

Communication Using Eigenvalues of Higher Multiplicity of the Nonlinear Fourier Transform

Javier García

Institute for Communications Engineering
 Technical University of Munich
 javier.garcia@tum.de

Abstract—Eigenvalues of higher multiplicity of the Nonlinear Fourier Transform (NFT) are considered for information transmission over fiber optic channels. The effects of phase, time or frequency shifts on this generalized NFT are derived, as well as an expression for the signal energy. These relations are used to design transmit signals and numerical algorithms to compute the direct and inverse NFTs, and to numerically demonstrate communication using a soliton with one double eigenvalue.

Index Terms—Inverse Scattering Transform, Nonlinear Fourier Transform, fiber optics, higher multiplicity eigenvalues, spectral efficiency

I. INTRODUCTION

Current optical transmission systems exhibit a peak in the achievable rate due to the Kerr nonlinearity of the Nonlinear Schrödinger Equation (NLSE) [1]. Several techniques have been proposed to attempt to overcome this limit, of which the Inverse Scattering Transform (IST) [2], or the Nonlinear Fourier Transform (NFT) [3], has attracted considerable attention. Numerous algorithms have been developed to compute the direct [4], [5] and inverse [5], [6], [7] NFT.

Information transmission using the NFT has been demonstrated both numerically and experimentally in several works, such as [8], [9], [10]. For purely discrete spectrum modulation, the spectral efficiencies obtained so far are not very high [11]. In this paper, eigenvalues of higher multiplicity in the discrete spectrum are considered for communication. The theory for these eigenvalues has been developed in [12], [13], but their applications to communications have to the best of our knowledge not been explored yet.

The paper is organized as follows. In Section II, we introduce the NLSE model. Section III briefly describes the NFT. In Section IV, we explain the theory of higher multiplicity eigenvalues from [12], [13], and we prove some properties of this generalized NFT. In Section V, we show how to compute the direct and inverse NFT with multiple eigenvalues. Section VI numerically demonstrates information transmission using a double eigenvalue, and Section VII concludes the paper.

II. SYSTEM MODEL

Consider the slowly varying component $Q(Z, T)$ of an electrical field propagating along an optical fiber, where Z

is distance and T is time. The field obeys the NLSE, which is expressed as [14, Eq. (2.3.46)]:

$$\frac{\partial}{\partial Z} Q(Z, T) = -j \frac{\beta_2}{2} \frac{\partial^2}{\partial T^2} Q(Z, T) + j\gamma |Q(Z, T)|^2 Q(Z, T) + N(Z, T) \quad (1)$$

where β_2 is the group velocity dispersion (GVD) parameter, and γ is the nonlinear coefficient. We neglect attenuation in (1) because we assume that fiber loss is exactly compensated by distributed Raman amplification. The noise term $N(Z, T)$ is the formal derivative of a band-limited Wiener process $W(Z, T)$, i.e., we have

$$\int_0^Z N(Z', T) dZ' = \sqrt{N_{\text{ase}}} W(Z, T) \quad (2)$$

where $N_{\text{ase}} = \alpha h \nu_s K_T$ is the distributed noise spectral density, α is the attenuation coefficient, $h \approx 6.626 \cdot 10^{-34} \text{ m}^2 \text{ kgs}^{-1}$ is Planck's constant, ν_s is the signal center frequency, and K_T is the phonon occupancy factor, which is approximately 1.13 for Raman amplification [1]. Note that, unlike [1], we do not include the distance in the definition of N_{ase} . The Wiener process $W(Z, T)$ may be defined as

$$W(Z, T) = \lim_{K \rightarrow \infty} \frac{1}{\sqrt{K}} \sum_{k=1}^{\lfloor KZ \rfloor} W_k(T) \quad (3)$$

where the $W_k(T)$ are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian processes with zero mean, bandwidth B , and autocorrelation

$$\text{E} [W_k(T) W_k^*(T')] = B \text{sinc}(B(T - T')) \quad (4)$$

where $\text{sinc}(x) \triangleq \sin(\pi x) / (\pi x)$.

III. THE NONLINEAR FOURIER TRANSFORM

In this section, we briefly introduce the steps involved in the NFT. For more detail, the reader is referred to [3].

By applying the following change of variables:

$$T = T_0 t, \quad Z = 2 \frac{T_0^2}{|\beta_2|} z, \quad Q(Z, T) = \frac{1}{T_0} \sqrt{\frac{|\beta_2|}{\gamma}} q(z, t) \quad (5)$$

$$\text{E} [N(Z, T) N^*(Z', T')] = \frac{\beta_2^2}{2\gamma T_0^3} \text{E} [n(z, t) n^*(z', t')] \quad (5)$$

the NLSE (1) is normalized to

$$\frac{\partial}{\partial z} q(z, t) = -j \operatorname{sign}(\beta_2) \frac{\partial^2}{\partial t^2} q(z, t) + j2 |q(z, t)|^2 q(z, t) + n(z, t) \quad (6)$$

and we choose $\beta_2 < 0$ to focus on the case of anomalous GVD [14, p. 131]. The parameter T_0 can be freely chosen, and provides an additional degree of freedom for the normalization. The nonlinear term $|q|^2 q$ causes inter-channel interference.

