Fast Numerical Nonlinear Fourier Transforms

Sander Wahls, *Member, IEEE,* and H. Vincent Poor, *Fellow, IEEE*

arXiv:1402.1605v3 [cs.IT] 10 Sep 2015

Abstract—The nonlinear Fourier transform, which is also known as the forward scattering transform, decomposes a periodic signal into nonlinearly interacting waves. In contrast to the common Fourier transform, these waves no longer have to be sinusoidal. Physically relevant waveforms are often available for the analysis instead. The details of the transform depend on the waveforms underlying the analysis, which in turn are specified through the implicit assumption that the signal is governed by a certain evolution equation. For example, water waves generated by the Korteweg-de Vries equation can be expressed in terms of cnoidal waves. Light waves in optical fiber governed by the nonlinear Schrödinger equation (NSE) are another example. Nonlinear analogs of classic problems such as spectral analysis and filtering arise in many applications, with information transmission in optical fiber, as proposed by Yousefi and Kschischang, being a very recent one. The nonlinear Fourier transform is eminently suited to address them – at least from a theoretical point of view. Although numerical algorithms are available for computing the transform, a "fast" nonlinear Fourier transform that is similarly effective as the fast Fourier transform is for computing the common Fourier transform has not been available so far. The goal of this paper is to address this problem. Two fast numerical methods for computing the nonlinear Fourier transform with respect to the NSE are presented. The first method achieves a runtime of $O(D^2)$ floating point operations, where D is the number of sample points. The second method applies only to the case where the NSE is defocusing, but it achieves an $O(D \log^2 D)$ runtime. Extensions of the results to other evolution equations are discussed as well.

Index Terms—Nonlinear Fourier Transform, Forward Scattering Transform, Nonlinear Schrödinger Equation, Fast Algorithms

I. INTRODUCTION

Consider a smooth signal $q : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{C}$ governed by the *nonlinear Schrödinger equation (NSE)*[1]

$$i\,\partial_t q + \partial_{xx} q + 2\kappa |q|^2 q = 0 \tag{1}$$
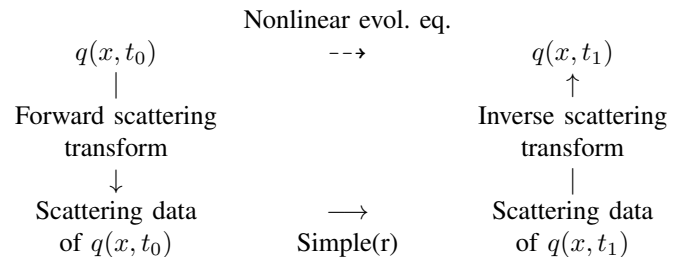
subject to a periodic boundary condition

$$q(x,t) \equiv q(x+\ell,t), \quad \ell > 0. \tag{2}$$

The constant $\kappa \in \{\pm 1\}$ determines whether the NSE is called *focusing* $(+)$ or *defocusing* $(-)$. The NSE describes several physically relevant phenomena. The complex envelope of a waveform in an optical fiber with perfect loss compensation

[1]Sometimes, the NSE (1) is given in the form $i\,\partial_t u = \partial_{xx} u + 2\kappa |u|^2 u$. This form is equivalent to (1) whenever $u = \bar{q}$. Furthermore, the roles of the arguments x and t are sometimes interchanged, e.g., when the NSE describes optical fiber. The spatial variable is then commonly denoted by z instead of x.

evolves according to the NSE [1, Ch. 6.1]. The focusing case corresponds to a fiber with anomalous dispersion and admits *bright solitons* ("particle-like waves") as solutions [1, Ch. 5.1]. Bright solitons are localized waves that remain unchanged after interactions with other bright solitons. They are often employed to encode information in optical communications [1, Ch. 4]. The defocusing case, which describes a fiber with normal dispersion, cannot be solved by bright solitons. However, it admits solutions in the form of "moving holes" in an otherwise constant signal, so-called *dark solitons* [1, Ch. 5.4]. The use of dark solitons for optical communications has been investigated much less than for bright solitons, but the possibility of using them for optical communications has been demonstrated experimentally [2], [3, p. 153ff]. The NSE also provides a model for the evolution of deep water waves [4].

It was a pleasant surprise when Zakharov and Shabat [5] showed that the NSE (for non-periodic signals that decay sufficiently rapidly as $|x| \to \infty$) can be solved in closed form using what is known as the inverse scattering method. This method was initially developed by Gardner et al. [6] for the solution of the Korteweg-de Vries equation. Not much later, Ablowitz et al. [7] extended the inverse scattering method to a wide range of evolution equations. The method is usually called the inverse scattering method because the main tools used in its derivation have their roots in physics, where they are used to analyze how particles behave based on their interactions with a scatterer. However, it can also be interpreted as a generalization of the Fourier method for the solution of linear evolution equations [7]. The direct time-evolution of signals governed by such equations can be complicated, but the time-evolution of their Fourier transforms often is simple. The inverse scattering method exploits the same principle:

$$
\begin{array}{ccc}
& \text{Nonlinear evol. eq.} & \\
q(x,t_0) & \dashrightarrow & q(x,t_1) \\
\big| & & \uparrow \\
\text{Forward scattering} & & \text{Inverse scattering} \\
\text{transform} & & \text{transform} \\
\downarrow & & \big| \\
\text{Scattering data} & \longrightarrow & \text{Scattering data} \\
\text{of } q(x,t_0) & \text{Simple(r)} & \text{of } q(x,t_1)
\end{array}
$$

The forward scattering transform, which represents the signal in an equivalent form called scattering data, can be seen as a generalization of the Fourier transform because the scattering data essentially reduces to the Fourier transform of the signal whenever the amplitude of the signal, and hence the nonlinear term in the NSE, becomes small. Therefore it is also known as the *nonlinear Fourier transform (NFT)* in the literature.

