Nonlinear Shannon Limit in Optical Fiber Transmission System

Akihiro MARUTA†, Senior Member

SUMMARY  The remaining issues in optical transmission technology are the degradation of optical signal to noise power ratio due to amplifier noise and the distortions due to optical nonlinear effects in a fiber. Therefore in addition to the Shannon limit, practical channel capacity is believed to be restricted by the nonlinear Shannon limit. The nonlinear Shannon limit has been derived under the assumption that the received signal points on the constellation map deviated by optical amplifier noise and nonlinear interference noise are symmetrically distributed around the ideal signal point and the sum of the noises are regarded as white Gaussian noise. The nonlinear Shannon limit is considered as a kind of theoretical limitation. However it is doubtful that its derivation process and applicable range have been understood well. In this paper, some fundamental papers on the nonlinear Shannon limit are reviewed to better understanding its meaning and applicable range.

key words: optical fiber communication, nonlinear Shannon limit, Gaussian noise model, four wave mixing

1. Introduction

The Shannon limit [1] which represents the theoretical upper limit of the channel capacity is the most important and most famous expression in communication engineering. The spectral efficiency $C/W \ [\text{bit/s/Hz}]$ is defined as the channel capacity $C \ [\text{bit/s}]$ per unit bandwidth under the influence of white Gaussian noise. The expression of the Shannon limit is given by

$$ C/W = \log_2 (1 + \text{SNR}) \ , $$

where SNR is the signal-to-noise ratio and $W \ [\text{Hz}]$ represents the bandwidth.

The information is transmitted by modulating amplitude, frequency and/or phase of lightwave following the temporal variation of information data in optical fiber communication system. In conventional systems, binary digital transmission is carried out by assigning the binary data, ‘0’ and ‘1’, to the status that the light is on and off, respectively. Recently, research and development on multi-level optical transmission in which amplitude and phase of lightwave are modulated have been actively carried out and their practical implementation has already begun [2] by the development of coherent optical communication technology [3], [4]. Optical transmission technologies, which can overcome fiber loss by optical amplification and overcome linear distortions due to group velocity dispersion and polarization mode dispersion by coherent detection assisted by digital signal processing, have remaining problems caused by the degradation of optical signal-to-noise power ratio (OSNR) and the distortions due to optical nonlinear effects in a fiber. Therefore in addition to the Shannon limit, practical channel capacity is believed to be limited by nonlinear Shannon limit [5]–[17]. The nonlinear Shannon limit has been derived under the assumption that the received signal points on the constellation map deviated by optical amplifier noise and nonlinear interference noise are assumed to be symmetrically distributed around the ideal signal point and the sum of the noises are regarded as white Gaussian noise. The nonlinear Shannon limit is considered as a kind of theoretical limitation. However it is doubtful that its derivation process and applicable range have been understood well.

In this paper, some fundamental papers [5]–[8] on the nonlinear Shannon limit are reviewed to better understanding its meanings and applicable range. In Sect. 2, we derive a model equation which describes a behavior of lightwave propagation in a fiber. In Sect. 3, the deductive derivation of the nonlinear Shannon limit starting from the model equation is presented. In Sect. 4, the application of the derived expressions to a practical system evaluation and design are discussed. Finally, Sect. 5 is devoted to conclude the discussion.

2. Lightwave Propagation in a Fiber

A model equation which describes lightwave propagation in a fiber is reviewed in this section [18].

2.1 Lightwave Propagation in a Linear Fiber

We consider an electric field of a linearly polarized (LP) mode propagating in a single mode fiber which can be treated with weakly guiding approximation. By setting the $z$-axis to the fiber axis, we assume that the electric field $E(r_\perp, z, t)$ of lightwave in a fiber can be expressed as

$$ E(r_\perp, z, t) = i \text{Re}[E(z, t) \phi(r_\perp) \exp[i(\omega_0 t - \beta_0 z)]], $$

where Re$[\ast]$ represents the real part of $[\ast]$. $r_\perp$ is the coordinate on the fiber’s cross section which is perpendicular to the $z$-axis. $t \ [\text{s}]$ expresses the time. $i$ represents a unit...
vector towards the direction of polarization, i.e., the direction of the electric field, and it is perpendicular to the \( z \)-axis. \( E(z, t) \) expresses a complex amplitude of the envelope of the electric field and its absolute value represents the pulse shape. \( \phi(r_z) \) represents the modal function of the fundamental mode (LP_{01} mode). \( i = \sqrt{-1} \) is the imaginary unit. \( \omega_0 \) [rad/s] represents the angular frequency of the carrier. \( \beta_0 \) [rad/m] represents the propagation constant of the fundamental mode at \( \omega_0 \).

Next, we calculate the Fourier transform of the optical pulse, \( E(z, t) \), with respect to the time \( t \). The Fourier transformation is defined by

\[
\tilde{E}(z, \Delta \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(z, t) \exp(i \Delta \omega t) \, dt. \tag{3}
\]

Equation (3) represents that the temporary localized optical pulse has finite width spectrum. By setting the frequency spectrum of the optical pulse at the input of the fiber as \( \tilde{E}(0, \Delta \omega) \), the frequency spectrum of the pulse, \( \tilde{E}(L, \Delta \omega) \) observed after propagation of the distance \( L \) in a linear fiber, the propagation constant of which is given by \( \beta(\omega) \), can be calculated as

\[
\tilde{E}(L, \Delta \omega) = \exp[i \Delta \beta(\Delta \omega)L] \tilde{E}(0, \Delta \omega), \tag{4}
\]

where \( \Delta \beta(\Delta \omega) = \beta(\omega) - \beta_0 \). The inverse Fourier transform of Eq. (3) is given by

\[
E(z, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{E}(z, \Delta \omega) \exp(-i \Delta \omega t) \, d(\Delta \omega). \tag{5}
\]

Then the optical pulse observed at the output of the fiber (\( z = L \)) can be calculated by substituting Eq. (4) into Eq. (5).

\[
E(L, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{E}(0, \Delta \omega) \exp[-i(\Delta \omega t - \Delta \beta L)] \, d(\Delta \omega). \tag{6}
\]

When \( \Delta \beta \) is an arbitrary function of \( \Delta \omega \), analytical integration of Eq. (6) is difficult. By applying Taylor expansion to the propagation constant \( \beta(\omega) \) which is a function of angular frequency \( \omega \), around the carrier frequency \( \omega_0 \), we have

\[
\beta(\omega) = \beta_0 + \frac{d\beta}{d\omega} \bigg|_{\omega=\omega_0} (\omega - \omega_0) + \frac{1}{2} \frac{d^2\beta}{d\omega^2} \bigg|_{\omega=\omega_0} (\omega - \omega_0)^2 + \frac{1}{6} \frac{d^3\beta}{d\omega^3} \bigg|_{\omega=\omega_0} (\omega - \omega_0)^3 + \cdots. \tag{7}
\]

Then\n
\[
\Delta \beta(\Delta \omega) = \beta_0^{(1)}(\Delta \omega) + \frac{1}{2} \beta_0^{(2)}(\Delta \omega)^2 + \frac{1}{6} \beta_0^{(3)}(\Delta \omega)^3 + \cdots, \tag{8}
\]

where

\[
\beta^{(m)}_0 = \left. \frac{d^m \beta}{d\omega^m} \right|_{\omega=\omega_0}, \quad (m = 1, 2, \ldots). \tag{9}
\]

The necessary number of expansion terms in Eq. (8) to approximate \( \Delta \beta \) is dependent on the pulse width at the input of the fiber. Namely, since the bandwidth is wider for narrower pulse width, higher order term needs to be considered.

