)\right\}\ \mathbf{U}(z,\omega_0) \tag{83}$$

$$= \left(e^{-i\tau(\omega-\omega_0)/2}\,|\tau\rangle\langle\tau\,| + e^{i\tau(\omega-\omega_0)/2}\,|\,\tau_\perp\rangle\,\langle\tau_\perp|\right)\mathbf{U}(z,\omega_0) \tag{84}$$

$$= e^{-i\tau(\omega-\omega_0)/2}|\tau\rangle\,\langle p\,| + e^{i\tau(\omega-\omega_0)/2}\,|\,\tau_\perp\rangle\langle p_\perp|\,, \tag{85}$$

where the second equality follows from Eq. (80) with $\vec{r} = \vec{\tau}(\omega-\omega_0)$, and where $|p\rangle = \mathbf{U}^\dagger(z,\omega_0)|\tau\rangle$ and $|p_\perp\rangle = \mathbf{U}^\dagger(z,\omega_0)|\tau_\perp\rangle$ are orthogonal Jones vectors that evolve into $|\tau\rangle$ and $|\tau_\perp\rangle$ at frequency $\omega=\omega_0$. To see the effect of PMD, consider an input waveform $f(t)$ transmitted on some polarization state characterized by the frequency-independent Jones vector $|u\rangle$. The effect of Eq. (85) can be described in the time domain as

$$f(t)|u\rangle \rightarrow \langle p|u|\rangle f(t+\tau/2)|\tau\rangle + \langle p_\perp|u|\rangle f(t-\tau/2)|\tau_\perp\rangle, \tag{86}$$

showing that two undistorted replicas polarized along $|\tau\rangle$ and $|\tau_\perp\rangle$ are generated and accumulate a differential delay τ during propagation. As a result, some distortion occurs in the field intensity waveform, which is the way in which PMD affects direct-detection systems. Clearly, if the input field is polarized along $|p\rangle$ or $|p_\perp\rangle$ (in which case $\langle p_\perp|u\rangle = 0$ or $\langle p|u\rangle = 0$, respectively), the effect of propagation is to change the field polarization and introduce a delay, while leaving the field intensity waveform undistorted. The delay τ is known as the *differential group delay* (DGD), while the polarization states $|\tau\rangle$ and $|\tau_\perp\rangle$ corresponding to the Stokes vectors $\pm\vec{\tau}/\tau$ are referred to as fast and slow *principal states of polarization* (PSPs), respectively. To avoid confusion, $|\tau\rangle$ and $|\tau_\perp\rangle$ are sometimes referred to as *output* PSPs, as opposed to $|p\rangle$ and $|p_\perp\rangle$, which are referred to as *input* PSPs.

The first-order picture of PMD described above relies on the assumption that the PMD vector is frequency independent within the signal bandwidth. The accuracy of this assumption can be assessed by characterizing the statistics of the PMD vector, which requires deriving the evolution equation of $\vec{\tau}$. To this end, we equate the derivative of Eq. (76) with respect to ω and that of Eq. (82) with respect to z, with the result

$$\frac{\partial\vec{\tau}}{\partial z}\cdot\vec{\sigma} = \frac{\partial\vec{\beta}}{\partial\omega}\cdot\vec{\sigma} + i\frac{-\left(\vec{\beta}\cdot\vec{\sigma}\right)\left(\vec{\tau}\cdot\vec{\sigma}\right) + \left(\vec{\tau}\cdot\vec{\sigma}\right)\left(\vec{\beta}\cdot\vec{\sigma}\right)}{2}, \tag{87}$$

which by tracing out the Pauli matrices yields the PMD dynamic equation [7],

$$\frac{\partial\vec{\tau}}{\partial z} = \frac{\partial\vec{\beta}}{\partial\omega} + \vec{\beta}\times\vec{\tau}. \tag{88}$$

Since $\vec{\beta}$ is random in nature, Eq. (88) is a stochastic equation. In particular, since the correlation length of the birefringence vector is of the order of meters to hundreds of meters [19], in fiber links of hundreds to thousands of kilometers the PMD vectors evolves like a 3-D Brownian motion, and therefore its three components can be modeled as independent and identically distributed zero-mean Gaussian variables, while the DGD (its length) has a Maxwellian probability density function [20]. The mean DGD, which turns out to be frequency independent, grows with the square root of propagation distance according to $\langle\tau(z)\rangle = \kappa_{\mathrm{PMD}}\sqrt{z}$, where the angled brackets denote ensemble averaging, and κ_{PMD} is the so-called PMD coefficient, whose value ranges between the order of 0.01 $\mathrm{ps}/\sqrt{\mathrm{km}}$ in low-PMD fibers to the order of 0.1 $\mathrm{ps}/\sqrt{\mathrm{km}}$ in deployed vintage fibers.

The answer to the original concern relating to the accuracy of the first-order PMD approximation comes from the study of the two-frequency correlation function of the PMD vector. This is typically performed under the assumption that $\vec{\beta}(z,\omega)$ can be expanded to first order with respect to ω in the band of interest, that is $\vec{\beta}(z,\ \omega+\Omega) \simeq \Omega\ \frac{\partial\vec{\beta}}{\partial\omega}(z,\omega)$, with the result [21–23]

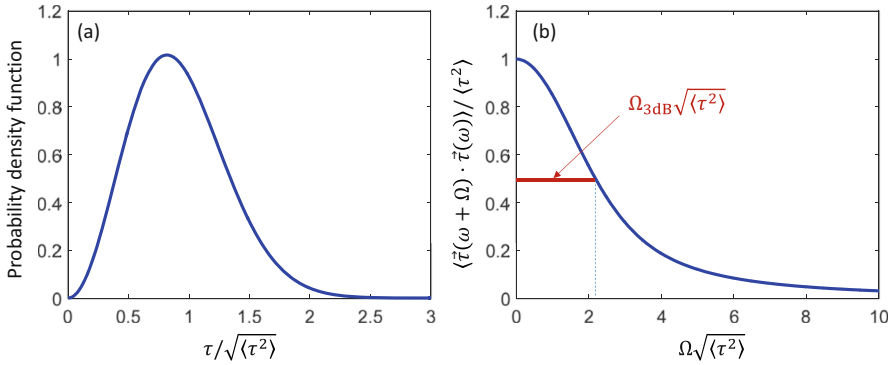

Fig. 7 (**a**) The probability density function of the DGD normalized to its root-mean-square values. (**b**) Normalized autocorrelation function of the PMD vector

$$\left\langle \vec{\tau}(z,\omega)\cdot\vec{\tau}(z,\omega+\Omega)\right\rangle = 3\frac{1-e^{-\frac{\Omega^2\langle\tau^2\rangle}{3}}}{\Omega^2} \tag{89}$$

where $\langle\tau^2\rangle = (8/3\pi)\,\langle\tau\rangle^2$. The PMD vector autocorrelation function (ACF) is plotted along with the Maxwellian probability density function of the DGD in Fig. 7. The PMD bandwidth is typically defined as the frequency at which the PMD vector ACF reduced to one-half of its maximum and is equal to $B_{\mathrm{PMD}} = \Omega_{3\mathrm{dB}}/2\pi \simeq 0.347/\sqrt{\langle\tau^2\rangle}$, as illustrated in Fig. 7b. As a reference, the PMD bandwidth of a 1000-km link of a low-PMD fiber with $\kappa_{\mathrm{PMD}} = 0.01$ ps/$\sqrt{\mathrm{km}}$ is $B_{\mathrm{PMD}} \simeq 1$ THz, and it reduces to $B_{\mathrm{PMD}} \simeq 100$ GHz for $\kappa_{\mathrm{PMD}} = 0.1$ ps/$\sqrt{\mathrm{km}}$. These values suggest that the first-order PMD approximation is very accurate within the bandwidth of WDM channels, typically of the order of a few tens of GHz. It should be noted that the situation is substantially different in the case of space-division multiplexed transmission systems [24].