The IST or NFT provides a domain in which the NLSE channel is multiplicative, i.e., there exists a *channel transfer function* $H(z, \lambda)$ such that

$$\tilde{Q}(z, \lambda) = H(z, \lambda) \tilde{Q}(0, \lambda) \quad (7)$$

where $\tilde{Q}(z, \lambda)$ is the NFT of $q(z, t)$. The NFT is based on the existence of a Lax pair (L, M) of operators that satisfies the following condition:

$$\frac{\partial L}{\partial z} = ML - LM. \quad (8)$$

As shown in [2, Section 1.4], the eigenvalues λ of L are invariant in z , and the eigenvectors v of L satisfy:

$$v_z = Mv \quad (9)$$

$$v_t = Pv \quad (10)$$

where a subscript indicates a derivative with respect to that variable. For the NLSE (6), L , M and P are given by:

$$L(z, t) = j \begin{pmatrix} \frac{\partial}{\partial t} & -q \\ -q^* & -\frac{\partial}{\partial t} \end{pmatrix} \quad (11)$$

$$M(z, t, \lambda) = \begin{pmatrix} 2j\lambda^2 - j|q|^2 & -2\lambda q - jq_t \\ 2\lambda q^* - jq_t^* & -2j\lambda^2 + j|q|^2 \end{pmatrix} \quad (12)$$

$$P(z, t, \lambda) = \begin{pmatrix} -j\lambda & q \\ -q^* & j\lambda \end{pmatrix}. \quad (13)$$

The NFT is calculated by solving the Zakharov-Shabat system (10). In the following, we often drop the dependence on z to simplify notation. Two solutions $v^1(t, \lambda)$ and $v^2(t, \lambda)$ that are bounded in the upper complex half plane ($\lambda \in \mathbb{C}^+$) are obtained using the boundary conditions

$$v^1(t, \lambda) \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{j\lambda t}, \quad t \rightarrow +\infty \quad (14a)$$

$$v^2(t, \lambda) \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda t}, \quad t \rightarrow -\infty. \quad (14b)$$

We define the *adjoint* of a vector $v = (v_1, v_2)^T$ as $\tilde{v} = (v_2^*, -v_1^*)^T$. Two additional solutions $\tilde{v}^1(t, \lambda^*)$ and $\tilde{v}^2(t, \lambda^*)$ of (10) are calculated by solving $v_t(t, \lambda^*) = P(t, \lambda^*)v(t, \lambda^*)$ using boundary conditions adjoint to (14), and taking the adjoint of the solutions. The four *canonical eigenvectors* $v^1(t, \lambda)$, $v^2(t, \lambda)$, $\tilde{v}^1(t, \lambda)$, and $\tilde{v}^2(t, \lambda)$ satisfy

$$v^2(t, \lambda) = a(\lambda) \tilde{v}^1(t, \lambda^*) + b(\lambda) v^1(t, \lambda) \quad (15a)$$

$$\tilde{v}^2(t, \lambda) = -b^*(\lambda^*) \tilde{v}^1(t, \lambda^*) + a^*(\lambda^*) v^1(t, \lambda) \quad (15b)$$

where $a(\lambda)$ and $b(\lambda)$ do not depend on t . The NFT of the signal $q(z, t)$ is made up of two spectra:

- the *continuous spectrum* $Q_c(\lambda) = \frac{b(\lambda)}{a(\lambda)}$, for $\lambda \in \mathbb{R}$;
- the *discrete spectrum* $Q_d(\lambda_k) = \frac{b(\lambda_k)}{a_\lambda(\lambda_k)}$, for the K eigenvalues $\{\lambda_k \in \mathbb{C}^+ : a(\lambda_k) = 0\}$

where $a_\lambda = da/d\lambda$. To compute the NFT, the following relations are useful

$$a(\lambda) = \lim_{t \rightarrow \infty} v_1^2(t, \lambda) e^{j\lambda t} \quad (16a)$$

$$b(\lambda) = \lim_{t \rightarrow \infty} v_2^2(t, \lambda) e^{-j\lambda t}. \quad (16b)$$

Given a signal $q(z, t)$ propagating according to the NLSE (6), its NFT evolves in z according to the following multiplicative relations:

$$Q_c(z, \lambda) = Q_c(0, \lambda) e^{4j\lambda^2 z} \quad (17a)$$

$$\lambda_k(z) = \lambda_k(0) \quad (17b)$$

$$Q_d(z, \lambda_k) = Q_d(0, \lambda_k) e^{4j\lambda_k^2 z}. \quad (17c)$$

IV. EIGENVALUES OF HIGHER MULTIPLICITY

The relations (15) are consistent only for $\lambda \in \mathbb{R} \cup \{\lambda_k\} \cup \{\lambda_k^*\}$, because the boundary condition (14a) on $\tilde{v}^1(t, \lambda^*)$ is unbounded outside this region. As the eigenvalues λ_k come in complex conjugate pairs, the spectrum at $\lambda \in \mathbb{R} \cup \{\lambda_k\}$ is enough to determine $q(t)$ uniquely. However, to the best of our knowledge, all the work on the NFT for optical communication assumes that all the eigenvalues λ_k have multiplicity 1, i.e., the zeros of $a(\lambda)$ are simple. There has been, however, some work [12], [13] on the mathematical theory of higher multiplicity eigenvalues, which we summarize in this section.

For a multiple zero λ_k of $a(\lambda)$, we have $a_\lambda(\lambda_k) = 0$, and the above definition of the discrete spectrum is not valid anymore. If the multiplicity of the eigenvalue λ_k is L_k , we need L_k constants $q_{k0}, \dots, q_{k, (L_k-1)}$ to determine the discrete spectrum. In [13], these *norming constants* are calculated in several intermediate steps.

- *Dependency constants*

The dependency constants $\gamma_{k\ell}$, $k \in \{1, \dots, K\}$, $\ell \in \{0, \dots, L_k - 1\}$ are defined by the following equation [13, Eq. (3.1)]:

$$v^{2, (m)}(t, \lambda_k) = \sum_{\ell=0}^m \binom{m}{\ell} \gamma_{k, (m-\ell)} v^{1, (\ell)}(t, \lambda_k) \quad (18)$$

where $v^{i, (m)}$ denotes the m -th derivative of v^i with respect to λ . Taking the m -th derivative of (15), and using $a^{(\ell)}(t, \lambda) = 0$ for $\ell \leq L_k - 1$, we have

$$\gamma_{k, \ell} = b^{(\ell)}(\lambda_k). \quad (19)$$

- *Generalized residues*

The generalized residues $t_{k\ell}$, $k \in \{1, \dots, K\}$, $\ell \in \{1, \dots, L_k\}$, are the coefficients of the expansion of $1/a(\lambda) - 1$ in inverse powers of $\lambda - \lambda_k$ [13, Eq. (4.3)]:

$$\frac{1}{a(\lambda)} - 1 = \frac{t_{kL_k}}{(\lambda - \lambda_k)^{L_k}} + \dots + \frac{t_{k1}}{(\lambda - \lambda_k)} + \mathcal{O}(1). \quad (20)$$