The dispersion parameter \( D \) [s/m\(^2\)] and the dispersion slope \( D_s \) [s/m\(^3\)] are frequently used to characterize a fiber. We here show the relation between \( D \), \( D_s \), and \( \beta_0^{(2)} \) [s\(^2\)/m], \( \beta_0^{(3)} \) [s\(^3\)/m]. \( D \) is defined by the differentiation of \( \beta^{(1)}(\omega) \) with respect to the wavelength \( \lambda \).

\[
D = \frac{d\beta^{(1)}}{d\lambda} \bigg|_{\omega=\omega_0} = \frac{d}{d\lambda} \left( \frac{d\beta}{d\omega} \right) \bigg|_{\omega=\omega_0} = \frac{d\omega}{d\lambda} \frac{d\beta}{d\omega} \bigg|_{\omega=\omega_0} = \frac{d\lambda}{d\omega} \beta_0^{(2)}. \tag{10}
\]

The relation between the angular frequency \( \omega \) and the wavelength \( \lambda \) is

\[
\omega = \frac{2\pi c}{\lambda}, \tag{11}
\]

where \( c \) represents the light speed in vacuum. Then the following equation can be derived by differentiating the above equation with respect to \( \lambda \) as

\[
\frac{d\omega}{d\lambda} = -\frac{2\pi c}{\lambda^2}. \tag{12}
\]

By substituting Eq. (12) into Eq. (10), we obtain

\[
D = -\frac{2\pi c}{\lambda^2} \beta_0^{(2)}. \tag{13}
\]

\( D_s \) is defined by the differentiation of \( D \) with respect to \( \lambda \) as

\[
D_s = \frac{dD}{d\lambda} \bigg|_{\omega=\omega_0} = -2\pi c \frac{d}{d\lambda} \left( \frac{\beta_0^{(2)}}{\lambda^2} \right) \bigg|_{\omega=\omega_0} = -2\pi c \frac{\beta_0^{(2)}}{\lambda^3} + \frac{1}{\lambda^2} \frac{d\beta_0^{(2)}}{d\lambda} \bigg|_{\omega=\omega_0} \]

\[
= -2\pi c \frac{\beta_0^{(2)}}{\lambda^3} + 2\pi c \frac{d\lambda}{d\omega} \frac{d}{d\omega} \left( \frac{d\beta}{d\omega^2} \right) \bigg|_{\omega=\omega_0} = -2\pi c \frac{\beta_0^{(2)}}{\lambda^3} + 2\pi c \left( \frac{2\pi c}{\lambda^3} \right)^2 \beta_0^{(3)} \]

\[
= -2\pi c \frac{\beta_0^{(2)}}{\lambda^3} + 2\pi c \left( \frac{2\pi c}{\lambda^3} \right)^2 \beta_0^{(3)}. \tag{14}
\]

### 2.2 Nonlinear Effect in a Fiber

Light propagating in a fiber is confined in a core which has higher refractive index than claddings and which diameter is \( \mu \text{m} \)-order, then the intensity of the electric field is locally high and the nonlinear polarization could not be ignored. Since the fiber loss is small and the light can propagate for
long distance with preserving its high intensity, the nonlinear optical effects are accumulated along the fiber and then obviously exhibited. In this subsection, the nonlinear optical effects in a fiber are briefly explained.

We expand the polarization vector $P$ of the medium by the electric field $E$ by power series as

$$P = ε_0(χ^{(1)} : E + χ^{(2)} : EEE + \ldots), \quad (15)$$

where $ε_0$ is the permittivity in vacuum, $χ^{(1)}$ is a linear susceptibility tensor (2nd-order tensor), $χ^{(j)} (j ≥ 2)$ is a $j$-th order nonlinear susceptibility tensor (($j + 1$)-th order tensor). Equation (15) shows that the polarization is proportional to the electric field and the medium can be considered as linear for weak intensity case, the nonlinear polarization can not be ignored and the nonlinear optical effects are exhibited for strong intensity case. Optical fiber is made from amorphous silica glass. Since the glass is not a crystal without noncentrosymmetry, it does not exhibit the second-order nonlinear optical effect induced by $χ^{(2)}$ in general. The third-order nonlinear optical effect induced by $χ^{(3)}$ is then the lowest-order effect.

Let us consider the nonlinear refractive index variation $Δn_{NL}$ in which the index is changed with proportional to the intensity of electric field due to optical Kerr effect in one of the third-order effects.

The variation of fiber’s propagation constant $β$, namely $Δβ$, arising from the third-order effect is proportional to $Δn_{NL}$ such like

$$Δβ = k_0 Δn_{NL}, \quad (16)$$

where $k_0 = 2π/λ$ is the wave number in vacuum. The transmitted power through a fiber, $P$ [W] can be calculated using Eq. (2) as

$$P(z,t) = \int S \vec{i}_z \cdot (E × H) \ dS = \frac{β_0 ϵ_0}{2k_0} |E(z,t)|^2 \int S |ϕ(r_⊥)|^2 \ dS, \quad (17)$$

where $i_⊥$ is a unit vector along the fiber axis, $H$ is the magnetic field, and $∫S \ dS$ represents the surface integration over the fiber’s cross section. By defining

$$E′(z,t) ≡ \sqrt{\frac{β_0 ϵ_0}{2k_0}} \int S |ϕ(r_⊥)|^2 \ dS E(z,t), \quad (18)$$

based on Eq. (17), the unit of $|E′(z,t)|^2$ results in [W] and then Eq. (16) can be rewritten as

$$Δβ = γ |E′|^2 E′. \quad (19)$$

Here, $γ [1/(m \cdot W)]$ is the nonlinear coefficient of optical fiber and can be represented as

$$γ = k_0 \frac{ n_2 }{ A_{ef}}, \quad (20)$$

by using nonlinear refractive index $n_2$ [m$^2$/W] and effective area of core $A_{ef}$ [m$^2$]. $A_{ef}$ is defined by

$$A_{ef} = \left( \frac{∫_S |ϕ(r_⊥)|^2 \ dS} {∫_S |ϕ(r_⊥)|^4 \ dS} \right)^{1/2}. \quad (21)$$

2.3 The Model Equation of Optical Pulse Propagation in a Fiber

In this subsection, we derive an equation that describes the optical pulse propagation in a fiber in which both of dispersion and nonlinear effects are considered. Firstly, we focus on the dispersion effect. We define the Fourier transform of the complex amplitude of the envelope of the electric field $E(z,t)$ by

$$\tilde{E}(Δβ, Δω) ≡ \frac{1}{2π} \int_{−∞}^{∞} \int_{−∞}^{∞} E(z,t) \exp[i(Δωt - Δβz)] \ dt \ dz, \quad (22)$$

then we have

$$\frac{1}{2π} \int_{−∞}^{∞} \int_{−∞}^{∞} \frac{∂^m E(z,t)}{∂t^m} \exp[i(Δωt - Δβz)] \ dt \ dz = (-iΔω)^m \tilde{E}, \quad (23)$$

$$\frac{1}{2π} \int_{−∞}^{∞} \int_{−∞}^{∞} \frac{∂^m E(z,t)}{∂z^m} \exp[i(Δωt - Δβz)] \ dt \ dz = (iΔβ)^m \tilde{E}. \quad (24)$$

By multiplying $\tilde{E}(Δβ, Δω)$ to Eq. (8), we obtain

$$Δβ \tilde{E} = β_0^{(1)} (Δω) \tilde{E} + \frac{1}{2} β_0^{(2)} (Δω)^2 \tilde{E} + \frac{1}{6} β_0^{(3)} (Δω)^3 \tilde{E} + \ldots. \quad (25)$$

By applying the inverse Fourier transform to Eq. (25) with considering Eqs. (23) and (24), we have

$$-i \frac{∂ E}{∂ z} = i β_0^{(1)} \frac{∂ E}{∂ t} - \frac{1}{2} β_0^{(2)} \frac{∂^2 E}{∂ t^2} - \frac{1}{6} β_0^{(3)} \frac{∂^3 E}{∂ t^3} + \ldots. \quad (26)$$

Hereafter, higher order dispersion terms than the fourth order are omitted. Since Eq. (26) is a linear equation, the replacement from $E(z,t)$ to $E′(z,t)$ defined by Eq. (18) does not change the form of equation. Next, considering the change of $Δβ$ due to the nonlinear effect shown in Eq. (19) and adding it to the right hand side of Eq. (26), we obtain

$$-i \frac{∂ E′}{∂ z} = i β_0^{(1)} \frac{∂ E′}{∂ t} - \frac{1}{2} β_0^{(2)} \frac{∂^2 E′}{∂ t^2} - \frac{1}{6} β_0^{(3)} \frac{∂^3 E′}{∂ t^3} + γ |E′|^2 E′. \quad (27)$$
By adding the fiber loss term to Eq. (27), we have

\[ \begin{align*}
&i \left( \frac{\partial E'}{\partial z} + \beta_0^{(1)} \frac{\partial E'}{\partial t} \right) - \frac{1}{2} \beta_0^{(2)} \frac{\partial^2 E'}{\partial t^2} \\
&- i \frac{1}{6} \beta_0^{(3)} \frac{\partial^3 E'}{\partial t^3} + \gamma |E'|^2 E' = -i \alpha E',
\end{align*} \]

where the loss coefficient \( \alpha \) [1/m] can be represented by

\[ \alpha = \frac{\log_{10} 10}{20} \tilde{\alpha}, \]

with using \( \tilde{\alpha} \) [dB/m]. Equation (28) represents an equation which describes a lightwave propagation in an optical fiber exhibiting the dispersion up to the third-order and the nonlinear change of refractive index. Note here that we should care about the applicable range of Eq. (28) under the assumption of single mode, single polarization, and quasi-monochromatic approximation when Eq. (28) has been derived. The quasi-monochromatic approximation means that the full width at half maximum of optical spectrum, \( (\Delta f)_{\text{FWHM}} \), is much smaller than the carrier frequency, \( f_0 \).