We conclude this section by noting that the magnitude of the fiber PMD (the mean DGD) determines the complexity of the digital signal processing that is required for its equalization [25, 26]. On the other hand, since PMD is a unitary propagation effect, no fundamental loss in terms of system information capacity is implied. That is contrary to the case of polarization-dependent loss, which is discussed in the following section.

Polarization-Dependent Loss

Polarization-dependent loss affects propagation by imposing different losses on different field polarizations and is fundamentally unavoidable in the presence of optical birefringence. To see why this is the case, we expand on the example of Eq. (66) by introducing an imaginary part in the perturbation tensor $\widetilde{\chi}_P$, which yields

$$|\widetilde{L}\rangle = \frac{i}{2}\begin{bmatrix}\delta\beta & 0\\ 0 & -\delta\beta\end{bmatrix} + \frac{1}{2}\begin{bmatrix}\delta\alpha & 0\\ 0 & -\delta\alpha\end{bmatrix}, \tag{90}$$

where the second matrix describes differential loss in the x and y polarizations. The need for this addition arises from the causality of the fiberglass response, which implies that the real and imaginary components of $\widetilde{\chi}_P$ are related to each other through the Kramers-Kronig relations [13]. Similar to the case of polarization-mode coupling, if the x and y polarizations are not the local polarization eigenstates, then the differential loss matrix in Eq. (90) must be replaced with $\mathbf{A} = \mathbf{U}\text{diag}\{\delta\alpha, -\delta\alpha\}\mathbf{U}^{\dagger}$, which is also Hermitian and traceless, so that

$$\frac{\partial|\widetilde{E}(z,\omega-\omega_0)\rangle}{\partial z} = \frac{\mathbf{A}+i\mathbf{B}}{2}|\widetilde{E}(z,\omega-\omega_0)\rangle, \tag{91}$$

and $\mathbf{A} = \vec{\alpha}(z,\omega)\cdot\vec{\sigma}$. The vector $\vec{\alpha}$ describes the local dependence of loss on the signal polarization.

Contrary to the case of pure birefringence, in the presence of PDL, propagation is not unitary, and since the field power spectral density is given by

$$\langle\widetilde{E}(z,\omega-\omega_0)\,|\,\widetilde{E}(z,\omega-\omega_0)\rangle = \langle\widetilde{E}(0,\omega-\omega_0)\,|\,\mathbf{T}^{\dagger}\mathbf{T}\,|\,\widetilde{E}(0,\omega-\omega_0)\rangle, \tag{92}$$

where $\mathbf{T}(z,\omega)$ is the system transfer matrix describing propagation from 0 to z, the relevant matrix in the study of PDL is $\mathbf{T}^{\dagger}\mathbf{T}$. This is a Hermitian matrix and therefore can be expressed as

$$\mathbf{T}^{\dagger}\mathbf{T} = \gamma_0\left(\mathbf{I} + \vec{\Gamma}\cdot\vec{\sigma}\right), \tag{93}$$

where γ_0 accounts for the polarization-averaged loss, and $\vec{\Gamma}$ is the so-called PDL vector [27]. Its physical meaning appears from noting that the intensities of polarization states with Stokes vectors parallel and antiparallel to $\vec{\Gamma}$ are attenuated by $\gamma_0(1+\Gamma)$ and $\gamma_0(1-\Gamma)$, respectively, with $\Gamma = |\vec{\Gamma}|$, which also implies that $\Gamma \leq 1$. The attenuation of the other polarization states ranges between these two values. The PDL vector obeys the evolution equation [28]

$$\frac{\partial\vec{\Gamma}}{\partial z} = \vec{\alpha} - \vec{\Gamma}\left(\vec{\alpha}\cdot\vec{\Gamma}\right) + \vec{\beta}\times\vec{\Gamma}. \tag{94}$$

The last term on the right-hand side of the equality is responsible for rotating $\vec{\Gamma}$ as it accumulates along the fiber, and therefore, it can be neglected with no consequences on the resulting statistics of the PDL vector. In turn, the second term is negligible in the regime of small PDL, as in this regime, it is small compared to the first one. The PDL vector can therefore be approximated as $\vec{\Gamma}(z,\omega) \simeq \int_0^z \vec{\alpha}(z',\omega)\,\mathrm{d}z'$, which clarifies why $\vec{\alpha}$ is often referred to as the local PDL vector.

It is worth pointing out that the PDL of a fiber-optic link often arises from fiber splices and system components rather than from the fiber itself. Therefore, it is useful to recall the concatenation rule giving the PDL vector of two cascaded PDL elements characterized by the PDL vectors $\vec{\Gamma}_1$ and $\vec{\Gamma}_2$ [27],

$$\begin{aligned}\vec{\Gamma} &= \frac{\sqrt{1-\left|\vec{\Gamma}_2\right|^2}}{1+\vec{\Gamma}_1\cdot\vec{\Gamma}_2}\vec{\Gamma}_1 + \frac{1+\vec{\Gamma}_1\cdot\vec{\Gamma}_2\left(1-\sqrt{1-\left|\vec{\Gamma}_2\right|^2}\right)\Big/\left|\vec{\Gamma}_2\right|^2}{1+\vec{\Gamma}_1\cdot\vec{\Gamma}_2}\vec{\Gamma}_2 \\ &\simeq \vec{\Gamma}_1+\vec{\Gamma}_2,\end{aligned} \tag{95}$$

where the second equality holds for small PDL.

A common figure of merit for the PDL of a fiber-optic link is the power ratio in dB between the least and most attenuated polarization states,

$$\rho = 10\log_{10}\left(\frac{1+\Gamma}{1-\Gamma}\right), \tag{96}$$

and typical values of ρ for communication systems using single-mode fibers rarely exceed a few dB. Clearly ρ is random in nature and its probability density function depends on the distribution of the PDL sources along the link. When these can be approximated as continuously distributed, if the mean PDL is sufficiently small, then the probability density function of ρ is well approximated by a Maxwellian distribution [29]; otherwise, an exact yet less simple expression for the probability density function is available in Ref. [30]. Conversely, when the PDL is dominated by a finite number of contributions, the distribution of ρ may be considerably different. This scenario is relevant in the context of long fiber routes with a few dominant PDL elements connected by long fiber spans with negligible PDL [31], as well as in the context of short fiber links, like data-center interconnects, or intra-data-center links with lumped PDL sources. The exact probability density function of ρ for these cases has been derived only recently [32]. It should be noted that since the PDL introduced by components typically has a negligible frequency dependence, the overall frequency dependence of the PDL vector is determined primarily by the PMD of the fiber spans connecting the PDL elements [33].