They can be computed as

$$t_{k\ell} = \frac{1}{(L_k - \ell)!} \lim_{\lambda \rightarrow \lambda_k} \frac{d^{L_k - \ell}}{d\lambda^{L_k - \ell}} \left[\frac{(\lambda - \lambda_k)^{L_k}}{a(\lambda)} \right]. \quad (21)$$

- *Norming constants*

The norming constants $q_{k\ell}$, $k \in \{1, \dots, K\}$, $\ell \in \{0, \dots, L_k - 1\}$, are given by [13, Eq. (4.1)]:

$$q_{k\ell} = j^\ell \sum_{m=0}^{L_k - \ell - 1} \frac{b^{(m)}(\lambda_k)}{m!} t_{k,(\ell+m+1)}. \quad (22)$$

The generalization of the distance evolution equation (17c) to the case $L_k > 1$ is given by [13, Eq. (4.9)]:

$$\begin{bmatrix} q_{k,(L_k-1)}(z) & \cdots & q_{k0}(z) \\ q_{k,(L_k-1)}(0) & \cdots & q_{k0}(0) \end{bmatrix} e^{-4j\Lambda_k^2 z} \quad (23)$$

for all $k \in \{1, \dots, K\}$, where

$$\Lambda_k = \begin{pmatrix} -j\lambda_k & -1 & 0 & \cdots & 0 \\ 0 & -j\lambda_k & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -j\lambda_k & -1 \\ 0 & 0 & \cdots & 0 & -j\lambda_k \end{pmatrix} \in \mathbb{C}^{L_k \times L_k}. \quad (24)$$

We write the NFT as

$$\text{NFT} \{q(t)\} = (Q_c(\lambda), \{\lambda_k\}, \{q_{k\ell}\}). \quad (25)$$

A. Properties of the NFT with Higher Multiplicity Eigenvalues

We prove the following properties in Appendix A.

1) Phase shift:

$$\text{NFT} \{q(t)e^{j\phi_0}\} = (Q_c(\lambda)e^{-j\phi_0}, \{\lambda_k\}, \{q_{k\ell}e^{-j\phi_0}\}). \quad (26)$$

2) Time shift: if $q'(t) = q(t - t_0)$ then

$$\text{NFT} \{q'(t)\} = (Q'_c(\lambda), \{\lambda'_k\}, \{q'_{k\ell}\}) \quad (27)$$

satisfies

$$Q'_c(\lambda) = Q_c(\lambda)e^{-2j\lambda t_0} \quad (28a)$$

$$\lambda'_k = \lambda_k \quad (28b)$$

$$\begin{bmatrix} q'_{k,(L_k-1)} & \cdots & q'_{k0} \end{bmatrix} = \begin{bmatrix} q_{k,(L_k-1)} & \cdots & q_{k0} \end{bmatrix} e^{2\Lambda_k t_0}. \quad (28c)$$

3) Frequency shift:

$$\text{NFT} \{q(t)e^{-2j\omega_0 t}\} = (Q_c(\lambda - \omega_0), \{\lambda_k + \omega_0\}, \{q_{k\ell}\}). \quad (29)$$

4) Time dilation: for $T > 0$

$$\text{NFT} \left\{ \frac{1}{T} q\left(\frac{t}{T}\right) \right\} = \left(Q_c(T\lambda), \left\{ \frac{\lambda_k}{T} \right\}, \left\{ \frac{q_{k\ell}}{T^{\ell+1}} \right\} \right). \quad (30)$$

5) Parseval's theorem:

$$\int_{-\infty}^{\infty} |q(t)|^2 dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 + |Q_c(\lambda)|^2) d\lambda + 4 \sum_{k=0}^K L_k \Im \{\lambda_k\}. \quad (31)$$

V. NUMERICAL COMPUTATION OF THE (I)NFT WITH HIGHER MULTIPLICITY EIGENVALUES

Using the theory from Section IV, we extend the existing numerical algorithms that compute the (I)NFT to include multiple eigenvalues.

A. Direct NFT

Most algorithms that compute the direct NFT discretize the Zakharot-Shabat system (10) to find $a(\lambda)$ and $b(\lambda)$ from (16). Let $u = (u_1, u_2)^T$, where $u_1(t, \lambda) = v_1^2(t, \lambda)e^{j\lambda t}$ and $u_2(t, \lambda) = v_2^2(t, \lambda)e^{-j\lambda t}$. Then from (10) we have

$$u_t(t, \lambda) = \begin{pmatrix} 0 & q(t)e^{2j\lambda t} \\ -q^*(t)e^{-2j\lambda t} & 0 \end{pmatrix} u(t, \lambda) \quad (32)$$

and from (16) we have

$$a(\lambda) = \lim_{t \rightarrow \infty} u_1(t, \lambda) \quad (33a)$$

$$b(\lambda) = \lim_{t \rightarrow \infty} u_2(t, \lambda). \quad (33b)$$

To compute the NFT of $q(t)$, we discretize the time axis in the interval $t \in [t_1, t_2]$. Let $t_n = t_1 + n\epsilon$, $q_n = q(t_n)$, where $n \in \{0, \dots, N-1\}$, N is the number of samples, and $\epsilon = (t_2 - t_1)/(N-1)$ is the step size. Similarly, let $u[n] = u(t_1 + n\epsilon, \lambda)$. Starting at $u[0] = (1, 0)^T$ (see (14)), the following update step is applied iteratively:

$$u[n+1] = A[n]u[n], \quad n \in \{0, \dots, N-2\} \quad (34)$$

and we have $a(\lambda) = u_1[N-1]$ and $b(\lambda) = u_2[N-1]$. The kernel $A[n]$ varies according to the discretization algorithm. A few options are given in [4]. In this work, we consider the trapezoidal kernel proposed in [5]:

$$A[n] = \begin{pmatrix} \cos(|q_n|\epsilon) & \sin(|q_n|\epsilon)e^{j(\theta_n+2\lambda t_n)} \\ -\sin(|q_n|\epsilon)e^{-j(\theta_n+2\lambda t_n)} & \cos(|q_n|\epsilon) \end{pmatrix} \quad (35)$$

where $\theta_n = \arg q_n$. However, the following analysis is valid for any kernel $A[n]$. To obtain the norming constants $q_{k\ell}$, we need to calculate higher order λ -derivatives of $a(\lambda)$ and $b(\lambda)$. More specifically, from (22) we need the first $L_k - 1$ derivatives of $b(\lambda)$. In the case of $a(\lambda)$, we obtain an upper bound on the order of the required derivatives.

Lemma 1. *The value of $t_{k\ell}$ in (21) depends on λ_k only through the functions $a^{(m)}(\lambda_k)$ for $m \in \{L_k, \dots, 2L_k - \ell\}$*

Proof. See Appendix B. \square

For an eigenvalue of multiplicity L_k , we need to compute the first $2L_k - 1$ derivatives of $u[N-1]$. We do this by setting the initial conditions

$$u^{(m)}[0] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad m \in \{1, \dots, 2L_k - 1\} \quad (36)$$

and applying the update steps

$$u^{(m)}[n+1] = \sum_{r=0}^m \binom{m}{r} A^{(r)}[n] u^{(m-r)}[n] \quad (37)$$

where $A^{(r)}[n]$, the r -th order λ -derivative of $A[n]$, is obtained in closed form. For the trapezoidal kernel (35) we have

$$A^{(r)}[n] = (2jt_n)^r \sin(|q_n|\epsilon) \cdot \begin{pmatrix} 0 & e^{j(\theta_n+2\lambda t_n)} \\ (-1)^{r+1} e^{-j(\theta_n+2\lambda t_n)} & 0 \end{pmatrix}. \quad (38)$$

Once we have the required values of a, b and their derivatives, we use (22) and (21) to compute the norming constants.

In (21), the derivative is evaluated in closed form, and then L'Hôpital's rule is applied repetitively to obtain an expression for $t_{k\ell}$ that depends only on nonzero derivatives of a . See (75)-(77) in Appendix B for details. For $L_k = 2$, this gives

$$q_{k1} = \frac{j2b(\lambda_k)}{a_{\lambda\lambda}(\lambda_k)} \quad (39a)$$

$$q_{k0} = \frac{2b_\lambda(\lambda_k)}{a_{\lambda\lambda}(\lambda_k)} - \frac{2}{3} \frac{b(\lambda_k)a_{\lambda\lambda\lambda}(\lambda_k)}{a_{\lambda\lambda}(\lambda_k)^2} \quad (39b)$$

and for $L_k = 3$ we have

$$q_{k2} = \frac{-6b(\lambda_k)}{a_{\lambda\lambda\lambda}(\lambda_k)} \quad (40a)$$

$$q_{k1} = \frac{j6b_\lambda(\lambda_k)}{a_{\lambda\lambda\lambda}(\lambda_k)} - jb(\lambda_k) \frac{3a_{\lambda\lambda\lambda\lambda}(\lambda_k)}{2a_{\lambda\lambda\lambda}(\lambda_k)^2} \quad (40b)$$

$$q_{k0} = \frac{6b_{\lambda\lambda}(\lambda_k)}{a_{\lambda\lambda\lambda}(\lambda_k)} - b_\lambda(\lambda_k) \frac{3a_{\lambda\lambda\lambda\lambda}(\lambda_k)}{2a_{\lambda\lambda\lambda}(\lambda_k)^2} + b(\lambda_k) \frac{15a_{\lambda\lambda\lambda\lambda}(\lambda_k)^2 - 12a_{\lambda\lambda\lambda}(\lambda_k)a_{\lambda\lambda\lambda\lambda\lambda}(\lambda_k)}{20a_{\lambda\lambda\lambda}(\lambda_k)^3}. \quad (40c)$$

Forward-Backward Method: This technique was proposed in [5] to improve numerical stability. We write (34) as

$$\begin{pmatrix} a(\lambda) \\ b(\lambda) \end{pmatrix} = A[N-1] \cdots A[1]A[0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} = RL \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (41)$$

where $R = A[N-1] \cdots A[n_0]$ and $L = A[n_0-1] \cdots A[0]$, and n_0 is chosen according to some criterion to minimize the numerical error. The iterative procedure (34) is run *forward* up to $n_0 - 1$ to obtain

$$\begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = L \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} L_{11} \\ L_{21} \end{pmatrix} \quad (42)$$

and *backward* from $r[N-1] = (0, 1)^T$ down to $r[n_0-1]$:

$$\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = R^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -R_{12} \\ R_{11} \end{pmatrix}. \quad (43)$$

The kernel $A[n]^{-1}$ is used to compute (43): for the trapezoidal case this amounts to replacing ϵ with $-\epsilon$ in (35). Note that (43) is valid only for kernels with unit determinant.