From Eqs. (13) and (14), \( \beta_0^{(2)} \) and \( \beta_0^{(3)} \) can be represented by using dispersion parameter \( D \) and dispersion slope \( D_s \) as,

\[ \begin{align*}
\beta_0^{(2)} &= \frac{\lambda^2}{2 \pi c} D, \\
\beta_0^{(3)} &= \frac{\lambda^3}{(2 \pi c)^2} (\alpha D_s + 2D).
\end{align*} \]

Especially, when \( \beta_0^{(2)} \) is not zero and the single channel optical pulse propagation in which pulse width is wider than a few picoseconds is considered, we can regard \( \beta_0^{(3)} \approx 0 \) and then the third order dispersion term can be omitted in Eq. (28).


A method to evaluate the influence to the transmission characteristics arising from the nonlinear optical effect in in-line dispersion uncompensated optical transmission line with using dispersion compensation by digital coherent transceiver, so-called the Gaussian noise model of nonlinear interference light is reviewed [5]–[8].

3.1 Bit Error Rate of Digital Modulation Schemes

In digital modulation/demodulation theory, the bit error rate (BER) can be given by a function of the (electrical) SNR [19].

\[ \text{BER} = f(\text{SNR}) \]

Definite form of a function \( f \) can be determined for each modulation/demodulation scheme. The SNR is defined as a ratio of averaged signal power \( P_s \) and noise power of amplified spontaneous emission (ASE) \( P_{\text{ASE}} \), and given by

\[ \text{SNR} = \frac{P_s}{P_{\text{ASE}}} \]

where the averaged signal power \( P_s \) can be obtained by

\[ P_s = \frac{P_{\text{ch}}}{R_s} \int_{-\infty}^{\infty} |H_{R_s}(\Delta \omega)|^2 \; d(\Delta \omega) \]

Here, \( P_{\text{ch}} \) is the received optical power per channel and \( R_s \) represents the symbol rate of the transmitted signal. \( H_{R_s}(\Delta \omega) \) is the baseband transfer function of the coherent receiver. The ASE noise power \( P_{\text{ASE}} \) can be obtained by

\[ P_{\text{ASE}} = \int_{-\infty}^{\infty} G_{\text{ASE}} |H_{R_s}(\Delta \omega)|^2 \; d(\Delta \omega) \]

where \( G_{\text{ASE}} \) is a power spectral density (PSD) of a single polarization component in dual polarized ASE noise which is received at the receiver.

The relation between SNR and OSNR is given by

\[ \text{OSNR} = \frac{P_s}{P_{\text{ASE}} + P_{\text{NL}}} \]

where \( P_{\text{NL}} \) is the OSNR bandwidth.

We assume a nonlinear interference (NLI) noise can be regarded as a Gaussian noise which is independent of ASE noise, then using the NLI noise power \( P_{\text{NL}} \), the SNR including the NLI noise can be represented as

\[ \text{SNR}_{\text{NL}} = \frac{P_s}{P_{\text{ASE}} + P_{\text{NL}}} \]

where

\[ P_{\text{NL}} = \int_{-\infty}^{\infty} G_{\text{NL}}(\Delta \omega) |H_{R_s}(\Delta \omega)|^2 \; d(\Delta \omega) \]

Here, \( G_{\text{NL}}(\Delta \omega) \) is the PSD of the NLI noise. The relation between SNR_{NL} and OSNR_{NL} is then

\[ \text{OSNR}_{\text{NL}} = \frac{R_s}{B_n} \text{SNR}_{\text{NL}} \]

3.2 Gaussian Noise Approximation of Transmitted Optical Signal

3.2.1 Complex Periodic White Gaussian Noise Process

A transmitted optical signal is modelled by a complex periodic white Gaussian noise (PWGN) process with the period of \( T_0 \). It is expressed with Karhunen-Loeve expansion as

\[ E_{\text{PWGN}}(t) = \frac{1}{\sqrt{T_0}} \sum_{n=-\infty}^{\infty} \xi_n \exp(-in\Omega_0 t) \]

where \( \Omega_0 = 2\pi/T_0 \). \( \xi_n \) is a random variable with the mean value of 0 and the dispersion of \( \sigma_{\xi_n}^2 = 1 \). By applying Fourier transform to Eq. (40), we have
By applying the Fourier transform to Eq. (45), we have

\[ \tilde{E}_{\text{PWGN}}(\Delta \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E_{\text{PWGN}}(t) \exp(i\Delta \omega t) \, dt \]
\[ = \frac{\sqrt{\Omega_0}}{2\pi} \sum_{n=-\infty}^{\infty} \xi_n \int_{-\infty}^{\infty} \exp[i(\Delta \omega - n\Omega_0)t] \, dt \]
\[ = \sqrt{\Omega_0} \sum_{n=-\infty}^{\infty} \xi_n \delta(\Delta \omega - n\Omega_0) , \tag{41} \]

where \( \delta(\omega) \) is Dirac’s delta function. Then the PSD for the PWGN is given by

\[ \tilde{G}_{\text{PWGN}}(\Delta \omega) = \langle |\tilde{E}_{\text{PWGN}}(\Delta \omega)|^2 \rangle \]
\[ = \Omega_0 \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle \xi_m \xi_n \rangle \delta(\Delta \omega - m\Omega_0) \delta(\Delta \omega - n\Omega_0) \]
\[ = \Omega_0 \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sigma_m^2 \delta_{nm} \delta(\Delta \omega - m\Omega_0) \delta(\Delta \omega - n\Omega_0) \]
\[ = \Omega_0 \sum_{n=-\infty}^{\infty} \delta(\Delta \omega - n\Omega_0) , \tag{42} \]

where \( \langle * \rangle \) represents the average and \( \delta_{nm} \) is Kronecker delta.

3.2.2 Gaussian Noise Approximation of Transmitted Optical Signal

We show how to apply the Gaussian noise approximation to a transmitted optical signal \( E_a(t) \). \( E_a(t) \) is assumed as a periodic function with the period of \( T_0 \) as

\[ E_a(t) = q(t) \quad \text{for} \quad 0 \leq t < T_0 . \tag{43} \]

\( M \) pulses in which each temporal waveform is represented by \( s(t) \) are allocated with the symbol period of \( T_s \) in \( 0 \leq t < T_0 \) and their amplitudes and/or phases are modulated by information data. \( q(t) \) can be then represented as

\[ q(t) = \sum_{m=0}^{M-1} a_m s(t - mT_s) . \tag{44} \]

Here, \( T_0 = MT_s \) and \( a_m \) is a complex number which represents the modulated amplitude and/or phase.

Next, let us calculate the Fourier transform of \( E_a(t) \), i.e., \( \tilde{E}_a(\Delta \omega) \). Since \( E_a(t) \) is a periodic function with the period of \( T_0 \), it can be represented with the Fourier series expansion.

\[ E_a(t) = \sum_{n=-\infty}^{\infty} q_n \exp(-in\Omega_0 t) , \tag{45} \]

where

\[ q_n = \frac{1}{T_0} \int_{0}^{T_0} q(t) \exp(in\Omega_0 t) \, dt . \tag{46} \]

By applying the Fourier transform to Eq. (45), we have

\[ \tilde{E}_a(\Delta \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E_a(t) \exp(i\Delta \omega t) \, dt \]
\[ = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} q_n \int_{-\infty}^{\infty} \exp[i(\Delta \omega - n\Omega_0)t] \, dt \]
\[ = \sqrt{2\pi} \sum_{n=-\infty}^{\infty} q_n \delta(\Delta \omega - n\Omega_0) . \tag{47} \]

By substituting Eq. (44) into Eq. (46), we obtain

\[ q_n = \frac{1}{T_0} \sum_{m=0}^{M-1} a_m \int_{0}^{T_0} s(t - mT_s) \exp(in\Omega_0 t) \, dt \]
\[ = \frac{1}{T_0} \sum_{m=0}^{M-1} a_m \exp(in\Omega_0 mT_s) \]
\[ \times \int_{-mT_s}^{T_0-mT_s} s(t) \exp(in\Omega_0 t) \, dt . \tag{48} \]

Here, we assume

\[ \bar{s}(n\Omega_0) = \frac{1}{\sqrt{2\pi}} \int_{-mT_s}^{T_0-mT_s} s(t) \exp(in\Omega_0 t) \, dt \]
\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(\tau) \exp(in\Omega_0 \tau) \, d\tau , \tag{49} \]

and use \( \Omega_0 = 2\pi/T_0 \) and \( T_s = T_0/M \), Eq. (48) can be represented as

\[ q_n = \sqrt{2\pi} \frac{1}{T_0} \bar{s}(n\Omega_0) \sum_{m=0}^{M-1} a_m \exp\left(\frac{2\pi n m}{M}\right) . \tag{50} \]