An important aspect of PDL is that it impairs the performance of a fiber communication system at a fundamental level by reducing its information capacity. This reduction is due to the fact that, contrary to PMD, PDL is a non-unitary propagation effect, and the mathematical framework for studying the impact of PDL in modern coherent systems has been introduced in Ref. [34]. As argued in Ref. [34], the mechanism through which PDL affects the system capacity is by unbalancing (in a random fashion) the signal and noise powers in any pair of orthogonal polarizations. This can be seen with reference to the link schematic of Fig. 4, by looking at the noise accumulated within some narrow band B, so that propagation from the h-th

amplifier to the receiver can be described in the time domain by means of the transmission matrix $\mathbf{T}_h$. Under this assumption, the noise impinging upon the receiver is expressible as $|n\rangle = \sum_{h=1}^{M} \mathbf{T}_h \, |n_h\rangle$, where $|n_h\rangle$ is the ASE noise introduced by the h-th amplifier. The noise coherency matrix, that in the absence of PDL would be $\mathbf{Q} = \mathrm{E}[|n\rangle \langle n|] = \frac{P_{\mathrm{ASE}}}{2}\mathbf{I}$, is given by $\mathbf{Q} = \frac{P_{\mathrm{ASE}}}{2N_s} \sum_{h=1}^{N_s} \mathbf{T}_h \mathbf{T}_h^{\dagger}$, as follows from the fact that the noise contributions added by the individual amplifiers are independent of each other and identically distributed, namely $\mathrm{E}[|n_h\rangle \langle n_k|] = \frac{P_{\mathrm{ASE}}}{2N_s}\mathbf{I}\delta_{h,k}$, where by the symbol $\delta_{h,k}$, we denote the Kronecker delta. In this case, the system capacity is given by $C = B[\log_2(1 + S\lambda_1) + \log_2(1 + S\lambda_2)]$, where S is the average power transmitted in each polarization signal, and λ_1 and λ_2 are the two eigenvalues of the (Hermitian) matrix $\mathbf{Q}^{-1/2}\mathbf{T}\mathbf{T}^{\dagger}\mathbf{Q}^{-1/2}$. The capacity loss follows from the fact that C is maximized for $\lambda_1 = \lambda_2$, as would be the case if $\mathbf{T}$ where unitary and $\mathbf{Q}$ proportional to the identity, while in general $\lambda_1 \neq \lambda_2$ in the presence of PDL. It is instructive to note that since $\mathbf{Q}$ is Hermitian by definition, it can be expressed as $\mathbf{Q} = \gamma_0' \left(\mathbf{I} + \vec{\Gamma}' \cdot \vec{\sigma}\right)$, where γ_0' describes the polarization-averaged noise power and $\vec{\Gamma}'$ is a real-valued vector accounting for the noise non-circularity [34].

Nonlinear Propagation

We conclude this chapter with an introduction to the study of nonlinear distortions experienced by a signal propagating in the fiber-optic channel. These distortions are caused by the dependence of the glass refractive index on optical power, and they become relevant at high levels of the propagating signal power. The interplay between linear and nonlinear propagation effects is highly nontrivial and is a unique characteristic of the fiber-optic channel.

The physical origin of nonlinear propagation is the nonlinear polarization vector, which consists of two contributions, one instantaneous, due to the electronic response of glass, and another, slower and delayed, associated with the response of the nuclei, which are much heavier than the electrons and therefore react more slowly to changes in the electric field. In what follows, we restrict ourselves to the instantaneous contribution, which produces the nonlinear polarization vector $\Delta\vec{p} = \varepsilon_0 \chi^{(3)} |\vec{e}|^2 \vec{e}$. The corresponding perturbation term is evaluated in the time domain, by first using Eq. (36), which yields the following expression for the positive-frequency content of $\Delta\vec{p}$,

$$\begin{aligned} \Delta\vec{P}_{\omega_0} = \frac{\varepsilon_0 \chi^{(3)} |F|^2 F}{8\mathcal{N}^3} \Big[&\Big(3|E_x|^2 E_x + 2|E_y|^2 E_x + E_y^2 E_x^*\Big)\hat{x} \\ &+ \Big(2|E_x|^2 E_y + 3|E_y|^2 E_y + E_x^2 E_y^*\Big)\hat{y}\Big], \end{aligned} \tag{97}$$

and then by entering this expression into Eq. (49). This procedure yields the result

$$|L\rangle = i\gamma\left[\left(|E_x|^2E_x + \frac{2}{3}\ |E_y|^2E_x + \frac{1}{3}E_y^2E_x^*)\hat{x} + \left(\frac{2}{3}|E_x|^2E_y + |E_y|^2E_y + \frac{1}{3}E_x^2E_y^*\right)\hat{y}\right], \tag{98}$$

where γ is the nonlinearity coefficient defined as

$$\gamma = \frac{\omega_0}{2\mathcal{N}}\ \frac{\varepsilon_0 3\chi^{(3)}}{8\mathcal{N}^3}\int |F|^4 \mathrm{d}x\mathrm{d}y = \frac{3\omega_0\varepsilon_0 Z_0^2\chi^{(3)}}{4n_{\mathrm{eff}}^2 A_{\mathrm{eff}}}. \tag{99}$$

Typically, in standard single-mode fibers, $\gamma \simeq 1.3\ \mathrm{W}^{-1}\ \mathrm{km}^{-1}$. It is also customary to express γ as

$$\gamma = \frac{n_2\omega_0}{cA_{\mathrm{eff}}}, \tag{100}$$

where n_2 is the so-called nonlinear-index coefficient,

$$n_2 = \frac{3\chi^{(3)}}{8n_{\mathrm{eff}}}\ \frac{2Z_0}{n_{\mathrm{eff}}} = \frac{3\chi^{(3)}}{8n_{\mathrm{eff}}}2Z_{\mathrm{eff}}, \tag{101}$$

whose physical significance is discussed in the following section.

As an intermediate summary, we note that the equations describing propagation in the presence of loss, chromatic dispersion, and instantaneous Kerr nonlinearity are often cast as follows,

$$\frac{\partial E_x}{\partial z} = -\frac{\alpha}{2}E_x - i\frac{\beta_2}{2}\frac{\partial^2 E_x}{\partial t^2} + i\gamma\left(|E_x|^2E_x + \frac{2}{3}|E_y|^2E_x + \frac{1}{3}E_y^2E_x^*\right), \tag{102}$$

$$\frac{\partial E_y}{\partial z} = -\frac{\alpha}{2}E_y - i\frac{\beta_2}{2}\frac{\partial^2 E_y}{\partial t^2} + i\gamma\left(|E_y|^2E_y + \frac{2}{3}|E_x|^2E_y + \frac{1}{3}E_x^2E_y^*\right). \tag{103}$$

Clearly, in the absence of polarization effects, the coupling between the x and y components of the field vector occurs only through the nonlinear terms, and it can be easily seen that if only one of the two components is excited at the fiber input, the other component is absent during propagation. In this situation, the equation describing nonlinear propagation of the nonzero field component simplifies to the following form

$$\frac{\partial E}{\partial z} = -\frac{\alpha}{2}E - i\frac{\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + i\gamma|E|^2E, \tag{104}$$

which is known as the nonlinear Schrödinger equation (NLSE). This denomination follows from the equivalence existing between the linear part of Eq. (104) (with $\alpha = 0$) and the Schrödinger equation describing the evolution of a quantum system. By analogy, Eqs. (102) and (103) are referred to as *coupled* NLSEs. It is instructive to point out that the nonlinear term of Eq. (104) can also be expressed as $i\frac{\omega_0}{c}n_{\mathrm{NL}}E$,

where $n_{\mathrm{NL}} = n_2|E|^2/A_{\mathrm{eff}} = n_2 I$ has the meaning of a nonlinear refractive index, whose magnitude is proportional to the average power density $I = |E|^2/A_{\mathrm{eff}}$. This clarifies the definition of the nonlinear-refractive index coefficient n_2 introduced in Eq. (101).