Our interest in the NFT stems from three recent papers of Yousefi and Kschischang on optical fiber communications [8], [9], [10]. The optical fiber channel suffers from several

nonlinearities, most of which are captured by the NSE. Current optical communication systems treat the interference between multiple users that is caused by the nonlinearities in a fiber as random noise. Motivated by the fact that the data rates in current optical communication systems are close to the capacity of that approach [11], [12], Kschischang and Yousefi proposed a new method[2] called *nonlinear frequency division multiplexing (NFDM)* [8]. NFDM can be interpreted as a nonlinear variant of common *orthogonal frequency-division multiplexing (OFDM)* [10]. The basic idea is to generate the information-bearing signal in the scattering domain in order to exploit the fact that the spatial evolution of the scattering data is simpler than that of the original signal in the time domain. One particular advantage of NFDM is that multi-user interference can be avoided. At the receiver, the NFT is used to recover the information. Yousefi and Kschischang investigated several numerical methods in order to realize the NFT [9], but found that a computationally efficient fast NFT is still lacking [9, Sec. VII]. The goal of this paper is to address this problem.

The authors have recently established the first fast nonlinear Fourier transform for non-periodic signals governed by the focusing NSE that decay sufficiently rapidly as $|x| \to \infty$ [15]. The same boundary conditions have been used by Yousefi and Kschischang, but periodic boundary condition actually seem to be more appropriate because one may use a cyclic prefix [16, p. 156]. In this paper, fast NFTs for signals governed by the periodic NSE will be developed. While NFTs can be derived for signals governed by many other different evolution equations as well [7], we will mainly restrain ourselves to the NSE for the sake of clarity. (Extensions to non-periodic signals and other evolution equations will also be discussed, but only briefly.) Finally, let us note that fast NFTs are of interest also in other areas. For example, they have been used to analyze and filter water waves [17], [18], [19], [20]. Measurements of oceanic data often contain up to 10.000 data points [17, p. 95]. The high complexity of the nonlinear Fourier transform has also been bemoaned in the analysis of plasma waves [21, Sec. 8].

The contributions of this paper are as follows. In Secs. II and III, finite band solutions to the NSE are introduced and a comprehensive survey of the relevant theoretical results on the nonlinear Fourier transform, which is a method for extracting the parameters of a finite band solution from measured data, is given. The so-called monodromy matrix, which is an analytic matrix-valued function, turns out to be the main tool in the derivation of the nonlinear Fourier transform. The fast algorithms in this paper will require a rational approximation of the monodromy matrix. The authors have recently proposed a framework for obtaining rational approximations of the analog of the monodromy matrix in the case of rapidly decaying signals [15]. In Sec. IV, this framework is carried over to the periodic case and extended by the introduction of some new coordinate transforms. Then, in Sec. V, the fast algorithms presented in [15] are carried over to the periodic case and

compared to other current numerical approaches to find the nonlinear Fourier transform. In Sec. VI, a new alternative fast algorithm for the defocusing NSE is introduced and compared. Some numerical examples are presented in Sec. VII. Extensions of our results to non-periodic signals and signals governed by other evolution equations are discussed in the Secs. VIII and IX, respectively. Finally, the paper is concluded in Sec. X.

Notation

Real numbers: $\mathbb{R}$; $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} : x \geq 0\}$; Complex numbers: $\mathbb{C}$; Integers: $\mathbb{Z}$; $\sqrt{\cdot}$: Canonical square root (i.e., positive signs are preferred); $i := \sqrt{-1}$; Euler's number: e; Real part: $\text{Re}(\cdot)$; Imaginary part: $\text{Im}(\cdot)$; Complex conjugate: $\overline{(\cdot)}$; Natural logarithm: $\ln(\cdot)$; Floor function: $\lfloor \cdot \rfloor$; Absolute value: $|\cdot|$; Adjoint matrix: $(\cdot)^*$; Matrix trace: $\text{tr}\,\cdot$; Matrix exponential: $\text{expm}(\cdot)$; Matrix product: $\prod_{k=1}^{K} \mathbf{A}_k := \mathbf{A}_K \mathbf{A}_{K-1} \times \cdots \times \mathbf{A}_1$; Matrix element in the ith column and jth row: $[\cdot]_{i,j}$; Derivative w.r.t. a variable u: ∂_u; $\partial_{uv} := \partial_u \partial_v$; Equal for all arguments: $\equiv$; Absolute value of the largest coefficient of a $\mathbb{C}^{m \times n}$-valued polynomial $\mathbf{p}(z) = \sum_{k=0}^{K} \mathbf{p}_k z^k$: $|\text{largest coefficient}(\mathbf{p})| = \max\{|[\mathbf{p}_k]_{i,j}| : k = 1, \ldots, K, i = 1, \ldots, m, j = 1, \ldots, n\}$; Degree of $\mathbf{p}$: $\deg(\mathbf{p}) = \max\{k = 1, \ldots, K : \mathbf{p}_k \neq \mathbf{0}\}$

II. FINITE BAND SOLUTIONS

In this section, the theory of so-called *finite-band solutions* to the NSE (1) will be reviewed [22]. Finite-band solutions have explicit descriptions which rely only on a finite number of parameters. These parameters will constitute the *scattering data*, and finding those parameters from an initial condition $q(x, t_0)$ will constitute the non-linear Fourier transform as discussed in the following section. Finite-band solutions can be loosely thought of as the nonlinear analog to a conventional Fourier series expansion with only finitely many non-zero terms. By restricting the exposition to finite-band solutions, many questions regarding convergence can be avoided. We remark that this means no significant loss of generality since any sufficiently smooth periodic solution to the NSE can be approximated arbitrarily well on any fixed finite time interval by a periodic finite-band solution [23]. In the literature, two different types of finite band solutions can be found. Since each type has its own advantages and disadvantages, both of them will be reviewed in the following.