Using (38), we obtain r_1, r_2, l_1, l_2 , and their derivatives up to order $2L_k - \ell$. From (41) we have

$$a(\lambda) = R_{11}L_{11} + R_{12}L_{21} \quad (44)$$

and we compute

$$a^{(\ell)}(\lambda_k) = \sum_{m=0}^{\ell} \binom{\ell}{m} \left(r_2^{(m)} l_1^{(\ell-m)} - r_1^{(m)} l_2^{(\ell-m)} \right). \quad (45)$$

To obtain $b^{(\ell)}(\lambda_k)$, note that

$$\begin{aligned} b(\lambda_k) &= R_{21}L_{11} + R_{22}L_{21} \\ &= \frac{R_{21}}{R_{11}} (R_{11}L_{11} + R_{12}L_{21}) + \frac{L_{21}}{R_{11}} \\ &= \frac{R_{21}}{R_{11}} a(\lambda_k) + \frac{L_{21}}{R_{11}} \end{aligned} \quad (46)$$

where we used $R_{22} = (1 + R_{12}R_{21})/R_{11}$. The ℓ -th derivative of the left summand in (46) is 0 for $\ell \leq L_k - 1$, because $a^{(\ell)}(\lambda_k) = 0$ for $\ell \leq L_k - 1$. Therefore, we have

$$b^{(\ell)}(\lambda_k) = \frac{d^\ell}{d\lambda^\ell} \frac{l_2}{r_2} \Big|_{\lambda=\lambda_k} \quad (47)$$

which can be written in closed form using (72) below. Equations (47) and (45), together with (22) (or (39) or (40)), let us compute the direct NFT with higher multiplicity eigenvalues from the forward-backward method.

B. Inverse NFT

The inverse NFT can be computed using the generalized version of the Gelfand-Levitan-Marchenko equation (GLME) [7]:

$$\begin{aligned} K(t, y) - \Omega^*(t + y) \\ + \int_t^\infty dx \int_t^\infty ds K(t, s)\Omega(s + x)^*\Omega(x + y) = 0. \end{aligned} \quad (48)$$

The kernel $\Omega(y)$ is given by [12]:

$$\Omega(y) = \frac{1}{2\pi} \int_{-\infty}^\infty Q_c(\lambda) e^{j\lambda y} d\lambda + \sum_{k=1}^K \sum_{\ell=0}^{L_k-1} q_{k\ell} \frac{y^\ell}{\ell!} e^{j\lambda_k y}. \quad (49)$$

The inverse NFT is then obtained as

$$q(t) = -2K(t, t). \quad (50)$$

A numerical procedure to solve the GLME (48) is given in [7, Section 4.2]: it suffices to replace $F(y)$ in the reference by $\Omega(y)$ from (49).

When there is no continuous spectrum, a closed-form expression is given in [12] for the generalized K -solitons:

$$\begin{aligned} q(z, t) \\ = -2B^H e^{-\Lambda^H t} (I + M(z, t)N(t))^{-1} e^{-\Lambda^H t + 4j(\Lambda^H)^2 z} C^H \end{aligned} \quad (51)$$

where

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \Lambda_K \end{pmatrix}, \quad (52)$$

Λ_k is given by (24),

$$B = (B_1 \cdots B_K)^T, \quad B_k = (0 \cdots 0 \ 1)^T \in \{0, 1\}^{L_k \times 1}, \quad (53)$$

$$C = (C_1 \cdots C_K), \quad C_k = (q_{k, (L_k-1)} \cdots q_{k0}), \quad (54)$$

I is an identity matrix of size $\sum_k L_k$, and

$$M(z, t) = \int_t^\infty e^{-\Lambda^H s + 4j(\Lambda^H)^2 z} C^H C e^{-\Lambda s - 4j\Lambda^2 z} ds \quad (55)$$

$$N(t) = \int_t^\infty e^{-\Lambda x} B B^H e^{-\Lambda^H x} dx. \quad (56)$$

We compute the integrals (55) and (56) numerically, and then apply (51) to obtain the inverse NFT of a purely discrete spectrum with higher multiplicity eigenvalues.

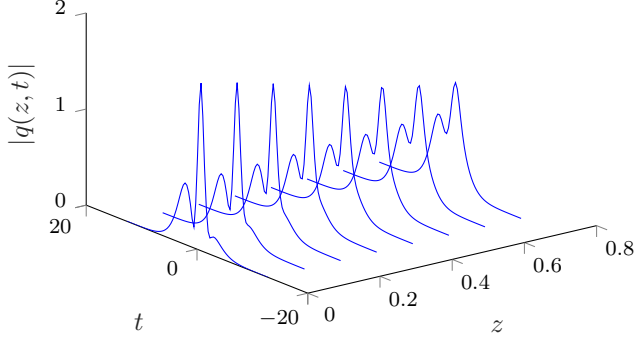


Figure 1. Propagation of a double soliton

C. Example: Double Soliton

A soliton $q(z, t)$ with a 2nd order eigenvalue at $\lambda = \xi + j\eta$, with norming constants q_{11} and q_{10} , can be derived in closed form from (51). The result is

$$q(z, t) = \frac{h(z, t)}{f(z, t)} \quad (57)$$

where

$$h(z, t) = -j4\eta e^{-j \arg q_{11}} e^{-j2\xi t} e^{-j4(\xi^2 - \eta^2)z} \left\{ e^{-X} \left[-|q_{11}|^2 (2\eta t + 8\eta(\xi + j\eta)z + 2) - \eta q_{11}^* q_{10} \right] + e^X \left[|q_{11}|^2 (2\eta t + 8\eta(\xi - j\eta)z) + \eta q_{11} q_{10}^* \right] \right\} \quad (58)$$

$$f(z, t) = |q_{11}|^2 [\cosh(2X) + 1] + 2|q_{10}\eta + q_{11}(2\eta t + 8\eta(\xi + j\eta)z + 1)|^2 \quad (59)$$

and

$$X = 2\eta t + 8\eta\xi z - \log \frac{|q_{11}|}{4\eta^2}. \quad (60)$$

We refer to this soliton as a *double soliton*. The evolution of the norming constants (23) reduces to

$$q_{11}(z) = q_{11}(0) e^{4j\lambda^2 z} \quad (61a)$$

$$q_{10}(z) = (q_{10}(0) + 8\lambda z q_{11}(0)) e^{4j\lambda^2 z}. \quad (61b)$$

Note from (61b) that a soliton with higher multiplicity eigenvalues does *not* exhibit periodic (breathing) behavior in z . Only the first norming constant $q_{k, (L_k - 1)}$ evolves periodically if the eigenvalue is purely imaginary. In particular, the ratio $|q_{10}|/|q_{11}|$ is unbounded with distance, which causes the unbounded pulse broadening that can be inferred from Figure 1.