By substituting Eq. (50) into Eq. (47), we have

\[ \tilde{E}_a(\Delta \omega) = \Omega_0 \sum_{n=-\infty}^{\infty} \bar{s}(n\Omega_0) \delta(\Delta \omega - n\Omega_0) \]
\[ \times \sum_{m=0}^{M-1} a_m \exp\left(\frac{2\pi n m}{M}\right) . \tag{51} \]

Considering the property of Dirac’s delta function, \( \bar{s}(n\Omega_0) \delta(\Delta \omega - n\Omega_0) \) can be replaced by \( \bar{s}(\Delta \omega) \delta(\Delta \omega - n\Omega_0) \) in Eq. (51). Therefore,

\[ \tilde{E}_a(\Delta \omega) = \Omega_0 \sum_{n=-\infty}^{\infty} \bar{s}(\Delta \omega) \delta(\Delta \omega - n\Omega_0) \]
\[ \times \sum_{m=0}^{M-1} a_m \exp\left(\frac{2\pi n m}{M}\right) . \tag{52} \]

Furthermore, we assume

\[ \bar{s}(\Delta \omega) = |\bar{s}(\Delta \omega)| \exp(i\varphi_s(\Delta \omega)) , \tag{53} \]

and rewrite Eq. (52) as

\[ \tilde{E}_a(\Delta \omega) = \sqrt{\Omega_0} |\bar{s}(\Delta \omega)| \]
\[ \times \sum_{n=-\infty}^{\infty} \zeta_n \delta(\Delta \omega - n\Omega_0) . \tag{54} \]

Here, \( \zeta_n \) is a random variable defined by
\[ \zeta_n = \sqrt{\Omega_0} \sum_{m=0}^{M-1} a_m \times \exp \left( i \left( 2\pi n \frac{m}{M} + \varphi_s(n\Omega_0) \right) \right). \quad (55) \]

By calculating the average value of \( \zeta_n \zeta_t^* \), we have

\[
\langle \zeta_n \zeta_t^* \rangle = \Omega_0 \left( \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} a_{m_1}^* a_{m_2} \exp \left( -i \left( 2\pi \frac{m_1 m_2}{M} + \varphi_s(\ell\Omega_0) \right) \right) \right) = \Omega_0 \sigma_a^2 \exp \left[ i \left( \varphi_s(n\Omega_0) - \varphi_s(\ell\Omega_0) \right) \right] \times \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \exp \left( i 2\pi \left( n - \ell \right) \frac{m}{M} \right), \quad (56) \]

where \( \sigma_a^2 \) is the variance of the Gaussian variable of \( \zeta_n \), and

\[
\sum_{m=0}^{M-1} \exp \left( i 2\pi \left( n - \ell \right) \frac{m}{M} \right) = \begin{cases} 0 & \text{if } n \neq \ell \text{ and } (n - \ell) \neq kM, \\ M & \text{if } (n - \ell) = pM. \end{cases} \quad (57) \]

### 3.2.3 Model of Transmitted Optical Signal

Based on the discussions in Sects. 3.2.1 and 3.2.2, using Eq. (41), the transmitted optical signal approximated by the Gaussian noise model is represented by

\[
\hat{G}_{\text{cos}}(\Delta\omega) = \sqrt{G_{T_s}(\Delta\omega)} \hat{E}_{\text{prop}}(\Delta\omega) = \sqrt{G_{T_s}(\Delta\omega)} \sum_{n=-\infty}^{\infty} \hat{E}_n(0) \delta(\Delta\omega - n\Omega_0) = \sum_{n=-\infty}^{\infty} \hat{E}_n(0) \delta(\Delta\omega - n\Omega_0), \quad (58) \]

where

\[
\hat{E}_n(0) = \xi_n \sqrt{\Omega_0 G_{T_s}(n\Omega_0)}. \quad (59) \]

Here, \( G_{T_s}(\Delta\omega) \) represents the envelope of modulated spectrum of each channel in the transmitter. Then the PSD of the transmitted signal is given by

\[
\hat{G}_{\text{cos}}(\Delta\omega) = \langle |\hat{E}_{\text{cos}}(\Delta\omega)|^2 \rangle = \Omega_0 G_{T_s}(\Delta\omega) \sum_{n=-\infty}^{\infty} \sum_{\ell=-\infty}^{\infty} (\xi_n \xi_{\ell n}) \delta(\Delta\omega - n\Omega_0) = \Omega_0 G_{T_s}(\Delta\omega) \sum_{n=-\infty}^{\infty} \sum_{\ell=-\infty}^{\infty} \delta_{n\ell} \sigma_a^2 \delta(\Delta\omega - n\Omega_0). \quad (60) \]

Next, the inverse Fourier transform of Eq. (58) leads

\[
E_{\text{cos}}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{E}_{\text{cos}}(\Delta\omega) \exp(-i\Delta\omega t) \, d(\Delta\omega)
\]

\[
= \sum_{n=-\infty}^{\infty} \hat{E}_n(0) \sqrt{\Omega_0 G_{T_s}(n\Omega_0)} \exp(-i\Delta\omega_0 t). \quad (61) \]

Using the input waveform to the transmission line shown in Eq. (61), let us calculate the electric field after propagating through the transmission line in the next subsections.

#### 3.3 Nonlinear Interference Noise Generated by Four Wave Mixing

##### 3.3.1 Non-degenerate Four Wave Mixing Model

For the simplicity, by picking up 4 different frequency continuous lights in the discrete spectrum shown in Eq. (61) and considering those lights propagation in a fiber, we set \( E(z,t) \) as

\[
E(z,t) = \sum_{k=1}^{4} \hat{E}_k(z) \exp(-i\Delta\omega_k t), \quad (62) \]

where \( \Delta\omega_k (k = 1, 2, 3, 4) \) is the difference between the angular frequency \( \omega_k \) and the reference angular frequency \( \omega_0 \) in Eq. (2), i.e., \( \Delta\omega_k = \omega_k - \omega_0 \). The frequency allocation is assumed to satisfy

\[
\omega_3 + \omega_4 = \omega_1 + \omega_2. \quad (63) \]

Under this condition, \( \Delta\omega_k \) satisfies

\[
\Delta\omega_3 + \Delta\omega_4 = \Delta\omega_1 + \Delta\omega_2, \quad (64) \]

as shown in Fig. 1. Hereafter, we consider the case in which \( \Delta\omega_2 + \Delta\omega_1 \neq 2\Delta\omega_1 \) is satisfied. By substituting Eq. (64) into Eq. (28), Eq. (28) can be separated to four equations as

\[
\begin{align*}
\frac{d\hat{E}_1}{dz} + \Delta\beta(\Delta\omega_1) & \hat{E}_1 \\
+ \frac{\gamma}{2\pi} & \left( |\hat{E}_1|^2 + 2 \sum_{i=2,3,4} |\hat{E}_i|^2 \right) \hat{E}_1 + 2\hat{E}_2 \hat{E}_3 \hat{E}_4 & = -ia \hat{E}_1,
\end{align*}
\]
Fig. 1 Frequency allocation of four continuous waves in which non-degenerate four wave mixing occurs.

The solution of Eq. (70) is

\[ \vec{E}_4(z) = \vec{E}_4(0) \exp\left[\left(-\frac{i}{\pi} \vec{E}_1(0) \vec{E}_2(0) \vec{E}_3(0)ight) \right] \times \exp\left[\left(-2\alpha + i\Delta B\right)z\right]. \]  

We assume that the solution of Eq. (71) has the following form,

\[ \vec{E}_4(z) = \vec{E}_4(0) \exp\left[\left(-\frac{i}{\pi} \vec{E}_1(0) \vec{E}_2(0) \vec{E}_3(0)ight) \right] \times \exp\left[\left(-2\alpha + i\Delta B\right)z\right]. \]

where \( \Delta B \) is the phase mismatching parameter. By using Eqs. (69) and (64), \( \Delta B \) can be represented by

\[ \Delta B = \Delta \beta_1 + \Delta \beta_2 - \Delta \beta_3 - \Delta \beta_4 \]

By setting \( \vec{E}_4(0) = 0 \), Eq. (74) can be integrated as

\[ \vec{E}_4(z) = \frac{1}{2\alpha - i\Delta B} \exp\left[\left(-2\alpha + i\Delta B\right)z\right]. \]

By substituting Eq. (76) into Eq. (73), \( \vec{E}_4(z) \) can be rewritten as

\[ \vec{E}_4(z) = \frac{1}{2\alpha - i\Delta B} \exp\left[\left(-2\alpha + i\Delta B\right)z\right]. \]

3.3.2 FWM Light Generated in Periodically Amplified Transmission System

Let us consider the FWM light generated in periodically amplified optical transmission system in which optical amplifiers are disposed with the spacing of \( za \) [20], [21]. By considering \( N \) amplifiers, the total system length is \( N za \). The transmission fiber is assumed to be homogeneous in which
the group velocity dispersion $\beta_0^{(2)}$, the third order dispersion $\beta_0^{(3)}$, the nonlinear coefficient $\gamma$, and the loss coefficient $\alpha$ are constant. Furthermore, optical amplifiers are placed just before the output end of each section and their power gain $G$ is set to $G = \exp(2\alpha z_a)$ with which the loss of one section with the length $z_a$ can be compensated.