A. The Finite-Band Solutions of Kotlyarov and Its

A solution $q(x, t)$ to the NSE (1) is called a *finite-band solution* in the sense of Kotlyarov and Its [24], [25] (also see [26], [27]) if there exist finitely many of the following parameters:

$$\textit{Main spectrum:} \quad \lambda_1, \ldots, \lambda_{2N} \in \mathbb{C}, \qquad (3)$$
$$\textit{Initial auxiliary spectrum:} \quad \mu_j(x_0, t_0) \in \mathbb{C}, \qquad (4)$$
$$\textit{Initial amplitude}[3]: \quad q(x_0, t_0) \in \mathbb{R}_{\geq 0}, \qquad (5)$$
$$\textit{Riemann sheet indices:} \quad \sigma_j(x_0, t_0) \in \{\pm 1\}, \qquad (6)$$
$$j = 1, \ldots, N-1;$$

---

[2]Hasegawa [13] had proposed to embed data exlusively in the solitonic part of the nonlinear Fourier spectrum already in 1993, but his proposal did not receive much attention. Recently, Prilepsky et al. [14] proposed a system that utilizes exclusively the non-solitonic part of the spectrum.

such that $q(x,t)$ is generated through the following system of coupled partial differential equations: [4]

$$\partial_x \ln q = 2\,\mathrm{i} \left(\sum_{j=1}^{N-1} \mu_j - \frac{1}{2} \sum_{k=1}^{2N} \lambda_k \right), \tag{7}$$

$$\partial_t \ln q = 2\,\mathrm{i} \left(\sum_{\substack{j,k=1 \\ j>k}}^{2N} \lambda_j \lambda_k - \frac{3}{4} \left(\sum_{k=1}^{2N} \lambda_k \right)^2 \right)$$
$$+ 4\,\mathrm{i} \left(\frac{1}{2} \left(\sum_{k=1}^{2N} \lambda_k \right) \left(\sum_{j=1}^{N-1} \mu_j \right) - \sum_{\substack{j,k=1 \\ j>k}}^{N-1} \mu_j \mu_k \right) \tag{8}$$

$$\partial_x \mu_j = \frac{-2\,\mathrm{i}\,\sigma_j \sqrt{\prod_{k=1}^{2N} (\mu_j - \lambda_k)}}{\prod_{\substack{m=1 \\ m \neq j}}^{N-1} (\mu_j - \mu_m)}, \tag{9}$$

$$\partial_t \mu_j = -2 \left(\sum_{\substack{m=1 \\ m \neq j}}^{N-1} \mu_m - \frac{1}{2} \sum_{k=1}^{2N} \lambda_k \right) \partial_x \mu_j. \tag{10}$$

The generated functions $\mu_j(x,t)$ are known as the *auxiliary spectrum*. The associated *Riemann sheet indices* $\sigma_j(x,t)$ stay constant most of the time. Changes of sign occur if a $\mu_j(x,t)$ reaches one of the λ_k.[5] These changes of sign result in a wave-like motion of the $\mu_j(x,t)$ that will oscillate between a pair of main spectral points λ_k and λ_l, $k \neq l$. The $\mu_j(x,t)$ are also known as *hyperelliptic modes* in the literature [19]. It may happen that a hyperelliptic mode is constant because it is trapped between two repeating points in the main spectrum: $\mu_j(x,t) \equiv \lambda_k \equiv \lambda_l$ for $k \neq l$. The contributions of μ_j, λ_k and λ_l to (7)–(10) cancel each other out in that case, which implies these parameters do not contribute to the shape of the function $q(x,t)$. Such parameters are called *degenerate*.

Remark 1 (An assumption). The representation (7)–(10) is not unique because it is possible to add degenerate parameters without changing the solution. In order to simplify the later exposition, we shall therefore assume (like, e.g, in [29]) that no point in the main spectrum appears more than once:

$$k \neq l \implies \lambda_k \neq \lambda_l. \tag{11}$$

This condition in particular precludes the existence of degenerate points in the main spectrum.

At this point, it is interesting to note that $q(x,t)$ generated through (7)–(10) does not depend on whether the NSE is defocusing or focusing (i.e., $\kappa = -1$ or $+1$). This suggests that not every choice of the parameters (3)–(5) will lead to a solution of the NSE, and this is indeed the case. The following theorem can be used to check whether a given set of parameters results in a (probably non-periodic) solution to the NSE.

Theorem 2 ([26]; [27], Thm. 2.1). *The function q generated by the Eqs. (7)–(10) solves the NSE (1), not necessarily subject to the boundary condition (2), if and only if the functions*

$$g_z(x,t) := \mathrm{i}\, q(x,t) \prod_{j=1}^{N-1} (z - \mu_j(x,t)), \tag{12}$$

$$h_z(x,t) := \mathrm{i}\, \kappa \bar{q}(x,t) \prod_{j=1}^{N-1} (z - \bar{\mu}_j(x,t)), \tag{13}$$

$$P(z) := \prod_{k=1}^{2N} (z - \lambda_k) \tag{14}$$

are such that the "function-valued function"

$$z \mapsto f_z(x,t) := \sqrt{P(z) + g_z(x,t) h_z(x,t)} \tag{15}$$

is a polynomial of finite degree.