The simpler form of the scalar NLSE facilitates developing a physical intuition of various nonlinear-propagation-related phenomena, whose analysis becomes less straightforward when using the coupled NLSEs. It should be noted, however, that even in the scalar case, the NLSE does not in general admit closed-form solutions. These can be only found for specific settings of the parameters involved and/or for specific input waveforms.

One relevant example is that of self-phase modulation, which can be conveniently studied by neglecting chromatic dispersion (i.e., by setting $\beta_2 = 0$). In this case, defining $E = |E| \exp(i\phi)$, one finds

$$E(z,t) = E(0,t)e^{-\frac{\alpha}{2}z}e^{i\gamma|E(0,t)|^2 L_{\mathrm{eff}}(z)}, \tag{105}$$

$$L_{\mathrm{eff}}(z) = \frac{1 - e^{-\alpha z}}{\alpha}. \tag{106}$$

The quantity $L_{\mathrm{eff}}(z)$ is known as the effective length and reduces to z in the absence of loss, whereas it saturates to $1/\alpha$ for propagation distances significantly exceeding $1/\alpha$. Equation (105) shows that the effect of the nonlinearity is to impose a phase modulation $\phi_{\mathrm{SPM}}(z, t) = \gamma|E(0, t)|^2 L_{\mathrm{eff}}(z)$ that is proportional to the instantaneous intensity of the propagating waveform, whose time dependence is responsible for an instantaneous frequency shift $\delta\omega(t) = -\mathrm{d}\phi_{\mathrm{SPM}}(z, t)/\mathrm{d}t$. For a pulse-like waveform, SPM produces a negative (red-) shift in the leading edge of the pulse and a positive (blue-) shift in the trailing edge of the pulse, which imply spectral broadening.

Although the above description is only rigorous in the absence of group-velocity dispersion, it still provides an intuitive understanding of how the nonlinear term also affects propagation in the presence of group-velocity dispersion, yet only through a sufficiently short fiber section, where the two effects may be treated as if they occurred one after the other. Therefore, owing to group-velocity dispersion, the blue- and red-shifted pulse edges experience a change in their group velocities. It is easy to see that in the regime of anomalous dispersion (where $\beta_2 < 0$), the leading and trailing edges of the pulse experience a reduction and an increase in their group velocities, respectively, yielding to some pulse compression. Beyond that point, dispersion causes the pulse to broaden again, but this time with the leading edge consisting of the blue-shifted frequencies (traveling faster in the case of anomalous dispersion) and the trailing edge consisting of the red-shifted frequencies (as they travel slower in the case of anomalous dispersion). In the absence of loss, the two mechanisms may balance each other perfectly, as is the case for the following waveform,

$$E(z,t) = \sqrt{\frac{|\beta_2|}{\gamma T_0^2}}\mathrm{sech}\left(\frac{t}{T_0}\right)e^{-i\frac{\beta_2}{2T_0^2}z}, \tag{107}$$

which is a special solution of the NLSE, characterized by a z-independent intensity profile (both in the time and frequency domains). The waveform in Eq. (107) is known as the *fundamental soliton*, and it belongs to the class of the soliton solutions of the NLSE. Indeed, in the case of ideally distributed amplification ($\alpha = 0$), the NLSE is integrable and admits closed-form solutions that can be obtained by means of the inverse scattering method [35]. Defining the nonlinear length as the propagation length for which $\phi_{\mathrm{SPM}} = 1$ rad, namely $L_{\mathrm{NL}}^{-1} = \gamma \max_t \left\{ |E(0,\ t)|^2 \right\}$, $L_{\mathrm{NL}} = L_D$ for the fundamental soliton. ($L_D = T_0^2/|\beta_2|$ is the dispersion length introduced in the section "Chromatic Dispersion"). That is, the balancing of chromatic dispersion and self-phase modulation results in the formation of the fundamental soliton when the corresponding characteristic lengths are matched.

Fiber-optic solitons were intensely researched in the 1990s [36, 37], when they were seen as a potential solution for long-haul transmission in intensity-modulation and direct-detection systems, owing to their immunity to the effect of the fiber nonlinearity and dispersion. However, by the end of the 1990s, this approach was abandoned primarily due to the excessive timing jitter caused by the Gordon-Haus effect [38]. Methods of compensating for it were proposed, e.g., Ref. [37], but they were never implemented in commercial systems. Instead, in the late 1990s, commercial systems relied on the dispersion-managed soliton paradigm [39], until the arrival of the coherent revolution [11]. The attempt to treat the fiber nonlinearity as a constructive factor, rather than as an impairment, has been recently revived in the context of modern coherent systems based on the modulation of the nonlinear Fourier spectrum of the transmitted signals, of which solitons are part. This novel transmission paradigm has been researched since the beginning of the past decade [40].

The Manakov Equation

The consideration of random polarization-mode coupling introduces a major simplification into the nonlinear propagation equations for the field vector. The reason for this simplification lies in the fact that random polarization coupling scrambles the instantaneous state of polarization of the propagating field on a length-scale much shorter than that on which nonlinear effects become appreciable. In fact, the length-scale of nonlinear effects is rarely shorter than a few tens of kilometers, while the instantaneous state of polarization of the propagating field covers the Poincaré sphere uniformly over propagation distances of the order of centimeters to meters. This simple observation suggests that the nonlinear terms of the coupled NLSEs can be averaged with respect to the random orientation of the instantaneous polarization state, under the assumption that it is uniformly distributed over the Poincaré sphere. Based on symmetry considerations, the nonlinear terms cannot have components orthogonal to the field vector itself and therefore must reduce to the simple form $i\kappa\gamma\ \langle E|E\rangle|E\rangle$. This can be seen conveniently by expressing $|L\rangle$ as [41]

$$|L\rangle = i\gamma\langle E\,|\,E\rangle|E\rangle - i\gamma\frac{S_3}{3}\sigma_3|E\rangle, \tag{108}$$

where $S_3 = 2\mathrm{Im}\{E_x^* E_y\}$ is the third component of the instantaneous Stokes vector associated with the propagating field. Owing to random polarization-mode coupling, S_3 covers uniformly the values ranging between $-\langle E|E\rangle$ and $\langle E|E\rangle$ over a short propagation distance, where the optical power $\langle E|E\rangle$ can be assumed to remain constant. The component of $|L\rangle$ along $|E\rangle$ can therefore be obtained by averaging $\langle E|L\rangle$ over the uniform distribution of S_3, and then dividing by $\langle E|E\rangle^{1/2}$, namely

$$\frac{1}{\langle E|E\rangle^{1/2}} \int_{-\langle E|E\rangle}^{\langle E|E\rangle} \frac{1}{2\langle E|E\rangle} i\gamma \left(\langle E|E\rangle^2 - \frac{S_3^2}{3} \right) \mathrm{d}S_3 = i\frac{8}{9} \langle E|E\rangle^{3/2} \tag{109}$$

which when multiplied by the unit vector $|E\rangle/\langle E|E\rangle^{1/2}$ yields the known nonlinear term $i\frac{8}{9}\langle E|E\rangle|E\rangle$. On the other hand, one can see that the projection of $|L\rangle$ on the field vector $|E_\perp\rangle = \left[E_y^*, -E_x^*\right]^t$, which is orthogonal to $|E\rangle$, is proportional to $E_x^* E_y^3 - E_x^3 E_y^* + E_x E_y |E_x|^2 - E_x E_y |E_y|^2$. Since E_x and E_y have independent and random phases, as a result of random polarization-mode coupling, they are all averaged out over a short propagation distance. Therefore, the equation describing nonlinear propagation in the presence of random polarization-mode coupling reduces to the simple form

$$\frac{\partial|E\rangle}{\partial z} = -\frac{\alpha}{2}|E\rangle - i\frac{\beta_2}{2}\frac{\partial^2|E\rangle}{\partial t^2} + i\frac{8}{9}\gamma\langle E|E\rangle|E\rangle. \tag{110}$$

This equation is known as the Manakov equation [42], and it was first introduced to study nonlinear propagation in optical fibers by Wai and Menyuk in 1991 [43]. More recently, generalized versions of the Manakov equation have been shown to describe nonlinear multiple-mode propagation in fibers for space-division multiplexing [44–46]. It is important to note that, contrary to the coupled NLSEs, its form is independent of the field representation basis. Note also that Eq. (110) is similar to the scalar NLSE in various ways. One similarity is the fact that the study of self-phase modulation can be extended by simply replacing E with $|E\rangle$ and $|E|^2$ with $\langle E|E\rangle = |E_x|^2 + |E_y|^2$ in Eqs. (105), which shows that each polarization component is phase-modulated also by the other polarization component, a phenomenon referred to as cross-polarization modulation. Another similarity is that the Manakov equation is also integrable in the absence of loss and therefore admits soliton solutions.