FWM light generated in each section independently propagates in the remaining transmission line linearly and reaches to the output. At the output, the total FWM light is obtained by coherently summing up the FWM lights generated in each section.

Firstly, the electric field of FWM light $\hat{E}_4^{(k)}(kz_a)$ at the output of the $k$-th section which is generated in the $k$-th section can be calculated by using Eq. (77) as

$$
\hat{E}_4^{(k)}(kz_a) = \frac{i\gamma}{\pi} \hat{E}_1(0) \hat{E}_2(0) \hat{E}_3^*(0)
\times \hat{E}_4^{(k-1)}(kz_a) \frac{1 - \exp((-2\alpha + i\Delta B)z_a)}{2\alpha - i\Delta B}
\times \exp[\alpha + i\Delta \beta(\Delta \omega_4)z_a] \hat{G}.
$$

The electric field $\hat{E}_j[(k-1)z_a]$ at the input of the $k$-th section can be calculated by using Eq. (72) as

$$
\hat{E}_j[(k-1)z_a] = \hat{E}_j(0) \times \exp[i\alpha + i\Delta \beta(\Delta \omega_4)(k-1)z_a].
$$

By substituting Eq. (79) into Eq. (78), we have

$$
\hat{E}_4^{(k)}(kz_a) = \frac{i\gamma}{\pi} \hat{E}_1(0) \hat{E}_2(0) \hat{E}_3^*(0) \hat{E}_4^{(k-1)}(kz_a)
\times \exp[i\alpha + i\Delta \beta(\Delta \omega_4)(k-1)z_a] \frac{1 - \exp((-2\alpha + i\Delta B)z_a)}{2\alpha - i\Delta B}
\times \exp[i\Delta \beta(\Delta \omega_4)z_a] \hat{G}.
$$

The FWM light $\hat{E}_4^{(k)}(kz_a)$ generated in the $k$-th section is assumed to propagate in the remaining transmission line linearly and reach the output end of the transmission line. Then the FWM light $\hat{E}_4^{(k)}(Nz_a)$ at the output $z = Nz_a$ can be calculated by using Eq. (72) as

$$
\hat{E}_4^{(k)}(Nz_a) = \hat{E}_4^{(k)(Nz_a)} \exp[i\Delta \beta(\Delta \omega_4)(N-k)z_a]
= \frac{i\gamma}{\pi} \hat{E}_1(0) \hat{E}_2(0) \hat{E}_3^*(0) \hat{E}_4^{(k-1)}(kz_a) \frac{1 - \exp((-2\alpha + i\Delta B)z_a)}{2\alpha - i\Delta B}
\times \exp[i\Delta \beta(\Delta \omega_4)Nz_a] \hat{G}.
$$

By summing up Eq. (81) for all $k$, the total FWM light $\hat{E}_4^{(wa)}(Nz_a)$ at the output $z = Nz_a$ can be obtained as

$$
\hat{E}_4^{(wa)}(Nz_a) = \sum_{k=1}^{N} \hat{E}_4^{(k)}(Nz_a)
= \frac{i\gamma}{\pi} \hat{E}_1(0) \hat{E}_2(0) \hat{E}_3^*(0) \frac{1 - \exp((-2\alpha + i\Delta B)z_a)}{2\alpha - i\Delta B}
\times \exp[i\Delta \beta(\Delta \omega_4)Nz_a] \sum_{k=1}^{N} \exp[i\Delta B(k-1)z_a]
= \frac{i\gamma}{\pi} \hat{E}_1(0) \hat{E}_2(0) \hat{E}_3^*(0) \frac{1 - \exp((-2\alpha + i\Delta B)z_a)}{2\alpha - i\Delta B}
\times \exp[i\Delta \beta(\Delta \omega_4)Nz_a] \frac{1 - \exp(i\Delta Bz_a)}{2\alpha - i\Delta B}
\times \sin \left( \frac{\Delta BNz_a}{2} \right)
\times \sin \left( \frac{\Delta Bz_a}{2} \right).
$$

3.4 Gaussian Noise Model

By setting $\Delta \omega_1 = \ell \Omega_0$, $\Delta \omega_2 = m \Omega_0$, and $\Delta \omega_3 = \ell + m - n \Omega_0$, we rewrite $\Delta B$ as $\Delta B_{\ell mn}$ in Eq. (75), we have

$$
\Delta B_{\ell mn} = \sqrt{\left( (\ell - \frac{\ell + m - n}{\pi}) \Omega_0^2 \right)}
\times \left\{ \beta_0^{(2)} + \beta_0^{(3)} (\ell + m) \Omega_0 \right\}.
$$

The FWM term $\gamma \hat{E}_1 \hat{E}_2 \hat{E}_4^*/\pi$ in Eq. (68) can be represented by the sum of exchanging the order of $\hat{E}_1$ and $\hat{E}_2$, as $\gamma(\hat{E}_1 \hat{E}_2 + \hat{E}_2 \hat{E}_1) \hat{E}_4^*/(2\pi)$). Therefore, changing the order of $\ell$ and $m$, and counting them separately, the right hand side of Eqs. (77) and (82) becomes half. By rewriting $\hat{E}_4(z)$ to $\hat{E}(z,n\Omega_0)$ in these equations and substituting Eq. (59) into them, we have

$$
\hat{E}(z,n\Omega_0) = \frac{i\gamma}{2\pi} \Omega_0^{3/2} \xi \ell \xi_m \xi_{m+n}^* \frac{G_{\ell mn}(\ell \Omega_0)G_{\ell mx}(m \Omega_0)G_{\ell mx}(\ell + m - n \Omega_0)}
\times 1 - \exp[(-2\alpha + i\Delta B_{\ell mn}z_a)]
\times \exp[-\alpha + i\Delta \beta(n \Omega_0)z_a]
\times \frac{\sin \left( \frac{\Delta B_{\ell mn}Nz_a}{2} \right)}{2\alpha - i\Delta B_{\ell mn}}
\times \frac{\sin \left( \frac{\Delta B_{\ell mn}z_a}{2} \right)}{2\alpha - i\Delta B_{\ell mn}},
$$

where $G_{\ell mn}(\ell \Omega_0)G_{\ell mx}(m \Omega_0)G_{\ell mx}(\ell + m - n \Omega_0)$.
3.4.1 Gaussian Noise Model for Incoherent FWM Light

When the FWM light generated in a single span (ss) is incoherent, the FWM light generated in all sections can be obtained by summing up the power of FWM light generated in N sections.

In periodically amplified optical transmission system, by representing the power gain of optical amplifier located at just before the output end of each section as $G = \exp(2\alpha z_a)$, nonlinear interference light $\hat{E}^{\text{(in)}}(z_a, \Delta\omega)$ generated in a single span can be expressed as

$$\hat{E}^{\text{(in)}}(z_a, \Delta\omega) = \sum_{n=-\infty}^{\infty} \hat{E}^{\text{(in)}}(z_a, n\Omega_0) \sqrt{G} \delta(\Delta\omega - n\Omega_0).$$

(86)

$\hat{E}^{\text{(in)}}(z_a, n\Omega_0)$ can be calculated by using Eq. (84) as

$$\hat{E}^{\text{(in)}}(z_a, n\Omega_0) = \sum_{\ell} \sum_{m} \hat{E}(z_a, n\Omega_0)$$

$$= i \frac{\gamma}{2\pi} \Omega_0^{3/2} \exp[\{-\alpha + i\Delta\beta(n\Omega_0)\}] z_a$$

$$\times \sum_{\ell} \sum_{m} \xi_{\ell} \xi_{m} e^{i(z_a \ell + m - n)}$$

$$\times \sqrt{G_{\ell}} (\ell\Omega_0) G_{\ell\ell}(m\Omega_0) G_{\ell\ell\ell}(\ell m + n\Omega_0)$$

$$\times \frac{1 - \exp[-2\alpha - i\Delta B_{\ell m n}] z_a}{2\alpha - i\Delta B_{\ell m n}}. \quad (87)$$