The functions (12), (13) and (15) are known as the *squared eigenfunctions* in the literature [29], [28]. They play a fundamental role in the analysis of finite band signals because their spatial and temporal evolution is remarkably simple.[6] Theorem 2 can also serve as a starting point for generating solutions to the NSE. This will be discussed later in Sec. II-C.

Let us now illustrate the concepts introduced in this section so far with a simple example.

Example 3 (Periodic one-band solution, $N = 1$). Theorem 2 shows that $q(x,t)$ will solve the NSE if and only if $f_z^2 = P(z) + g_z h_z$ for some unknown polynomial $z \mapsto f_z = \beta z - \gamma$:

$$(\beta z - \gamma)^2 = (z - \lambda_1)(z - \lambda_2) - \kappa |q|^2. \tag{16}$$

By comparing the coefficients of the polynomials (with respect to z) on both sides of (16), one finds $\beta = \pm 1$ and $\gamma = \beta(\lambda_1 + \lambda_2)/2$. For both choices of γ, a comparison of the constant terms in (16) results in the condition

$$\left(\frac{\lambda_1 - \lambda_2}{2} \right)^2 = -\kappa |q|^2. \tag{17}$$

Since we are in the one-band case, this condition can also be obtained directly. Solving (7) and (8) for $N = 1$ leads to

$$q(x,t) = q(x_0,t_0)\, \mathrm{e}^{-\mathrm{i}(\lambda_1+\lambda_2)(x-x_0)}\, \mathrm{e}^{2\,\mathrm{i}\left(\lambda_1\lambda_2 - \frac{3}{4}(\lambda_1+\lambda_2)^2\right)(t-t_0)}.$$

Direct substitution of this function into (1) shows again that it solves the NSE if and only if (17) is satisfied. Also note that the function is periodic in x with period ℓ [see (2)] if and only if $\mathrm{e}^{-\mathrm{i}(\lambda_1+\lambda_2)\ell} = 1$, or, equivalently, $\frac{\ell}{2\pi}(\lambda_1 + \lambda_2) \in \mathbb{Z}$.

Finally, we note that finite-band solutions have a closed-form expression which closely resembles the expression which arises when *Hirota's method* is used to solve the NSE with vanishing boundary conditions [10, Sec. III.B].

---

[3] The restriction of $q(x_0,t_0)$ to be non-negative incurs no loss of generality because if q solves the NSE (1) then so does $\mathrm{e}^{\mathrm{i}\,\varphi} q$ for any $\varphi \in [0, 2\pi)$.

[4] For $N = 1$, empty sums in (7)–(10) imply zeros and empty products ones.

[5] Technically, the μ_j evolve on a two-sheeted Riemann surface specified by the λ_k [28, Apdx.]. The sheet index indicates the current sheet.

[6] There exist matrices $\boldsymbol{\Theta}_z(x,t)$ and $\boldsymbol{\Psi}_z(x,t)$, independent of $\boldsymbol{\theta}_z^T := \begin{bmatrix} f_z & g_z & h_z \end{bmatrix}$, such that $\partial_x \boldsymbol{\theta}_z = \boldsymbol{\Theta}_z \boldsymbol{\theta}_z$ and $\partial_t \boldsymbol{\theta}_z = \boldsymbol{\Psi}_z \boldsymbol{\theta}_z$ [24], [26, Eqs. 9+12]. An interesting consequence is that a finite band signal $q(x,t)$ satisfies the NSE (1) if and only $\partial_{xt} \boldsymbol{\theta}_z = \partial_{tx} \boldsymbol{\theta}_z$ [29].

Remark 4 (Closed-form solution). Finite-band solutions of the form (7)–(10) can be given in closed form [25]. The *Riemann theta function* with respect to a $d \times d$ matrix $\boldsymbol{\tau}$ is given by

$$\Theta_{\boldsymbol{\tau}}(\mathbf{z}) := \sum_{\mathbf{m} \in \mathbb{Z}^d} e^{2\pi\,\mathrm{i}\,\mathbf{m}^T \mathbf{z} + \pi\,\mathrm{i}\,\mathbf{m}^T \boldsymbol{\tau} \mathbf{m}}, \quad \mathbf{z} \in \mathbb{C}^d.$$

The series converges absolutely for all $\mathbf{z}$ whenever $\boldsymbol{\tau}$ is symmetric with positive definite imaginary part. With a suitably chosen $\boldsymbol{\tau}$, vectors $\mathbf{k}$, $\boldsymbol{\omega}$, and $\boldsymbol{\delta}^{\pm}$, and scalars k_0 and ω_0, any finite band solution can be written as [28, Eq. (A31)]

$$q(x,t) = q(x_0, t_0) e^{\mathrm{i}\,k_0 x - \mathrm{i}\,\omega_0 t} \frac{\Theta_{\boldsymbol{\tau}}(\frac{\pi}{2}(\mathbf{k}x + \boldsymbol{\omega}t + \boldsymbol{\delta}^-))}{\Theta_{\boldsymbol{\tau}}(\frac{\pi}{2}(\mathbf{k}x + \boldsymbol{\omega}t + \boldsymbol{\delta}^+))}. \quad (18)$$

The parameters (3)–(6) are the starting point if one wants to compute this representation [28, Apdx.] (also see [19], [27].) A large family of parameters such that (18) leads to periodic solutions of the NSE (1) has been given in [30]. An explicit parametrization of all parameters that result in a periodic solution is given in [31, Ch. 3]. A necessary and sufficient condition for periodic solutions with respect to a discrete version of the NSE can be found in [32, Thm. 5.2]. It should be mentioned that the naive evaluation of (18) becomes infeasible for larger d due to the curse of dimensionality. Numerical methods for the efficient evaluation of (certain) theta functions have been discussed in [19], [33], [34].