The Manakov equation can be extended to account for the effect of PMD by adding a linear term which results from expanding $\vec{\beta}(z,\ \omega) \simeq \vec{\beta}_\omega(z)(\omega - \omega_0)$ in the frequency-domain propagation equation, as discussed in the introduction of Eq. (68). This yields the following equation,

$$\frac{\partial|E\rangle}{\partial z} = -\frac{\alpha}{2}|E\rangle - i\frac{\beta_2}{2}\frac{\partial^2|E\rangle}{\partial t^2} + i\frac{8}{9}\gamma\langle E|E\rangle|E\rangle + \frac{\vec{\beta}_\omega \cdot \vec{\sigma}}{2}\frac{\partial|E\rangle}{\partial t}, \tag{111}$$

which is known as the Manakov-PMD equation [47]. For the sake of completeness, we point out that both Eqs. (110) and (111) are missing the random-coupling term.

However, while the effect of this term is very important and it underpins the derivation of the Manakov equation itself, its presence in the Manakov equation is immaterial, as it can be removed by means of a z-dependent change of reference, and therefore, it is omitted. The use of suitable reference frames is common practice in the modeling of polarization effects [48].

Cross-Phase Modulation

When multiple information-carrying signals are transmitted in different frequency channels, various nonlinear interchannel interference effects may occur. In this section, we introduce the simplest and the most important of these effects, known as cross-phase modulation.

We start by considering two co-propagating channels $E_0(t)$ and $E_1(t)$ separated in frequency by Ω so that the overall complex envelope at the fiber input can be expressed as

$$E(0,t) = E_0(t) + e^{-i\Omega t}E_1(t), \tag{112}$$

where we considered scalar fields for simplicity. In order to describe the central idea, we express the solution of the propagation equation in the absence of dispersion and loss, to be consistent with Eq. (105)

$$E(z,t) = \left[E_0(t) + e^{-i\Omega t}E_1(t)\right]e^{i\gamma|E_0(t)+\exp(-i\Omega t)E_1(t)|^2 z}. \tag{113}$$

Under the assumption of weak nonlinearity the above can be approximated by expanding the exponential function to first order with respect to γ,

$$\begin{aligned} E(z,t) &\simeq E_0(t)\left[1 + i\gamma\left(|E_0|^2 + 2|E_1|^2\right)z\right] + E_1(t)\left[1 + i\gamma\left(|E_1|^2 + 2|E_0|^2\right)z\right]e^{-i\Omega t} \\ &\quad + i\gamma E_0^2 E_1^* e^{i\Omega t} + i\gamma E_0^* E_1^2 e^{-i2\Omega t} \\ &\simeq E_0(t)e^{i\gamma\left(|E_0|^2+2|E_1|^2\right)z} + E_1(t)e^{i\gamma\left(2|E_0|^2+|E_1|^2\right)z}e^{-i\Omega t} + i\gamma E_0^2 E_1^* e^{i\Omega t} + i\gamma E_0^* E_1^2 e^{-i2\Omega t}. \end{aligned} \tag{114}$$

The first two terms on the right-hand side of the second equality show that each signal is phase-modulated not only by itself but also by the other signal, with twice as large a coefficient. The phase modulation imposed by one signal onto the other is referred to as *cross-phase* modulation (XPM). The third and fourth terms on the right-hand side of the same equality reside at different frequencies, and they are referred to as *four-wave mixing* (FWM) products. They are typically responsible for a smaller perturbation to the transmitted signal, compared to that produced by XPM, but their study is not covered in this chapter.

We now move to see how XPM is included in the Manakov equation. To this end, we replace the field vector $|E(z,\, t)\rangle$ with the ansatz $|E(z,\, t)\rangle = |E_0(z,\, t)\rangle + |E_1(z,\, t)\rangle$

$\exp(-i\Omega t)$, and extract two equations, one for $|E_0(z,\, t)\rangle$ and another for $|E_1(z,\, t)\rangle$, by isolating the terms at zero-frequency and at frequency Ω. This procedure yields the following pair of coupled equation

$$\frac{\partial|E_0\rangle}{\partial z} = -\frac{\alpha}{2}|E_0\rangle - i\frac{\beta_2}{2}\frac{\partial^2|E_0\rangle}{\partial t^2} + i\frac{8}{9}\gamma(\langle E_0|E_0\rangle + \langle E_1|E_1\rangle + |E_1\rangle\langle E_1|)|E_0\rangle \tag{115}$$

$$\begin{aligned}\frac{\partial|E_1\rangle}{\partial z} = &-\frac{\alpha}{2}|E_1\rangle + i\Omega^2|E_1\rangle - \beta_2\Omega\frac{\partial|E_1\rangle}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2|E_1\rangle}{\partial t^2} \\ &+ i\frac{8}{9}\gamma(\langle E_1|E_1\rangle + \langle E_0|E_0\rangle + |E_0\rangle\langle E_0|)|E_1\rangle.\end{aligned} \tag{116}$$

The terms proportional to $\langle E_1|E_1\rangle|E_0\rangle$ and $\langle E_0|E_0\rangle|E_1\rangle$ describe XPM, as can be seen by comparison with the scalar case. On the other hand, the terms $(|E_j\rangle\langle E_j|)|E_i\rangle$ are absent in the scalar analysis and describe polarization rotation of one field vector about the other. This can be seen by considering the Stokes representation of the propagation equations. In fact, by expressing $|E_j\rangle\langle E_j| = \frac{1}{2}\left(\mathbf{I} + \vec{S}_j \cdot \vec{\sigma}\right)$, one can see that $\vec{S}_j$ plays the role of a local birefringent vector, and therefore, the term $i(|E_j\rangle\langle E_j|)|E_i\rangle$, which becomes $\vec{S}_j \times \vec{S}_i$ in the Stokes representation, describes a rotation of $\vec{S}_i$ about $\vec{S}_j$. The rotation takes place while the two signals overlap in time, and for pulsed waveforms, this occurs for the time that the faster pulse takes to surpass the slower pulse. This dynamic is often referred to as pulse collision [49, 50]. The third term on the right-hand side of Eq. (116) accounts for the difference in group velocity between the two signals, and $\beta_2\Omega$ is the differential delay per unit propagation distance between them. Finally, the second term on the right-hand side of Eq. (116) is responsible for an immaterial phase term $\exp(-i\Omega^2 z)$.