We remark that the following closed form expression for the *center time* of the double soliton seems to be valid based on our experiments, although we have not found a proof:

$$\frac{\int_{-\infty}^{\infty} t |q(t)|^2}{\int_{-\infty}^{\infty} |q(t)|^2} = \frac{1}{2\eta} \log \left(\frac{|q_{11}|}{4\eta^2} \right). \quad (62)$$

VI. INFORMATION TRANSMISSION USING HIGHER MULTIPLICITY EIGENVALUES

We simulated a communications system with the parameters given in Table I. Three different launch signals were compared:

- a double-soliton with an eigenvalue at $\lambda = 1.25i$,
- a 2-soliton with eigenvalues $\lambda_1 = 1.5i$ and $\lambda_2 = 1i$,
- and a 1-soliton with an eigenvalue at $\lambda_0 = 2.5i$.

From (31), the three signals have the same pulse energy in the normalized domain. The 1-soliton uses multi-ring modulation on $q_0 = Q_d(\lambda_0)$ with 32 rings and 128 phases per ring. The 2-soliton has the two spectral amplitudes $q_1 = Q_d(\lambda_1)$ and $q_2 = Q_d(\lambda_2)$, while the double-soliton system has the two norming constants q_{11} and q_{10} . Both the 2-soliton and the double-soliton have 4 rings and 16 phases per spectral amplitude. With this design, the three launch signals can transmit up to 12 bits per channel use. The amplitudes of the rings have been heuristically optimized to obtain a small time-bandwidth product (TBP) of 10.5 sec·Hz that is approximately the same for the three signals. The ring amplitudes for the 1-soliton are given by

$$|q_0| \in \{0.088754 \cdot 1.6142^k : k \in \{0, \dots, 31\}\}. \quad (63)$$

The ring amplitudes for the two-soliton and double-soliton are given in Tables II and III, respectively. The phases are uniformly spaced in $[0, 2\pi)$, starting at 0 for q_0 and q_{11} , at $\pi/128$ for q_2 , and at $\pi/16$ for q_{10} . The optimal criterion for choosing ring amplitudes has not yet been found, but expressions such as (62) suggest that geometric progressions are better suited than arithmetic progressions.

The free normalization parameter T_0 in (5) was used to obtain the desired powers. This means that the pulse duration and bandwidth are different for each power value. This is unavoidable in soliton systems. The TBP of 10.5 is, however, constant for all values of power.

Propagation according to (1) was simulated using the split-step Fourier method. In all systems, the transmitter used closed-form expressions to generate the solitons, and the receiver used forward-backward computation with the trapezoidal kernel to obtain the norming constants. Equalization was performed by inverting (23). The mutual information of the transmitted and received symbols was measured and normalized by the TBP to obtain the spectral efficiency. In the 2-soliton and double-soliton systems, the mutual information was computed jointly for the two spectral amplitudes,

$$I(q_1^{(TX)}, q_2^{(TX)}; q_1^{(RX)}, q_2^{(RX)}) \quad (64)$$

where $q_k^{(TX)}$ refers to the transmitted symbols and $q_k^{(RX)}$ refers to the received symbols after equalization. The effective number of transmitted symbols was 614440.

Figure 2 shows the spectral efficiency for the three systems. The double soliton performs better than the 1-soliton, but worse than the 2-soliton. This is expected: the higher order derivatives of $a(\lambda)$ in (39) lead to a loss of accuracy. However, the experiment demonstrates that the generalized NFT with multiple zeros can be used to transmit information.

Table I
SIMULATION PARAMETERS

Parameter	Symbol	Value
Dispersion coefficient	β_2	$-21.667 \text{ ps}^2/\text{km}$
Nonlinear coefficient	γ	$1.2578 \text{ W}^{-1}\text{km}^{-1}$
Fiber length	z	4000 km
Noise spectral density	N_{ase}	$6.4893 \cdot 10^{-24} \text{ Ws/m}$

Table II
RING AMPLITUDES FOR THE 2-SOLITON SYSTEM

$ q_1 (\lambda = 1.5i)$	2.5355	2.8364	3.1730	3.5496
$ q_2 (\lambda = 1i)$	0.2662	1.0211	3.9173	15.0283

VII. CONCLUSION

We started from the theory in [12], [13], and proved some properties of the generalized NFT that are useful for communications. We designed and implemented algorithms to compute the generalized NFT, and we numerically demonstrated the potential of higher multiplicity eigenvalues for information transmission. With this, we extend the class of signals that admit an NFT, and we show that there are additional degrees of freedom for NFT-based optical communications systems.

There are several directions for future work. For example, an extension of the Darboux algorithm to the generalized NFT would speed up the computation of the INFT. More insight into the duration, bandwidth and robustness to noise of multiple eigenvalue signals would be useful.

APPENDIX A

PROOF OF THE PROPERTIES OF THE NFT WITH HIGHER MULTIPLICITY EIGENVALUES

In the following, all primed variables (a') refer to the spectral functions of the shifted signal $q'(t)$.

1) *Phase shift*: replacing q with $qe^{j\phi_0}$ in (32), we have $a'(\lambda) = a(\lambda)$ and $b'(\lambda) = b(\lambda)e^{-j\phi_0}$. The property then follows from (17a) and (22).