Therefore the NLI power $\hat{P}^{\text{(in)}}(z_a, n\Omega_0)$ is represented as

$$\hat{P}^{\text{(in)}}(z_a, n\Omega_0) = \langle |\hat{E}^{\text{(in)}}(z_a, n\Omega_0)|^2 \rangle$$

$$= \left( \frac{\gamma}{2\pi} \right)^2 \Omega_0^3 \sum_{\ell} \sum_{m} \xi_{\ell} \xi_{m} e^{i(z_a \ell + m - n)}$$

$$\times \sqrt{G_{\ell}} (\ell\Omega_0) G_{\ell\ell}(m\Omega_0) G_{\ell\ell\ell}(\ell m + n\Omega_0)$$

$$\times \frac{1 - \exp[-2\alpha - i\Delta B_{\ell m n}] z_a}{2\alpha - i\Delta B_{\ell m n}}$$

$$\times \frac{1 - \exp[-2\alpha + i\Delta B_{\ell m n}] z_a}{2\alpha + i\Delta B_{\ell m n}}.$$

(88)

In the above, a single polarized electric field has been considered. Considering dual polarized field [22]–[24], we have

$$\hat{P}^{\text{(in)}}(z_a, n\Omega_0) = \frac{8}{27} \langle |\hat{E}^{\text{(in)}}(z_a, n\Omega_0)|^2 \rangle$$

$$= \frac{8}{27} \left( \frac{\gamma}{2\pi} \right)^2 \Omega_0^3 \sum_{\ell} \sum_{m} G_{\ell \Omega_0} (\ell\Omega_0) G_{\ell\ell\ell}(m\Omega_0) G_{\ell\ell\ell\ell}(\ell m + n\Omega_0)$$

$$\times \frac{1 - 2\exp(-2\alpha z_a) \cos(\Delta B_{\ell m n} z_a) + \exp(-4\alpha z_a)}{4\alpha^2 + (\Delta B_{\ell m n})^2}. \quad (89)$$

Therefore, the NLI power at the output end of the transmission line, $\hat{P}_{\text{NLI}}(N z_a, n\Omega_0)$ is represented as

$$\hat{P}_{\text{NLI}}(N z_a, n\Omega_0) = N \hat{P}^{\text{(in)}}(z_a, n\Omega_0)$$

$$= \frac{2}{27} \left( \frac{\gamma}{\pi} \right)^2 N \Omega_0^3 \sum_{\ell} \sum_{m} G_{\ell \Omega_0} (\ell\Omega_0) G_{\ell\ell\ell}(m\Omega_0) G_{\ell\ell\ell\ell}(\ell m + n\Omega_0)$$

$$\times \frac{1 - 2\exp(-2\alpha z_a) \cos(\Delta B_{\ell m n} z_a) + \exp(-4\alpha z_a)}{4\alpha^2 + (\Delta B_{\ell m n})^2}. \quad (90)$$

3.4.2 Gaussian Noise Model for Coherent FWM Light

When the FWM light generated in each span is coherent, the total nonlinear interference light at the output end $z = N z_a$,
\[ \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, \Delta \omega) \text{ can be represented as} \]
\[ \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, \Delta \omega) = \sum_{n=-\infty}^{\infty} \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) \delta(\Delta \omega - n\Omega_0). \]  
(91)

\[ \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) \text{ can be calculated by using Eq. (85) as} \]
\[ \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) = \sum_{\ell} \sum_{m} \hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) \]
\[ = \left( \frac{\gamma}{2\pi} \right)^{3/2} \Omega_0^{3/2} \exp(i\Delta \beta(n\Omega_0)Nz_a) \sum_{\ell} \sum_{m} \xi_{\ell}' \xi_{m} \xi_{\ell+\ell-m-n} \]
\[ \times \sqrt{G_{T_s}(\ell\Omega_0) G_T(m\Omega_0) G_T[(\ell + m - n)\Omega_0]} \]
\[ \times 1 - \exp[-(2\alpha - i\Delta B_{\text{fmm}})z_a] \]
\[ \times \frac{2\alpha - i\Delta B_{\text{fmm}}}{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)} \]
\[ \times \frac{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)}{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)} \]
\[ \exp \left[ i \frac{\Delta B_{\text{fmm}}(N-1)z_a}{2} \right]. \]  
(92)

Therefore the NLI power at the output end of the transmission line, \( \hat{P}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) \), is represented as
\[ \hat{P}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) = \langle |\hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0)|^2 \rangle \]
\[ = \left( \frac{\gamma}{2\pi} \right)^{3/2} \Omega_0^{3/2} \sum_{\ell} \sum_{m} \xi_{\ell}' \xi_{m} \xi_{\ell+\ell-m-n} \]
\[ \times \sqrt{G_{T_s}(\ell\Omega_0) G_T(m\Omega_0) G_T[(\ell + m - n)\Omega_0]} \]
\[ \times 1 - \exp[-(2\alpha - i\Delta B_{\text{fmm}})z_a] \]
\[ \times \frac{2\alpha - i\Delta B_{\text{fmm}}}{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)} \]
\[ \times \frac{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)}{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)} \]
\[ \exp \left[ i \frac{\Delta B_{\text{fmm}}(N-1)z_a}{2} \right] \]
\[ \times \exp \left[ \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right] \]
\[ \times \frac{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)}{\sin \left( \frac{\Delta B_{\text{fmm}}Nz_a}{2} \right)} \]
\[ \exp \left[ i \frac{\Delta B_{\text{fmm}}(N-1)z_a}{2} \right]. \]  
(93)

For the dual polarized field, we have
\[ \hat{P}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0) = \frac{8}{27} \langle |\hat{E}_{\text{NLI}}^{\text{total}}(Nz_a, n\Omega_0)|^2 \rangle \]
\[ = \frac{2}{27} \left( \frac{\gamma}{2\pi} \right)^{3/2} \Omega_0^{3/2} \sum_{\ell} \sum_{m} G_{T_s}(\ell\Omega_0) G_T(m\Omega_0) \]
\[ \times G_T[(\ell + m - n)\Omega_0] \]
\[ \times \frac{1 - 2 \exp(-2\alpha z_a) \cos(\Delta B_{\text{fmm}}z_a) + \exp(-4\alpha z_a)}{4\alpha^2 + (\Delta B_{\text{fmm}})^2} \]
Here, respectively. By setting \( \Delta \rho \) of the Nyquist WDM signal,

\[
\sin^2\left(\frac{\Delta B_{\text{mm}} N z_a}{2}\right) \times \frac{\sin^2\left(\frac{\Delta B_{\text{mm}} z_a}{2}\right)}{2}. \tag{94}
\]

### 3.5 Power Spectral Density of Nonlinear Interference Noise

Since the following expression is satisfied for any function of \( \hat{f}(\Delta \omega) \),

\[
\int_{-\infty}^{\infty} \hat{f}(\Delta \omega) \, d(\Delta \omega) = \lim_{\Omega_0 \to 0} \int_{-\infty}^{\infty} \Omega_0 \hat{f}(n \Omega_0), \tag{95}
\]

by setting \( \Omega_0 \to 0 \) (\( T_0 \to \infty \)) in Eq. (90), \( \ell \Omega_0 = \Delta \omega_1 \), \( m \Omega_0 = \Delta \omega_2 \), and \( n \Omega_0 = \Delta \omega \), we have

\[
G_{\text{NL}}(N z_a, \Delta \omega) = \lim_{\Omega_0 \to 0} \frac{\Omega_0}{\pi} G_{\text{NL}}(N z_a, n \Omega_0) = \frac{2}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_T(\Delta \omega_1) G_T(\Delta \omega_2) \times \frac{1 - 2 \exp(-2\alpha z_a) \cos(\Delta B z_a) + \exp(-4\alpha z_a)}{4\alpha^2 + (\Delta B)^2} \times |d(\Delta \omega_1) \, d(\Delta \omega_2)|. \tag{96}
\]

From Eq. (94), we also have

\[
G_{\text{NL}}^{(n)}(N z_a, \Delta \omega) = \lim_{\Omega_0 \to 0} \frac{\Omega_0}{\pi} G_{\text{NL}}^{(n)}(N z_a, n \Omega_0) = \frac{2}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_T(\Delta \omega_1) G_T(\Delta \omega_2) \times \frac{1 - 2 \exp(-2\alpha z_a) \cos(\Delta B z_a) + \exp(-4\alpha z_a)}{4\alpha^2 + (\Delta B)^2} \times |d(\Delta \omega_1) \, d(\Delta \omega_2)|. \tag{97}
\]

\( \Delta B \) in Eqs. (96) and (97), is given by a formula in which \( \Delta \omega_4 \) in Eq. (75) is replaced with \( \Delta \omega \).