Nonlinear Interference Noise in the Context of Digital Coherent Systems

Modern fiber communication links rely on coherent transmission and digital compensation of the main propagation impairments. In particular, contrary to the case of dispersion-managed systems, where chromatic dispersion is compensated for periodically along the link in the optical domain, in digital coherent systems, chromatic dispersion is compensated only at the receiver in the digital domain. The performance of these systems is therefore limited mainly by ASE noise and nonlinear distortions, and its analytical study is hindered primarily by the nonlinear signal propagation dynamics. However, in the relevant range of system parameters, the optical signal propagates in the so-called pseudo-linear regime [12], where the nonlinearity – which otherwise would be difficult to treat analytically – can be treated as a perturbation. This section provides an introduction to this analysis.

The main idea is that the signal propagates as if it were affected by loss and chromatic dispersion only, and the nonlinearity introduces an additive perturbation term that is treated as noise – the *nonlinear interference noise* (NLIN) – which adds to the amplification noise. The starting point is the Manakov equation written in the following form

$$\frac{\partial|E\rangle}{\partial z} = -\frac{a(z)}{2}|E\rangle - i\frac{\beta_2}{2}\frac{\partial^2|E\rangle}{\partial t^2} + i\frac{8}{9}\gamma\langle E|E\rangle|E\rangle, \tag{117}$$

where $a(z)$ describes loss and amplification along the link. The loss/gain term can be removed by redefining the field vector as $\sqrt{f(z)}|E(z,t)\rangle$, where $f(z) = \exp\left(-\int_0^z a(z')\ \mathrm{d}z'\right)$ is the loss/gain profile function, typically modeled with the saw-tooth profile of Fig. 4. This changes the Manakov equation into

$$\frac{\partial|E\rangle}{\partial z} = -i\frac{\beta_2}{2}\frac{\partial^2|E\rangle}{\partial t^2} + i\frac{8}{9}\gamma f(z)\langle E|E\rangle|E\rangle. \tag{118}$$

We then express the solution of the above as $|E(z,\ t)\rangle = |E^{(0)}(z,\ t)\rangle + |\Delta\ E(z,\ t)\rangle$, where $|E^{(0)}(z,\ t)\rangle$ is the solution for $\gamma = 0$ given by

$$|E^{(0)}(z,t)\rangle = \exp\left\{-iz\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\right\}|E(0,t)\rangle, \tag{119}$$

and the perturbation term is obtained by solving the (linear) equation

$$\frac{\partial|\Delta E\rangle}{\partial z} = -i\frac{\beta_2}{2}\frac{\partial^2|\Delta E\rangle}{\partial t^2} + i\frac{8}{9}\gamma f(z)\langle E^{(0)}\,|\,E^{(0)}\rangle|E^{(0)}\rangle, \tag{120}$$

where the last term on the right-hand side of the equality is the result of replacing $|E\rangle$ with the zero-th order solution $|E^{(0)}\rangle$ in the nonlinear term of the Manakov equation [12, 9]. The solution of (120) is given by

$$|\Delta E(z,t)\rangle = i\frac{8}{9}\gamma\int_0^z e^{-i(z-z')\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}} f(z')\langle E^{(0)}(z',t)\,|\,E^{(0)}(z',t)\rangle|E^{(0)}(z',t)\rangle \mathrm{d}z'. \tag{121}$$

To move forward with the analysis and describe the procedure for the evaluation of the NLIN power, we consider the transmission of two wavelength-division multiplexed channels spaced by Ω and assume that the input waveforms are pulse-code modulated signals, so that $|E(0,\ t)\rangle = |E_0(t)\rangle + \exp\ (-i\Omega t)|E_1(t)\rangle$, with

$$|E_0(t)\rangle = \sum_n \vec{a}_n g_0(t - nT) \tag{122}$$

$$|E_1(t)\rangle = \sum_n \vec{b}_n g_0(t - nT). \tag{123}$$

We refer to the channel described by $|E_0\rangle$ as the channel of interest (COI) and to that described by $|E_1\rangle$ as the interfering channel (IC). Here by $\vec{a}_n$ and $\vec{b}_n$, we denote two-dimensional column vectors whose elements are the constellation symbols transmitted in the two field polarizations during the n-th symbol-time slot. The symbols are generated periodically in time with a period T, and they are used to modulate the amplitude of delayed replicas of the fundamental pulse waveform $g_0(t)$. The solution of the linear propagation equation can be expressed in this case as

$$|E^{(0)}(z,t)\rangle = \sum_n \vec{a}_n g_z(t-nT) + e^{-i\Omega t} e^{i\frac{\beta_2}{2}\Omega^2 z} \sum_n \vec{b}_n g_z(t-nT-\beta_2\Omega z), \tag{124}$$

where $g_z(t) = \exp\left\{-iz\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\right\} g_0(t)$ is the fundamental pulse waveform propagated in the presence of chromatic dispersion only. We assume that the transmitted symbols are extracted at the receiver by using a matched filter, so that inter-symbol interference would be absent in the linear transmission regime. For this reason, the fundamental waveform must fulfill the condition $\int g_0^*(t-nT)g_0(t-mT)\,\mathrm{d}t = \delta_{n,m}$, while matched filtering is implemented by using the propagated fundamental waveform $g_L(t)$, where L is the link length. This is equivalent to compensate digitally for chromatic dispersion first and then use the original waveform $g_0(t)$ for matched filtering. Note that this procedure relies on the assumption that the coherent receiver produces a perfect replica of the optical field complex envelope in the electrical domain. Therefore, the NLIN affects the reception of the n-th symbol by adding a perturbation term with the following expression

$$\Delta\vec{a}_n = \int_{-\infty}^{\infty} g_L^*(t-nT)|\Delta E(L,t)\rangle \mathrm{d}t. \tag{125}$$

The characterization of the statistical properties of $\Delta\vec{a}_n$ and its dependence on the system parameters is not straightforward. However, the form of $\Delta\vec{a}_n$ can be identified by inspecting Eqs. (121) and (125) jointly. This shows that $\Delta\vec{a}_n$ consists of triplets of the form $\vec{a}_j^{\dagger}\vec{a}_k\vec{a}_l$, $\vec{b}_j^{\dagger}\vec{b}_k\vec{b}_l$, or $\vec{b}_j^{\dagger}\vec{a}_k\vec{b}_l$, and therefore, it can be expressed as

$$\Delta\vec{a}_n = i\frac{8}{9}\gamma \sum_{l,k,m} S_{l,k,m}\vec{a}_k^{\dagger}\vec{a}_l\vec{a}_m + i\frac{8}{9}\gamma \sum_{l,k,m} X_{l,k,m}\left(\vec{b}_k^{\dagger}\vec{b}_m\mathbf{I} + \vec{b}_m\vec{b}_k^{\dagger}\right)\vec{a}_l, \tag{126}$$

where the coefficients $S_{l,k,m}$ and $X_{l,k,m}$ are given by

$$S_{l,k,m} = i\frac{8}{9}\gamma\int_0^L \mathrm{d}z f(z) \int_{-\infty}^{\infty} \mathrm{d}t\, g_z^*(t-nT)g_z(t-lT)g_z^*(t-kT)g_z(t-mT), \tag{127}$$

$$\begin{aligned} X_{l,k,m} = i\frac{8}{9}\gamma\int_0^L \mathrm{d}z f(z) \int_{-\infty}^{\infty} \mathrm{d}t\, g_z^*(t-nT)g_z(t-lT)g_z^*(t-kT-\beta_2\Omega z) \\ \times g_z(t-mT-\beta_2\Omega z), \end{aligned} \tag{128}$$

and their numerical evaluation is discussed in Ref. [51]. The various triplets appearing in Eq. (126) reflect the nonlinear effects described in the previous sections. In particular, the triplets $\vec{a}_k^{\dagger}\vec{a}_l\vec{a}_m$ are the result of SPM, the triplets $\left(\vec{b}_k^{\dagger}\vec{b}_m\right)\vec{a}_l$ are produced by the cross-phase modulation imposed on the COI by the IC, and finally the triplets $\left(\vec{b}_m\vec{b}_k^{\dagger}\right)\vec{a}_l$ are associated to the nonlinear polarization rotation.