2) *Time shift*: the change of variable $t \rightarrow t - t_0$ in (10) proves that $a'(\lambda) = a(\lambda)e^{j\lambda t_0}$, and $b'(\lambda) = b(\lambda)e^{-j\lambda t_0}$. The expressions (28a) and (28b) follow immediately. Defining

$$t_{k\ell}(\lambda) = \frac{1}{(L_k - \ell)!} \frac{d^{L_k - \ell}}{d\lambda^{L_k - \ell}} \left[\frac{(\lambda - \lambda_k)^{L_k}}{a(\lambda)} \right] \quad (65)$$

we have

$$t_{kL_k}(\lambda) = \frac{(\lambda - \lambda_k)^{L_k}}{a(\lambda)} \quad (66a)$$

$$t_{k\ell}(\lambda) = \frac{1}{(L_k - \ell)!} \frac{d^{L_k - \ell}}{d\lambda^{L_k - \ell}} t_{kL_k}(\lambda). \quad (66b)$$

Using $a' = ae^{j\lambda t_0}$, we have $t'_{kL_k}(\lambda) = t_{kL_k}(\lambda)e^{-j\lambda t_0}$ and, from (66b),

$$t'_{k\ell}(\lambda) = e^{-j\lambda t_0} \sum_{i=0}^{L_k - \ell} \frac{1}{i!} (-jt_0)^i t_{k,\ell+i}(\lambda). \quad (67)$$

From $b' = be^{-j\lambda t_0}$ we have

$$(b')^{(\ell)}(\lambda) = e^{-j\lambda t_0} \sum_{i=0}^{\ell} \binom{\ell}{i} (-jt_0)^{\ell-i} b^{(i)}(\lambda). \quad (68)$$

Table III
RING AMPLITUDES FOR THE DOUBLE-SOLITON SYSTEM

$ q_{11} $	5.3785	5.9449	6.5708	7.2627
$ q_{10} $	34.3750	39.3496	45.0440	51.5625

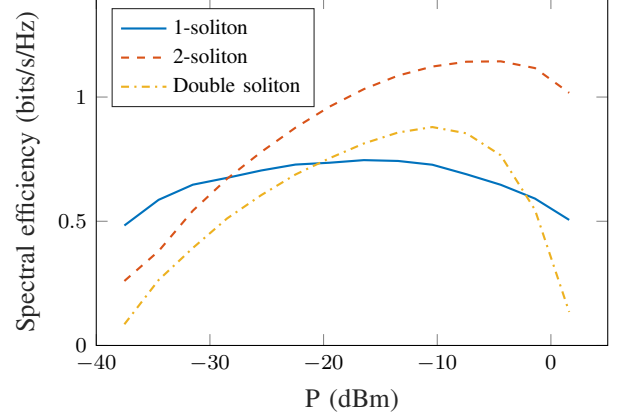


Figure 2. Spectral efficiency of three solitonic signals

Using (67) and (68) in (22), we have

$$q'_{k\ell} = j^\ell e^{-2j\lambda_k t_0} \sum_{m=0}^{L_k - \ell - 1} \frac{1}{m!} \sum_{i=0}^m \binom{m}{i} b^{(i)} \cdot (-jt_0)^{m-i} \sum_{r=0}^{L_k - \ell - m - 1} \frac{1}{r!} (-jt_0)^r t_{k,\ell+m+r+1}. \quad (69)$$

The change of variable $m = u + i - r$ yields

$$\begin{aligned} q'_{k\ell} &= e^{-2j\lambda_k t_0} \sum_{u=0}^{L_k - \ell - 1} (-t_0)^u \left(\sum_{r=0}^u \frac{1}{r! (u-r)!} \right) \\ &\quad \sum_{i=0}^{L_k - \ell - u - 1} \frac{j^{\ell+u}}{i!} b^{(i)} t_{k,\ell+u+i+1} \\ &= e^{-2j\lambda_k t_0} \sum_{u=0}^{L_k - \ell - 1} \frac{1}{u!} (-2t_0)^u q_{k,\ell+u} \end{aligned} \quad (70)$$

which is the same as (28c).

3) *Frequency shift*: using the change of variable $\lambda \rightarrow \lambda - \omega_0$ in (32), we have $a'(\lambda) = a(\lambda - \omega_0)$ and $b'(\lambda) = b(\lambda - \omega_0)$, from which the property follows.

4) *Time dilation*: the change of variable $t \rightarrow t/T$ in (32) proves that $a'(\lambda) = a(T\lambda)$ and $b'(\lambda) = b(T\lambda)$. Using this in (21) and changing $\lambda \rightarrow T\lambda$ shows that $t'_{k\ell} = t_{k\ell}/T^\ell$. Together with (22), this proves (30).

5) *Parseval's theorem*: this is a particular case ($n = 0$) of the more general trace formula:

$$\begin{aligned} C_n &= \frac{1}{\pi} \int_{-\infty}^{\infty} (2j\lambda)^n \log \left(1 + |Q_c(\lambda)|^2 \right) d\lambda \\ &\quad + \frac{4}{n+1} (2j)^n \sum_{k=0}^K L_k \Im \{ \lambda_k^{n+1} \} \end{aligned} \quad (71)$$

where C_n are the constants of motion, of which C_0 is the signal energy (31). The proof for simple eigenvalues is given

in [2, Sec. 1.6]. The result is extended to multiple eigenvalues by allowing several ζ_m in [2, Eq. (1.6.18)] to be equal.

APPENDIX B
PROOF OF LEMMA 1

Applying Faà di Bruno's formula [15, pp. 43-44] to $1/a(\lambda)$, and then the generalized product rule to $c(\lambda) \cdot (1/a(\lambda))$, for an arbitrary $c(\lambda)$, we can write a quotient rule for higher order derivatives:

$$\frac{d^n c}{d\lambda^n a} = \sum_{m=0}^n \left[\binom{n}{m} c^{(n-m)} \sum_{\mathbf{p} \in \mathcal{P}(m)} \frac{(-1)^{|\mathbf{p}|} m!}{p_1! 1!^{p_1} \dots p_m! m!^{p_m}} \frac{|\mathbf{p}|!}{a^{|\mathbf{p}|+1}} \prod_{i=1}^m \left(a^{(i)} \right)^{p_i} \right]. \quad (72)$$