B. The Finite Band Solutions of Ma and Ablowitz

Finite band solutions as defined in Ma and Ablowitz [35] (also see [36]) are specified through the following parameters:

$$\begin{aligned}
\textit{Main spectrum:} \quad & \lambda_1, \ldots, \lambda_{2N} \in \mathbb{C}, & (19) \\
\textit{Initial auxiliary spectra:} \quad & \varrho_j(x_0, t_0) \in \mathbb{C}, & (20) \\
& \xi_j(x_0, t_0) \in \mathbb{C}, & (21) \\
\textit{Riemann sheet indices:} \quad & \nu_j(x_0, t_0) \in \{\pm 1\}, & (22) \\
& \vartheta_j(x_0, t_0) \in \{\pm 1\}, & (23) \\
& j = 1, \ldots, N.
\end{aligned}$$

The auxiliary spectra evolve according to the following system of coupled differential equations:

$$\partial_x \varrho_j = \frac{2\,\mathrm{i}\,\nu_j \sqrt{\prod_{k=1}^{2N}(\varrho_j - \lambda_k)}}{\prod_{\substack{i=1 \\ i \neq j}}^{N}(\varrho_j - \varrho_i)} \left(\sum_{\substack{i=1 \\ i \neq j}}^{N} \varrho_i - \frac{1}{2} \sum_{k=1}^{2N} \lambda_k \right), \quad (24)$$

$$\partial_x \xi_j = \frac{2\,\mathrm{i}\,\vartheta_j \sqrt{\prod_{k=1}^{2N}(\xi_j - \lambda_k)}}{\prod_{\substack{i=1 \\ i \neq j}}^{N}(\xi_j - \xi_i)} \left(\sum_{\substack{i=1 \\ i \neq j}}^{N} \xi_i - \frac{1}{2} \sum_{k=1}^{2N} \lambda_k \right). \quad (25)$$

There are also differential equations that govern the temporal evolution of the auxiliary spectra [35, Eqs. (4.2)+(8.4)], but they are quite complicated and will therefore not be given here. As before, the auxiliary spectra change their sheet indices if and only if one of them reaches a point in the main spectrum. The corresponding finite-band signal is given by

$$q = \frac{\kappa + \mathrm{i}}{2} \sum_{k=1}^{2N} \lambda_k - \kappa\sqrt{-\kappa} \sum_{j=1}^{N} \varrho_j - \sqrt{\kappa} \sum_{j=1}^{N} \xi_j. \quad (26)$$

Again, not every choice of initial parameters will lead to a solution of the NSE. In contrast to the finite-band solutions of Kotlyarov and Its, no condition for that seems to be known. The reconstruction formula (26) is now simpler, but the evolution of the auxiliary spectra (20), (21) is more complicated. The main spectrum coincides with that in Sec. II-A.

Example 5 (One band; [35], Sec. 2.4)**.** In the Ma-Ablowitz case, finding the auxiliary spectra ϱ_j and ξ_j is complicated even in the one-band case $N = 1$. The general form of $q(x,t)$ for $N = 1$ again turns out to be [35, p. 133ff]

$$q(x,t) = A e^{\kappa\,\mathrm{i}\,\frac{c_0}{c_1}x} e^{-\mathrm{i}((\frac{c_0}{c_1})^2 + 2\kappa A^2)t}.$$

The free parameters A, c_0, c_1 are related to the spectral parameters as follows:

$$\lambda_1 = \frac{c_0}{2c_1} - \sqrt{-\kappa}A, \quad \lambda_2 = \frac{c_0}{2c_1} + \sqrt{-\kappa}A,$$

$$\varrho_1(x,t) = \frac{c_0}{2c_1} - \sqrt{-\kappa}A \sin\left(\frac{c_0}{c_1}x + \left(-\kappa(\frac{c_0}{c_1})^2 + 2A^2 \right)t \right),$$

$$\xi_1(x,t) = \frac{c_0}{2c_1} - \sqrt{-\kappa}A \cos\left(\frac{c_0}{c_1}x + \left(-\kappa(\frac{c_0}{c_1})^2 + 2A^2 \right)t \right).$$

C. Construction of Finite-Band Solutions for Information Transmission in Optical Fiber

The efficient construction of signals with prescribed scattering data is fundamental in optical communication systems based on nonlinear Fourier transforms, where this problem corresponds to generating the input to the fiber on the transmitter side. The first papers addressing this problem have been published only very recently [10], [37], [38], [39], [14], [40], [41], [42], all for the NSE with vanishing boundary conditions. We remark that in all these works, not all available degrees of freedom are exploited in order to reduce the computational complexity of the problem. It has also been observed that some degrees of freedom seem to be ill-suited for information transmission due to sensitivity issues [9, VIII]. Another problem in these works arises due to the chosen boundary conditions: it is difficult to control the temporal spread of the generated signals. The only established method so far seems to be pruning of the signal set [10, Sec. V.C], [37, Sec. III], which is only feasible if the number of degrees of freedom is small.