Within the perturbation approach, the impact of the NLIN on system performance is assessed by looking at the NLIN power, which is the variance of $\Delta\vec{a}_n$, where the ensemble averaging is performed with respect to the constellation symbols statistics. To this end, it is customary to neglect the contribution of SPM, as this is an intra-channel propagation effect and, at least in principle, it could be removed by digitally back-propagating the signal received in the COI [52]. Therefore, assuming that the transmitted constellation symbols are uncorrelated in time and between polarization channels, a tedious, yet straightforward calculation yields for the NLIN variance

$$P_{\mathrm{NLIN}} = P^3\chi_1 + P^3\chi_2\left(\frac{E\left[|b|^4\right]}{E\left[|b|^2\right]^2} - 2\right), \tag{129}$$

where P is the average signal power transmitted in each channel, E[ ] denotes ensemble averaging with respect to the constellation symbol probability distribution, and where the coefficients χ_1 and χ_2 are given by

$$\chi_1 = \frac{32}{27}\gamma^2\sum_{l,m,n}|X_{l,m,n}|^2 \tag{130}$$

$$\chi_2 = \frac{80}{81}\gamma^2\sum_{l,m}|X_{l,m,n}|^2. \tag{131}$$

Their computation is discussed in Ref. [51]. The first term on the right-hand-side of Eq. (129) coincides with the famous GN model result [53], whereas the second term is responsible for the dependence of the NLIN power on the modulation format and was first reported in Ref. [9]. This term vanishes in the case of Gaussian modulation, for which $E[|b|^4] = 2E[|b|^2]^2$.

An important implication of Eq. (129) is that the transmission quality cannot be improved by increasing the transmit power indefinitely. This can be seen by defining a generalized (effective) SNR as the ratio between the signal power and the sum of the ASE noise power and NLIN power, namely

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{P}{P_{\mathrm{ASE}} + P_{\mathrm{NLIN}}} = \frac{P}{P_{\mathrm{ASE}} + \xi P^3}, \tag{132}$$

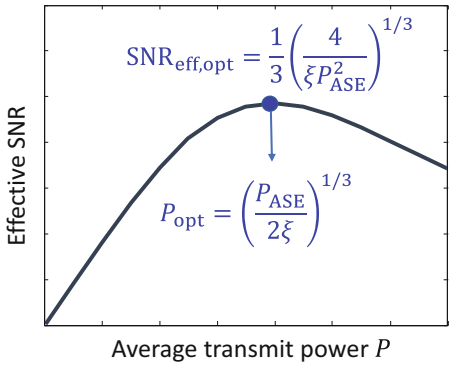

Fig. 8 The effective SNR of Eq. (132), as a function of the launched average optical power P

where the proportionality coefficient between P_{NLIN} and P^3 can be obtained from Eq. (129) in the two-channel case examined here. It is easy to see from Eq. (132) that the effective SNR peaks at the optimal signal power $P_{\mathrm{opt}} = (P_{\mathrm{ASE}}/2\xi)^{1/3}$, whereas it degrades for larger signal powers. This is illustrated in Fig. 8. We notice that while the extension of the NLIN characterization presented here to the case of larger channel counts is rather involved, the conclusion drawn from Fig. 8 remains valid.

Conclusions and Outlook

The study of lightwave propagation in optical fibers continues to pose new and fascinating problems, as new transmission regimes are explored in response to the ever growing demand for data traffic that today's global fiber-optic network is experiencing. The modeling of nonlinear transmission in ultrawide band systems that make use of spectral regions beyond the C-band is a notable example of new problems to be addressed, as well as the study of various propagation effects in new-generation fibers for space-division multiplexing [24]. These studies are key for exploring the ultimate transmission-capacity limits of the fiber-optic channel. This chapter provides an introduction to the modeling of some of the main propagation effects that are relevant in today's fiber-based transmission systems. These are chromatic dispersion, attenuation and amplification, polarization-mode dispersion, polarization-dependent loss, and nonlinear interference.

References

1. D. Gloge, Weakly guiding fibers. Appl. Opt. **10**(10), 2252–2258 (1971) [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-10-10-2252
2. D.B. Keck, Chapter 1 – Optical fiber waveguides, in *Fundamentals of Optical Fiber Communications*, ed. by M.K. Barnoski, 2nd edn., (Academic Press, 1981), pp. 1–107. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780120791514500054

3. A. Yariv, Chapter 3 – Propagation of optical beams in fibers, in *Optical Electronics*, Optics and photonics, ed. by A. Yariv, 4th edn., (Oxford University Press, New York, 1991), pp. 74–109
4. D. Marcuse, Gaussian approximation of the fundamental modes of graded-index fibers. J. Opt. Soc. Am. **68**(1), 103–109 (1978) [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=josa-68-1-103
5. T. Tamir, E. Garmire, J. Hammer, H. Kogelnik, F. Zernike, *Integrated Optics*, Topics in applied physics (Springer, Berlin/Heidelberg, 2013) [Online]. Available: https://books.google.it/books?id=vfXxCAAAQBAJ
6. C. Antonelli, A. Mecozzi, M. Shtaif, P.J. Winzer, Nonlinear propagation equations in fibers with multiple modes—Transitions between representation bases. APL Photon. **4**(2), 022806 (2019) [Online]. Available: https://doi.org/10.1063/1.5084118
7. J.P. Gordon, H. Kogelnik, PMD fundamentals: Polarization mode dispersion in optical fibers. Proc. Natl. Acad. Sci. **97**(9), 4541–4550 (2000) [Online]. Available: http://www.pnas.org/content/97/9/4541
8. Sellmeier, Zur erklärung der abnormen farbenfolge im spectrum einiger substanzen. Ann. Phys. **219**(6), 272–282 (1871) [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.18712190612
9. R. Dar, M. Feder, A. Mecozzi, M. Shtaif, Properties of nonlinear noise in long, dispersion-uncompensated fiber links. Opt. Express **21**(22), 25685–25699 (2013) [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-21-22-25685
10. D. Anderson, M. Lisak, Analytic study of pulse broadening in dispersive optical fibers. Phys. Rev. A **35**, 184–187 (1987) [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.35.184
11. P.J. Winzer, D.T. Neilson, A.R. Chraplyvy, Fiber-optic transmission and networking: The previous 20 and the next 20 years. Opt. Express **26**(18), 24190–24239 (2018) [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-18-24190
12. A. Mecozzi, R.-J. Essiambre, Nonlinear shannon limit in pseudolinear coherent systems. J. Lightwave Technol. **30**(12), 2011–2024 (2012)
13. J.S. Toll, Causality and the dispersion relation: Logical foundations. Phys. Rev. **104**, 1760–1770 (1956) [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.104.1760
14. C.M. Caves, Quantum limits on noise in linear amplifiers. Phys. Rev. D. **26**, 1817–1839 (1982) [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.26.1817
15. J. Bromage, Raman amplification for fiber communications systems. J. Lightwave Technol. **22**(1), 79 (2004) [Online]. Available: http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-22-1-79
16. H. Sunnerud, C. Xie, M. Karlsson, R. Samuelsson, P.A. Andrekson, A comparison between different PMD compensation techniques. J. Lightwave Technol. **20**(3), 368 (2002) [Online]. Available: http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-20-3-368
17. F. Buchali, H. Bülow, Adaptive pmd compensation by electrical and optical techniques. *J. Lightwave Technol.* **22**(4), 1116 (2004) [Online]. Available: http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-22-4-1116
18. H. Kogelnik, R.M. Jopson, L.E. Nelson, Chapter 15 – Polarization-mode dispersion, in *Optical Fiber Telecommunications IV-B*, Optics and photonics, ed. by I.P. Kaminow, T. Li, 4th edn., (Academic Press, Burlington, 2002), pp. 725–861. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123951731500153
19. A. Galtarossa, L. Palmieri, M. Schiano, T. Tambosso, Measurement of birefringence correlation length in long, single-mode fibers. Opt. Lett. **26**(13), 962–964 (2001) [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-26-13-962
20. M. Shtaif, A. Mecozzi, *Modelling of Polarization Mode Dispersion in Optical Communications Systems* (Springer, New York, 2005), pp. 34–51. [Online]. Available: https://doi.org/10.1007/0-387-26307-1_2
21. M. Karlsson, J. Brentel, Autocorrelation function of the polarization-mode dispersion vector. Opt. Lett. **24**(14), 939–941 (1999) [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-24-14-939