### 3.6 Nyquist Limit

Here we focus on a Nyquist WDM transmission system. In this case, the symbol rate \( R_s \) is equal to the channel bandwidth \( B_{\text{ch}} \). We assume that \( G_T(\Delta \omega) \) is given by

\[
G_T(\Delta \omega) = \begin{cases} G_{\text{WDM}} & |\Delta \omega| \leq \pi B_{\text{WDM}}, \\ 0 & \text{otherwise}. \end{cases} \tag{98}
\]

Here, \( G_{\text{WDM}} \) and \( B_{\text{WDM}} \) represent the transmission power spectral density and total bandwidth of the Nyquist WDM signal, respectively. By setting \( \Delta \omega = 0 \) in Eq. (96), we have

\[
G_{\text{NL}}(N z_a, 0) = \frac{2}{27} \left(\frac{\gamma}{\pi}\right)^2 N G_{\text{WDM}}^3 \int_{\mathcal{D}} \left[1 - 2 \exp(-2\alpha z_a) \right. \times \frac{\sin^2\left(\frac{\Delta B N z_a}{2}\right)}{2} \times \frac{\sin^2\left(\frac{\Delta B z_a}{2}\right)}{2} \times |d(\Delta \omega_1) \, d(\Delta \omega_2)|. \tag{99}
\]

Here, we assume the third order dispersion \( \beta_3^{(3)} = 0 \) in Eq. (75). In this case, the domain of integration \( \mathcal{D} \) in the \( \Delta \omega_1 - \Delta \omega_2 \) plane is determined by Eq. (98) and it satisfies

\[
|\Delta \omega_1|, |\Delta \omega_2|, |\Delta \omega_1 + \Delta \omega_2| \leq \pi B_{\text{WDM}}. \tag{100}
\]

The domain of integration \( \mathcal{D} \) is shown in Fig. 2. Analytical integration in the domain of \( \mathcal{D} \) is difficult. Then the domain of integration is approximated by a circular domain \( \mathcal{C} \) with the radius of \( \pi B_{\text{WDM}} \) which is surrounded by a dotted line in Fig. 2. The numerator of the integrand is also approximated as

\[
1 - 2 \exp(-2\alpha z_a) \cos(||\beta_0^{(2)}|| \Delta \omega_1 \Delta \omega_2 z_a) + \exp(-4\alpha z_a) \times \frac{|\beta_0^{(2)}||\Delta \omega_1 \Delta \omega_2 z_a|}{2} \tag{101}
\]

\[
= 1 - 2 \exp(-2\alpha z_a) + \exp(-4\alpha z_a) \times \frac{4 \alpha^2 + |\beta_0^{(2)}||\Delta \omega_1 \Delta \omega_2 z_a|}{2} \tag{102}
\]

\[
= 1 - \exp(-2\alpha z_a) |\beta_0^{(2)}||\Delta \omega_1 \Delta \omega_2 z_a| \tag{103}
\]

Furthermore, the effective length of optical amplifier spacing \( L_{\text{eff}} \) and the asymptotic effective length \( L_{\text{as}} \), are respectively defined by

\[
L_{\text{eff}} = \int_0^{z_a} \exp(-2\alpha z) \, dz = \frac{1 - \exp(-2\alpha z_a)}{2 \alpha}, \tag{104}
\]

\[
L_{\text{as}} = \lim_{z_a \to \infty} L_{\text{eff}} = \frac{1}{2 \alpha}. \tag{105}
\]

Using Eqs. (101)–(103), Eq. (99) becomes
In this case, the integrand in Eq. (107) has four simple poles in the domain surrounded by $C$.

$$z = \pm \lambda, \pm i \lambda, \quad \lambda = \sqrt{\frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}} + 1 - \frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}. \quad (108)$$

Then Eq. (107) can be represented as

$$I_\varphi = i \frac{16r}{(|\beta_0^{(2)}| L_{\text{eff},r}^2)^2} \int_C f(z) \, dz, \quad (109)$$

where

$$f(z) = z^3(z - \lambda)^{-1}(z + \lambda)^{-1}(z - i\lambda)^{-1}(z + i\lambda)^{-1}.$$

$$z^4 - \left\{\frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}\right\}^2 + 1 + \frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}.$$ 

The complex integral of Eq. (109) becomes

$$I_\varphi = i \frac{16r}{(|\beta_0^{(2)}| L_{\text{eff},r}^2)^2} \times (2i\pi)$$

$$\times \{\text{Res}[f(z), \lambda] + \text{Res}[f(z), -\lambda]$$

$$+ \text{Res}[f(z), i\lambda] + \text{Res}[f(z), -i\lambda]\}.$$ 

Here,

$$\text{Res}[f(z), \lambda] = \text{Res}[f(z), -\lambda]$$

$$= \text{Res}[f(z), i\lambda] = \text{Res}[f(z), -i\lambda]$$

$$= \lambda^3(\lambda + i\lambda)^{-1}(\lambda - i\lambda)^{-1}(\lambda + i\lambda)^{-1}$$

$$= \frac{\lambda^3}{1 + \frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}}.$$ 

$$z^4 - \left\{\frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}\right\}^2 + 1 + \frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}.$$ 

$$= -\frac{|\beta_0^{(2)}| L_{\text{eff},r}^2}{32\left\{\frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}\right\}^2 + 1}.$$ 

$$= -\frac{|\beta_0^{(2)}| L_{\text{eff},r}^2}{64 \left\{\frac{2}{|\beta_0^{(2)}| L_{\text{eff},r}^2}\right\}^2}.$$ 

Therefore the complex integral of Eq. (109) leads

$$I_\varphi = \frac{32\pi r}{(|\beta_0^{(2)}| L_{\text{eff},r}^2)^2} \times 4.$$
\[
\times \frac{(|\beta_0(2)|L_{\text{at}},r^2)^2}{64 \pi r} \left( 1 + \frac{(|\beta_0(2)|L_{\text{at}},r^2)^2}{2} \right) \\
= \frac{2 \pi r}{4 \pi} \sqrt{1 + \frac{(|\beta_0(2)|L_{\text{at}},r^2)^2}{2}}.
\]

Substituting Eq. (113) into Eq. (106), we have
\[
G_{\text{NL}}(Nz,0) = \frac{2}{27} \left( \frac{\gamma}{\pi^2} \right)^2 N^3 G^3_{\text{WDM}} L_{\text{eff}}^2 \\
\times \int_0^\infty \frac{2 \pi}{\sqrt{1 + R^2}} \frac{dR}{|\beta_0(2)|L_{\text{at}},r^2} \\
= \frac{4}{27} \pi |\beta_0(2)|L_{\text{at}},r^2 \sinh^{-1} \left( \frac{\sqrt{2} |\beta_0(2)|L_{\text{at}},r^2}{2} \right).
\]

Furthermore, by changing variable with \( R = |\beta_0(2)|L_{\text{at}},r^2/2 \) and integrating it, we obtain
\[
G_{\text{NL}}(Nz,0) = \frac{2}{27} \left( \frac{\gamma}{\pi^2} \right)^2 N^3 G^3_{\text{WDM}} L_{\text{eff}}^2 \\
\times \int_0^\infty \frac{2 \pi}{\sqrt{1 + R^2}} \frac{dR}{|\beta_0(2)|L_{\text{at}},r^2} \\
= \frac{4}{27} \pi |\beta_0(2)|L_{\text{at}},r^2 \sinh^{-1} \left( \frac{\sqrt{2} |\beta_0(2)|L_{\text{at}},r^2}{2} \right).
\]

For \( x \gg 1 \),
\[
\sinh^{-1}(x) = \log_e(x + \sqrt{x^2 + 1}) = \log_e(2x).
\]

Applying the approximation of Eq. (116) to Eq. (115), we obtain
\[
G_{\text{NL}}(Nz,0) = \frac{4}{27} \frac{\gamma^2 N^3 G^3_{\text{WDM}} L_{\text{eff}}^2}{|\beta_0(2)|L_{\text{at}},r^2} \log_e(\pi^2 |\beta_0(2)|L_{\text{at}},r^2 B_{\text{WDM}}^2).
\]

The NLI noise power \( P_{\text{NL}} \) can be calculated by using Eq. (38) as
\[
P_{\text{NL}}(Nz) = \int_{-\infty}^{\infty} G_{\text{NL}}(Nz, \Delta \omega)|H_{R_s}(\Delta \omega)|^2 d(\Delta \omega).
\]

By approximating \( G_{\text{NL}}(Nz, \Delta \omega) \approx G_{\text{NL}}(Nz,0) \) and assuming the bandwidth of the receiver is equal to the OSNR bandwidth \( B_n \), we set \( H_{R_s}(\Delta \omega) \) as
\[
H_{R_s}(\Delta \omega) = \begin{cases} 1 & |\Delta \omega| \leq \pi B_n, \\ 0 & \text{otherwise}, \end{cases}
\]