It seems that the construction of periodic solutions to the NSE with prescribed scattering data for optical communication has not yet been discussed in the literature. We therefore now review quickly a few potential starting points. The construction of periodic finite-band signals appears at first more complicated than for vanishing boundary conditions because the parameters (3)–(6) are coupled through the condition in Theorem 2. However, they offer the important advantage that their temporal support is fixed. Theorem 2 can serve as a starting point for the construction of parameters (3)–(6) such that the function $q(x,t)$ generated by (7)–(10) actually solves the NSE (1) [26], [27], but it seems difficult to enforce the periodic boundary condition (2), especially if a specific period

is desired. However, several methods based on the *Darboux-Bäcklund transforms* (which have also been employed for vanishing boundary conditions [10]) are available for the construction of periodic finite-gap solutions to the NSE with pre-specified (complex) main spectrum: see [43], [44, Ch. 3.1] and [45, Sec. 4.2], and [46, Thm. 6.15]. The theta function representation discussed in Remark 4 offers another potential way to generate finite-band solutions. The advantage of these approaches it that the period of the constructed solution, and therefore its temporal spread, can be controlled.

III. THE NONLINEAR FOURIER TRANSFORM

The NFT of a *periodic* finite-band solution q, taken at a reference time t_0 and with respect to a reference point x_0, is the mapping from $q(\cdot, t_0)$ and x_0 to the scattering data, which is given by either (3)–(6), or (19)–(23). In this section, the computation of the scattering data is reviewed as a preparation for the derivation of the numerical algorithms. First, the main points of the Lax pair formalism are explained in order to motivate the spectral analysis of the differential operator

$$\mathbf{L}_{t_0} = \mathrm{i} \begin{bmatrix} \frac{d}{dx} & q(\cdot, t_0) \\ \kappa \bar{q}(\cdot, t_0) & -\frac{d}{dx} \end{bmatrix}. \tag{27}$$

Second, the spectral theory for the operator $\mathbf{L}_{t_0}$ will be reviewed. In particular, the monodromy matrix and the Floquet discriminant will be introduced. Finally, the relation between the spectrum of $\mathbf{L}_{t_0}$ and the scattering data will be established.

Remark 6. In the literature, $\mathbf{L}_{t_0}$ is sometimes replaced with

$$\begin{bmatrix} 1 & \\ & \zeta \end{bmatrix} \mathbf{L}_{t_0} \begin{bmatrix} 1 & \\ & \zeta \end{bmatrix}^{-1} = \begin{bmatrix} \mathrm{i} \frac{d}{dx} & \frac{\mathrm{i}}{\zeta} q(\cdot, t_0) \\ \mathrm{i} \zeta \kappa \bar{q}(\cdot, t_0) & -\mathrm{i} \frac{d}{dx} \end{bmatrix}.$$

All these operators are similar. Their eigenvalues coincide.

A. The Lax Pair Formalism

The relation between the operator $\mathbf{L}_{t_0}$ and the NSE (1) is as follows. One can find a second differential operator $\mathbf{B}$, which also depends on q, such that the condition

$$\partial_{t_0} \mathbf{L}_{t_0} = \mathbf{B} \mathbf{L}_{t_0} - \mathbf{L}_{t_0} \mathbf{B} \tag{28}$$

is equivalent to q being a solution to the NSE [5]. The details of how to find a suitable operator $\mathbf{B}$ will not be given here because this procedure is not important in this paper. See [47, Ch. 6.1] for an in-depth derivation. Any two operators $\mathbf{L}_{t_0}$ and $\mathbf{B}$ that satisfy the condition (28) are said to form a *Lax pair*. The main point about Lax pairs is that the eigenvalues of $\mathbf{L}_{t_0}$ are independent of t_0 [48]. Furthermore, the time-evolution of any eigenfunction $\phi_{t_0}(x)$ of $\mathbf{L}_{t_0}$ is simply $\partial_{t_0} \phi_{t_0} = \mathbf{B} \phi_{t_0}$. This relation is the reason why the time-evolution of the scattering data, which will be derived from the eigenstructure of $\mathbf{L}_{t_0}$, is usually simpler than that of the original signal.

B. Monodromy Matrix and Floquet Discriminant

The eigenproblem $\mathbf{L}_{t_0} \boldsymbol{v} = z \boldsymbol{v}$ can be rearranged to

$$\frac{d}{dx} \boldsymbol{v} = \begin{bmatrix} -\mathrm{i} z & -q(\cdot, t_0) \\ \kappa \bar{q}(\cdot, t_0) & \mathrm{i} z \end{bmatrix} \boldsymbol{v}. \tag{29}$$

Equation (29) has a unique non-trivial solution for any initial condition of the form $\boldsymbol{v}(x_0) = \boldsymbol{v}_0 \neq \mathbf{0}$ [49, Thm. 3.9]. However, although the coefficients in (29) are periodic, solutions to (29) will not be periodic in general. The choice $q \equiv 0$ and $z = \mathrm{i}$, for example, leads to $\boldsymbol{v}(x)^T = \begin{bmatrix} \mathrm{e}^x & \mathrm{e}^{-x} \end{bmatrix}$. In the following, only eigenvalues that admit bounded quasi-periodic eigenfunctions (i.e., $\boldsymbol{v}(x + \ell) \equiv m \boldsymbol{v}(x)$ with $|m| = 1$) will be of interest. The analysis of differential equations with periodic coefficients is the subject of Floquet theory [49, Ch. 3.6]. Following [28, Apdx.], we now outline how Floquet theory allows us to identify these eigenvalues in a way that is similar to how eigenvalues with finite-energy eigenfunctions are found for vanishing boundary conditions [8].