22. M. Shtaif, A. Mecozzi, J.A. Nagel, Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states. IEEE Photon. Technol. Lett. **12**(1), 53–55 (2000)
23. M. Shtaif, A. Mecozzi, Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion. Opt. Lett. **25**(10), 707–709 (2000) [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-25-10-707
24. R. Ryf, C. Antonelli, *Space-Division Multiplexing* (Springer, Cham, 2020), pp. 353–393. [Online]. Available: https://doi.org/10.1007/978-3-030-16250-4_10
25. S.J. Savory, Digital filters for coherent optical receivers. Opt. Express **16**(2), 804–817 (2008) [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-2-804
26. M.S. Faruk, K. Kikuchi, Adaptive frequency-domain equalization in digital coherent optical receivers. Opt. Express **19**(13), 12789–12798 (2011) [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-13-12789
27. N. Gisin, Statistics of polarization dependent losses. Opt. Commun. **114**(5), 399–405 (1995) [Online]. Available: https://www.sciencedirect.com/science/article/pii/0030401895006874
28. N. Gisin, B. Huttner, Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers. Opt. Commun. **142**(1), 119–125 (1997) [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401897002368
29. A. Mecozzi, M. Shtaif, The statistics of polarization-dependent loss in optical communication systems. IEEE Photon. Technol. Lett. **14**(3), 313–315 (2002)
30. A. Galtarossa, L. Palmieri, The exact statistics of polarization-dependent loss in fiber-optic links. IEEE Photon. Technol. Lett. **15**(1), 57–59 (2003)
31. L.E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, L. Rapp, Statistics of polarization dependent loss in an installed long-haul wdm system. Opt. Express **19**(7), 6790–6796 (2011) [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-19-7-6790
32. S. Zarkosvky, M. Shtaif, Statistical distribution of polarization-dependent loss in systems characterized by the hinge model. Opt. Lett. **45**(5), 1224–1227 (2020) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-45-5-1224
33. C. Antonelli, A. Mecozzi, L.E. Nelson, P. Magill, Autocorrelation of the polarization-dependent loss in fiber routes. Opt. Lett. **36**(20), 4005–4007 (2011) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-36-20-4005
34. M. Shtaif, Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss. Opt. Express **16**(18), 13918–13932 (2008) [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-16-18-13918
35. G.P. Agrawal, Chapter 5 – Optical solitons, in *Nonlinear Fiber Optics*, Optics and photonics, ed. by G. Agrawal, 5th edn., (Academic Press, Boston, 2013), pp. 129–191. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B978012397023700005X
36. L.F. Mollenauer, K. Smith, Demonstration of soliton transmission over more than 4000 km in fiber with loss periodically compensated by raman gain. Opt. Lett. **13**(8), 675–677 (1988) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-13-8-675
37. A. Mecozzi, J.D. Moores, H.A. Haus, Y. Lai, Soliton transmission control. Opt. Lett. **16**(23), 1841–1843 (1991) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-16-23-1841
38. J.P. Gordon, H.A. Haus, Random walk of coherently amplified solitons in optical fiber transmission. Opt. Lett. **11**(10), 665–667 (1986) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-11-10-665
39. S.K. Turitsyn, B.G. Bale, M.P. Fedoruk, Dispersion-managed solitons in fibre systems and lasers. Phys. Rep. **521**(4), 135–203 (2012) [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157312002657
40. S.K. Turitsyn, J.E. Prilepsky, S.T. Le, S. Wahls, L.L. Frumin, M. Kamalian, S.A. Derevyanko, Nonlinear fourier transform for optical data processing and transmission: Advances and

perspectives. Optica **4**(3), 307–322 (2017) [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-4-3-307
41. S. Evangelides, L. Mollenauer, J. Gordon, N. Bergano, Polarization multiplexing with solitons. J. Lightwave Technol. **10**(1), 28–35 (1992)
42. S.V. Manakov, On the theory of two-dimensional stationary self-focusing of electromagnetic waves. Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki **65**, 505–516 (1973)
43. P.K.A. Wai, C.R. Menyuk, H.H. Chen, Stability of solitons in randomly varying birefringent fibers. Opt. Lett. **16**(16), 1231–1233 (1991) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-16-16-1231
44. A. Mecozzi, C. Antonelli, M. Shtaif, Nonlinear propagation in multimode fibers in the strong coupling regime. Opt. Express **20**(11), 11673–11678 (2012) [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-11-11673
45. A. Mecozzi, C. Antonelli, M. Shtaif, Coupled Manakov equations in multimode fibers with strongly coupled groups of modes. Opt. Express **20**(21), 23436–23441 (2012) [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-21-23436
46. S. Mumtaz, R.J. Essiambre, G.P. Agrawal, Nonlinear propagation in multimode and multicore fibers: Generalization of the Manakov equations. J. Lightwave Technol. **31**(3), 398–406 (2013)
47. P. Wai, C. Menyak, Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence. J. Lightwave Technol. **14**(2), 148–157 (1996)
48. J.P. Gordon, *Statistical Properties of Polarization Mode Dispersion* (Springer, New York, 2005), pp. 52–59. [Online]. Available: https://doi.org/10.1007/0-387-26307-1_3
49. L.F. Mollenauer, J.P. Gordon, F. Heismann, Polarization scattering by soliton–soliton collisions. Opt. Lett. **20**(20), 2060–2062 (1995) [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-20-20-2060
50. R. Dar, M. Feder, A. Mecozzi, M. Shtaif, Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications. J. Lightwave Technol. **34**(2), 593–607 (2016)
51. R. Dar, M. Feder, A. Mecozzi, M. Shtaif, Inter-channel nonlinear interference noise in wdm systems: Modeling and mitigation. J. Lightwave Technol. **33**(5), 1044–1053 (2015)
52. E. Ip, J.M. Kahn, Compensation of dispersion and nonlinear impairments using digital back-propagation. J. Lightwave Technol. **26**(20), 3416–3425 (2008) [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-26-20-3416
53. P. Poggiolini, The gn model of non-linear propagation in uncompensated coherent optical systems. J. Lightwave Technol. **30**(24), 3857–3879 (2012) [Online]. Available: http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-30-24-3857