Then we have
\[
P_{\text{NL}}(Nz) = 2\pi B_n G_{\text{NL}}(Nz,0) \\
= \frac{8}{27} \frac{\gamma^2 N^3 G^3_{\text{WDM}} L_{\text{eff}}^2 B_n}{|\beta_0(2)|L_{\text{at}},r^2} \log_e(\pi^2 |\beta_0(2)|L_{\text{at}},r^2 B_{\text{WDM}}^2).
\]

By setting \( L_{\text{at}},r \sim L_{\text{at}},r \), we obtain
\[
P_{\text{NL}}(Nz) = \left( \frac{2}{3} \right)^3 \frac{\gamma^2 N^3 G^3_{\text{WDM}} L_{\text{eff}} B_n}{|\beta_0(2)|L_{\text{at}},r^2} \\
\times \log_e(\pi^2 |\beta_0(2)|L_{\text{at}},r^2 B_{\text{WDM}}^2).
\]

The total bandwidth of WDM signal \( B_{\text{WDM}} \), the transmission power spectral density \( G_{\text{WDM}} \), and the total transmission power \( P_{\text{WDM}} \) are respectively expressed by using the total channel number \( N_\text{ch} \) and transmission power per channel \( P_\text{ch} \) as
\[
B_{\text{WDM}} = N_\text{ch} B_{\text{ch}}, \\
G_{\text{WDM}} = P_\text{ch}, \\
P_{\text{WDM}} = N_\text{ch} P_\text{ch} = G_{\text{WDM}} B_{\text{WDM}}.
\]

By substituting these expressions into Eq. (121), we have
\[
P_{\text{NL}}(Nz) = \left( \frac{2}{3} \right)^3 \frac{\gamma^2 N^3 P_{\text{ch}} L_{\text{eff}} B_n}{|\beta_0(2)|L_{\text{at}},r^2} \\
\times \log_e(\pi^2 |\beta_0(2)|L_{\text{at}},r^2 B_{\text{WDM}}^2) \times F_N G = F_N \exp(2\alpha z_a).
\]

4. Application to Evaluation and Design a Transmission System

Let us consider to apply Eq. (125) to evaluate and design a Nyquist WDM transmission system. By setting \( R_s = B_n \) in Eq. (39), Eq. (37) becomes
\[
\text{OSNR}_{\text{NL}} = \text{SNR}_{\text{NL}} = \frac{P_{\text{ch}}}{P_{\text{ASE}} + P_{\text{NL}}}.
\]

By defining a parameter \( F_N \) which is independent of fiber characteristics such as noise figure of optical amplifier, the total ASE noise power after propagation through the total length of the transmission line \( Nz_a \) is represented as
\[
P_{\text{ ASE}}(Nz_a) = \frac{N_{\alpha} P_{\text{ch},\alpha}}{B_{\text{ch},\alpha}} \times \log_e(\pi^2 |\beta_0(2)|L_{\text{at}},r^2 B_{\text{WDM}}^2).
\]
\[
\frac{\partial \text{OSNR}_{\text{NL}}}{\partial P_{\text{ch}}} = \frac{P_{\text{ASE}} - 2hP_{\text{ch}}^3}{(P_{\text{ASE}} + hP_{\text{ch}})^2}.
\]

By setting the right hand side of Eq. (132) to 0, the optimum value of \(P_{\text{ch}}\), i.e., \(P_{\text{ch}}^{(opt)}\), can be determined. By using Eqs. (127) and (130), we have

\[
P_{\text{ch}}^{(opt)} = \left( \frac{P_{\text{ASE}}}{2h} \right)^{1/3} = \left( \frac{F}{2\eta} \right)^{1/3} \left( \frac{Q_{\text{NL}}}{\eta^2 L_{\text{eff}}} \right)^{1/3}.
\]

By representing \(P_{\text{ch}}^{(opt)}\) in Eq. (133) with the unit of [dBm], we have

\[
P_{\text{ch}}^{(opt)}[\text{dBm}] = 10 \log_{10} \left( \frac{F}{2\eta} \right) = C_1 - \frac{10}{3} \log_{10} \left( \frac{\gamma^2 L_{\text{eff}}}{|\beta_{0}^{(2)}|} \right) + \frac{1}{3} \tilde{a}_{z_{\text{eff}}}. \tag{134}
\]

Here, Eq. (29) was used to derive the above expression. \(C_1\) is given by

\[
C_1 = \frac{10}{3} \log_{10} \left( \frac{F}{2\eta} \right). \tag{135}
\]

and it is independent of the transmission fiber. By substituting Eq. (133) into Eq. (131) and using Eqs. (127) and (130), we find the maximum value of OSNR_{NL}.

\[
\text{OSNR}_{\text{NL}}^{(max)} = \frac{1}{3h(|P_{\text{ch}}^{(opt)}|)^2} = \left( \frac{4}{27F^2 \eta} \right)^{1/3} \left( \frac{|\beta_{0}^{(2)}| \exp(-4\alpha z_{\text{eff}})}{\gamma^2 L_{\text{eff}}} \right)^{1/3}. \tag{136}
\]

By representing OSNR_{NL}^{(max)} with the unit of decibel, we obtain

\[
\text{OSNR}_{\text{NL}}^{(max)}[\text{dB}] = 10 \log_{10} \text{OSNR}_{\text{NL}}^{(max)} = C_2 - \frac{10}{3} \log_{10} \left( \frac{\gamma^2 L_{\text{eff}}}{|\beta_{0}^{(2)}|} \right) - \frac{2}{3} \tilde{a}_{z_{\text{eff}}} - 10 \log_{10} N,
\]

where

\[
C_2 = \frac{10}{3} \log_{10} \left( \frac{4}{27F^2 \eta} \right),
\]

and \(C_2\) is independent of the transmission fiber, too.

Next, we consider the relation between OSNR_{NL} and \(P_{ch}\) when the transmission power per channel \(P_{ch}\) is not optimum. By setting \(P_{ch} = rP_{ch}^{(opt)}\) with a constant \(r\), OSNR_{NL} can be represented by using Eqs. (131), (133), and (136) as

\[
\text{OSNR}_{\text{NL}}^{(r)} = \frac{P_{\text{ASE}}}{P_{\text{ASE}} + h r^2 P_{\text{ch}}^{(opt)}}, \quad \text{OSNR}_{\text{NL}}^{(max)} = \frac{3r}{r^3 + 2} \text{OSNR}_{\text{NL}}^{(max)}.
\]

By representing OSNR_{NL}^{(max)} with the unit of decibel, we find

\[
\text{OSNR}_{\text{NL}}^{(max)}[\text{dB}] = 10 \log_{10} \left( \frac{3r}{r^3 + 2} \right) + \text{OSNR}_{\text{NL}}^{(max)}[\text{dB}] = 10 \log_{10} \left( \frac{3r}{r^3 + 2} \right). \tag{140}
\]

We show Eq. (139) in Fig. 4. The optimum transmission optical power \(P_{ch}^{(opt)}\) and the maximum value of OSNR, \(\text{OSNR}_{\text{NL}}^{(max)}\), can be obtained in Eqs. (134) and (137), by the parameters of the transmission line except \(C_1\) in Eq. (135) and \(C_2\) in Eq. (138), i.e., only by the second order dispersion coefficient \(\beta_{0}^{(2)}\), nonlinear coefficient \(\gamma\), loss coefficient \(\alpha\), optical amplifier spacing \(z_{\text{eff}}\), number of sections \(N\). The OSNR when the transmission optical power per channel \(P_{ch}\) deviates from its optimal value can be calculated by using Eq. (140). These results can be applied to evaluate and design a Nyquist WDM transmission system.

SNR_{NL} is equal to OSNR_{NL} in Nyquist WDM transmission system as shown in Eq. (126) and then OSNR_{NL}^{(max)} in Eq. (136) gives the upper limit of the spectral efficiency shown in Eq. (1). This is the nonlinear Shannon limit.

5. Conclusion

The nonlinear Shannon limit has been considered as a kind of theoretical limitation. To clarify its meanings and applicable range, some fundamental papers on the nonlinear Shannon limit have been reviewed. The Gaussian noise model on the nonlinear Shannon limit has been experimentally validated for various fiber types in [8]. But it does not give the theoretical limitation of the channel capacity. Recently, various attempts to break the limitation by nonlinear compensation schemes and nonlinear transmission technologies have been reported [25–48]. The nonlinear Shannon limit...
is a transitional limitation which can be overcome.

References


Akihiro Maruta received the B.E., M.E., and Ph.D. degrees from Osaka University, Osaka, Japan, in 1988, 1990, and 1993, respectively, all in communications engineering. He was with the Department of Communications Engineering, Osaka University, in 1993. He has been an Professor in the Department of Information and Communication Technology, Osaka University, since 2016. His current research interests include optical fiber communication systems and all-optical signal processing. Dr. Maruta is a member of the IEEE Photonics Society and the Optical Society of America.