Recall that $a^{(i)}$ denotes an i -th order derivative. Here, $\mathcal{P}(m)$ denotes the set of partitions \mathbf{p} of m :

$$\mathbf{p} = [p_1, \dots, p_m], \quad \sum_{i=1}^m i p_i = m, \quad p_i \in \mathbb{N} \cup \{0\} \quad (73)$$

and $|\mathbf{p}| = \sum_{i=1}^m p_i$ is the *cardinality* of \mathbf{p} . Using (72) in (21) we have

$$t_{k\ell} = \lim_{\lambda \rightarrow \lambda_k} \frac{g(\lambda)}{a(\lambda)^{L_k - \ell + 1}} \quad (74)$$

where

$$g(\lambda) = \sum_{m=0}^{L_k - \ell} \left[\binom{L_k - \ell}{m} \frac{L_k!}{(\ell + m)!} (\lambda - \lambda_k)^{\ell + m} \cdot \sum_{\mathbf{p} \in \mathcal{P}(m)} \frac{(-1)^{|\mathbf{p}|} m!}{p_1! 1!^{p_1} \dots p_m! m!^{p_m}} |\mathbf{p}|! a^{L_k - \ell - |\mathbf{p}|} \prod_{i=1}^m \left(a^{(i)} \right)^{p_i} \right]. \quad (75)$$

Note that a has a zero of order L_k , and therefore $a^{(m)}(\lambda_k) = 0$ for $m \in \{0, \dots, L_k - 1\}$. To compute $t_{k\ell}$, we repeatedly apply L'Hôpital's rule until the numerator and the denominator become nonzero in the limit:

$$t_{k\ell} = \frac{g^{(r)}(\lambda_k)}{[d^r a(\lambda)^{L_k - \ell + 1} / d\lambda^r]_{\lambda = \lambda_k}}. \quad (76)$$

The number r of times we need to differentiate is equal to the order of the zero in the denominator:

$$r = L_k (L_k - \ell + 1). \quad (77)$$

The summands in $g^{(r)}(\lambda)$ are of the form

$$g_s(\lambda) = K_s \frac{d^{(L_k - \ell + 1)L_k}}{d\lambda^{(L_k - \ell + 1)L_k}} (\lambda - \lambda_k)^{\ell + m} a^{L_k - \ell - |\mathbf{p}|} \prod_{i=1}^m \left(a^{(i)} \right)^{p_i} \quad (78)$$

where s is an index, and K_s is a constant independent of λ . If we apply the generalized product rule to (78), we see that the only summands that will become nonzero for $\lambda = \lambda_k$ are those that apply $\ell + m$ differentiations on the factor $(\lambda - \lambda_k)^{\ell + m}$. This leaves $(L_k - \ell + 1)L_k - \ell - m$ differentiations for the other factors. Note that the derivative of a term such as

$$a^{L_k - \ell - |\mathbf{p}|} \prod_{i=1}^m \left(a^{(i)} \right)^{p_i} \quad (79)$$

is a sum of terms of the same form. Each new term has the same amount $\sum_i p_i$ of a -factors as the original (an a -factor here refers to a or one of its derivatives), and where the number of differentiations $\sum_i i p_i$ in the a -factors is increased by 1. We conclude that $g_s(\lambda)$ is made up of summands that contain

- $L_k - \ell - |\mathbf{p}| + \sum_i p_i = L_k - \ell$ a -factors that have
- $(L_k - \ell + 1)L_k - \ell - m + \sum_i i p_i = (L_k - \ell + 1)L_k - \ell$ differentiations.

Only the summands where all the a -factors are differentiated at least L_k times will be nonzero. We want to know the highest-order derivative of a that appears in a nonzero term. The worst case occurs when we have $L_k - \ell - 1$ a -factors with an L_k -th order derivative. The remaining a -factor must have a derivative of order $[(L_k - \ell + 1)L_k - \ell] - [L_k(L_k - \ell - 1)] = 2L_k - \ell$.

ACKNOWLEDGMENT

The author wishes to thank Prof. G. Kramer and B. Leible for useful comments and proofreading the paper.

REFERENCES

- [1] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity Limits of Optical Fiber Networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb 2010.
- [2] M. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. Society for Industrial and Applied Mathematics, 1981. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/1.9781611970883>
- [3] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Part I: Mathematical Tools," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4328, July 2014.
- [4] —, "Information Transmission Using the Nonlinear Fourier Transform, Part II: Numerical Methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, July 2014.
- [5] V. Aref, "Control and Detection of Discrete Spectral Amplitudes in Nonlinear Fourier Spectrum," *ArXiv e-prints*, May 2016.
- [6] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Part III: Spectrum Modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, July 2014.
- [7] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," *Opt. Express*, vol. 22, no. 22, pp. 26720–26741, Nov 2014. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-22-26720>
- [8] M. I. Yousefi and X. Yangzhang, "Linear and Nonlinear Frequency-Division Multiplexing," in *2016 Eur. Conf. Optical Commun. (ECOC)*, Sept 2016, pp. 1–3.
- [9] S. T. Le, H. Buelow, and V. Aref, "Demonstration of 64 0.5Gbaud nonlinear frequency division multiplexed transmission with 32QAM," in *2017 Optical Fiber Commun. Conf. and Exhib. (OFC)*, March 2017, pp. 1–3.
- [10] V. Aref, H. Blow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in *2015 Eur. Conf. Optical Commun. (ECOC)*, Sept 2015, pp. 1–3.
- [11] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue Communication," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3110–3117, July 2016.
- [12] T. Aktosun, F. Demontis, and C. van der Mee, "Exact solutions to the focusing nonlinear Schrödinger equation," *Inverse Problems*, vol. 23, no. 5, p. 2171, 2007. [Online]. Available: <http://stacks.iop.org/0266-5611/23/i=5/a=021>
- [13] T. N. B. Martinez, "Generalized inverse scattering transform for the nonlinear Schrödinger equation for bound states with higher multiplicities," *Electronic J. Differential Equations*, vol. 2017, no. 179, pp. 1–15, July 2017.
- [14] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. Academic Press, October 2012.
- [15] L. Arbogast, *Du calcul des dérivations*, 1800.