Let $\boldsymbol{\phi}_{x_0, t_0, z}$ and $\tilde{\boldsymbol{\phi}}_{x_0, t_0, z}$ denote the solutions of Eq. (29) with respect to the canonical initial conditions

$$\boldsymbol{\phi}_{x_0, t_0, z}(x_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \tilde{\boldsymbol{\phi}}_{x_0, t_0, z}(x_0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{30}$$

and the argument z. The *monodromy matrix*

$$\mathbf{M}_{x_0, t_0}(z) := \begin{bmatrix} \boldsymbol{\phi}_{x_0, t_0, z}(x_0 + \ell) & \tilde{\boldsymbol{\phi}}_{x_0, t_0, z}(x_0 + \ell) \end{bmatrix} \tag{31}$$

captures the evolution of these two solutions over one period. It can be thought of as the equivalent of the transfer matrix used with vanishing boundary conditions [8]. The monodromy matrix allows us to identify the z that admit quasi-periodic eigenfunctions with desired period transitions as follows.

Lemma 7 ([28], p. 831f.)**.** *Fix two arbitrary complex constants* $z, m \in \mathbb{C}$. *Then, the eigenproblem* $\mathbf{L}_{t_0} \boldsymbol{v} = z \boldsymbol{v}$ *admits a quasi-periodic eigenfunction* $\boldsymbol{v} \neq \mathbf{0}$ *in the sense that* $\boldsymbol{v}(x + \ell) \equiv m \boldsymbol{v}(x)$ *if and only if* $\Delta(z) := \frac{1}{2} \operatorname{tr} \mathbf{M}_{x_0, t_0}(z)$ *satisfies*

$$m^2 - 2m\Delta(z) + 1 = 0. \tag{32}$$

The function Δ is known as *Floquet discriminant* in the literature. As the notation suggests, Δ indeed does not depend on the reference points x_0 and t_0.

Lemma 8 ([35], Eqs. (3.8)+(8.3))**.** *The Floquent discriminant* Δ *is independent of the reference points* x_0 *and* t_0.

The determinant of the monodromy matrix is also invariant.

Lemma 9 ([35], Eqs. (1.5)+(6.3c))**.** *The monodromy matrix satisfies* $\det \mathbf{M}_{x_0, t_0}(z) \equiv 1$ *for all* x_0 *and* t_0.

The monodromy matrix possesses some symmetries.

Lemma 10 ([35], Secs. I.1+II.1)**.** *We have* $[\mathbf{M}_{x_0, t_0}(z)]_{2,2} \equiv [\bar{\mathbf{M}}_{x_0, t_0}(\bar{z})]_{1,1}$ *and* $[\mathbf{M}_{x_0, t_0}(z)]_{2,1} \equiv -\kappa [\bar{\mathbf{M}}_{x_0, t_0}(\bar{z})]_{1,2}$.

C. Connection between Kotlyarov-Its Finite-Band Solutions and the Monodromy Matrix

In this subsection, we assume that $q(x, t)$ is a periodic finite-band solution in the sense of Kotlyarov and Its. We develop expressions which allow us to compute the corresponding parameters (3)–(6) with the help of the monodromy matrix.

1) Squared Eigenfunctions in Terms of the Monodromy Matrix: The following theorem will enable us to evaluate the squared eigenfunctions given in Eqs. (12), (13) and (15) indirectly through evaluations of the monodromy matrix up to an unknown but commonly shared non-zero factor.

Theorem 11 ([24]; [27], Sec. 4.2, Apdx. I; [50], Sec. 4.6)**.** *Fix any choice of squared eigenfunctions (12), (13) and (15) such that the associated parameters (3)–(6) generate $q(x,t)$. Then, there exists a function $C : \mathbb{C} \to \mathbb{C}$ such that*

$$C(z)f_z(x_0,t_0) = -\,\mathrm{i}(\Delta(z) - [\mathbf{M}_{x_0,t_0}(z)]_{1,1}),$$
$$C(z)g_z(x_0,t_0) = [\mathbf{M}_{x_0,t_0}(z)]_{1,2}\,,$$
$$C(z)h_z(x_0,t_0) = [\mathbf{M}_{x_0,t_0}(z)]_{2,1}\,.$$

Proof: We only sketch the main idea. One defines

$$\tilde{f}_z(x,t) := -\frac{\mathrm{i}}{2}([\boldsymbol{\phi}_{x,t,z}]_1[\tilde{\boldsymbol{\phi}}_{x,t,z}]_2 + [\boldsymbol{\phi}_{x,t,z}]_2[\tilde{\boldsymbol{\phi}}_{x,t,z}]_1),$$
$$\tilde{g}_z(x,t) := [\boldsymbol{\phi}_{x,t,z}]_1[\tilde{\boldsymbol{\phi}}_{x,t,z}]_1, \quad \tilde{h}_z := -\kappa[\boldsymbol{\phi}_{x,t,z}]_2[\tilde{\boldsymbol{\phi}}_{x,t,z}]_2.$$

Then, one uses the fact that $\boldsymbol{\phi}_{x_0,t_0,z}$ and $\tilde{\boldsymbol{\phi}}_{x_0,t_0,z}$ are eigenfunctions of $\mathbf{L}_{t_0}$ to show that the spatio-temporal evolution of $\tilde{\boldsymbol{\theta}}_z^T := \begin{bmatrix} \tilde{f}_z & \tilde{g}_z & \tilde{h}_z \end{bmatrix}$ is also governed by same differential equations that were mentioned earlier for $\boldsymbol{\theta}_z$ in Footnote 6. Uniqueness arguments then show that $\boldsymbol{\theta}_z$ and $\tilde{\boldsymbol{\theta}}_z$ differ only by a normalization factor. Eventually, one connects $\tilde{f}_z$, $\tilde{g}_z$ and $\tilde{h}_z$ to the monodromy matrix using (31). ∎

2) Main Spectrum : Solving for $m = \pm 1$ in Eq. (32) shows that the operator $\mathbf{L}_{t_0}$ admits an (anti-)periodic eigenfunction for some z if and only if the Floquet discriminant satisfies $\Delta(z) \in \{\pm 1\}$. Under some weak assumptions, Theorem 2 implies that the main spectrum of a finite-band solution corresponds to the simple (anti-)periodic eigenvalues of $\mathbf{L}_{t_0}$.

Lemma 12. *Assume that there is a finite-band representation for $q(x,t)$ that satisfies (11), and fix it. Furthermore, assume that the roots of $1 - \Delta^2$ are at most double. The main spectrum (3) then corresponds exactly to the simple roots of $1 - \Delta^2$:*

$$\{\lambda_k\}_{k=1}^{2N} = \left\{ \zeta \in \mathbb{C} : \Delta(\zeta) \in \{\pm 1\}, \frac{d\Delta}{dz}(\zeta) \neq 0 \right\}. \quad (33)$$

Proof: The main spectrum corresponds to the simple roots of $P = f_z^2 - g_z h_z$ [cf. (11), (14) and (15)]. With the help of Lemma 9, Theorem 11 can be used to show that $C^2 P = (1 - \Delta^2)$. Now, every simple root of $1 - \Delta^2$ must be a root of P because the roots of C^2 are at least double. On the other hand, if P has a root then C^2 cannot have a root because the roots of $1 - \Delta^2$ are at most double. Thus, every simple root of P is also a simple root of $1 - \Delta^2$. ∎

A natural question arising at this point is whether the non-simple roots of $1 - \Delta^2$ are of any importance. The answer is yes; they are essential for analyzing the impact of perturbations.

Remark 13 (Non-simple roots). The non-simple roots[7] of $1 - \Delta^2$ can be interpreted as "canonical" degenerate points in the main spectrum of a finite-band solution. Since these roots are eigenvalues of the operator $\mathbf{L}_{t_0}$, perturbation theory

---

[7]i.e., $\zeta \in \mathbb{C}$ such that $\Delta(\zeta) \in \{\pm 1\}$ and $\frac{d\Delta}{dz}(\zeta) = 0$.

shows that double-roots will in general split up into two simple roots, leading to new non-degenerate points in the main spectrum. The corresponding hyperelliptic mode μ_j is no longer trapped. Although the impact of small perturbations on the roots themselves is small, they can nevertheless change the trajectories of the formerly trapped hyperelliptic mode significantly [27], [28], [29], [31]. As the solution evolves, this can lead to instabilities known as *rogue (or freak) waves* [34], [51]. Whether or not a double root can lead to an instability depends on its location in the complex plane. We note that real double roots cannot cause instabilities [29, Thm. 5.3].

Let us now illustrate matters with another example.

Example 14 (Plane wave; e.g. [28], Sec. II.A)**.** Consider the following periodic solution to the focusing NSE:

$$q(x,t) = q_0\, \mathrm{e}^{2\,\mathrm{i}\, q_0^2 t}, \quad q_0 \geq 0.$$

This is a special case of Ex. 3. We therefore know that a one-band representation for $q(x,t)$ that satisfies (11) exists, and can use Lem. 12 to analyze this particular representation. The Floquet discriminant is $\Delta(z) = \cos(\ell\sqrt{z^2 + q_0^2})$, which leads to infinitely many roots for $\Delta(z) \pm 1$:

$$\zeta_n^{\pm} := \pm\sqrt{n^2\pi^2/\ell^2 - q_0^2}, \quad n \in \mathbb{N}.$$

By taking the limit $\zeta \to \zeta_n^{\pm}$ with respect to

$$\frac{d\Delta}{dz}(\zeta) = -2\zeta\ell \frac{\sin\left(\ell\sqrt{\zeta^2 + q_0^2}\right)}{\sqrt{\zeta^2 + q_0^2}},$$

one finds that only the roots at $\zeta_0^{\pm} = \pm\,\mathrm{i}\, q_0$ are simple. Thus, the main spectrum is $\lambda_1 = \zeta_0^+$ and $\lambda_2 = \zeta_0^-$. The non-simple roots correspond to the $\zeta_n^{\pm}$ with $n \geq 1$. They are imaginary if $n < q_0\ell/\pi$ and real otherwise. The effect of perturbations on the non-simple roots is discussed e.g. in [28], [29].

In this example, it is interesting to note that the non-degenerate main spectrum is symmetric with respect to the real axis, and that the number of non-real double roots is finite. These two properties can be generalized as follows.

Lemma 15 ([29], [36])**.** *If the main spectrum is given by (33), it must consist of complex conjugate pairs. Furthermore, $\{\lambda_j\}_{j=1}^{2N} \subset \mathbb{R}$ in the defocusing case $\kappa = -1$.*

Proof: Lemma 10 implies $\bar{\Delta}(\bar{z}) \equiv \Delta(z)$, and thus (33) is symmetric with respect to the real axis. The λ_j are real if $\kappa = -1$ because $\mathbf{L}_{t_0}$ then is self-adjoint with respect to $\langle \phi, \tilde{\phi} \rangle = \int_{x_0}^{x_0+\ell} \tilde{\phi}(x)^* \phi(x) dx$. ∎

Lemma 16 ([28], [29])**.** *The functions $\Delta(z) \pm 1$ have only finitely many non-simple non-real roots. That is, $1 - \Delta(\zeta)^2 = \frac{d\Delta}{dz}(\zeta) = 0$ for only finitely many complex points $\zeta \notin \mathbb{R}$.*

Lemma 16 is important because, as mentioned in Remark 13, only non-real double roots can lead to instabilities [28], [29]. Thus, only finitely many degenerate points have to be taken into account during a stability analysis.

3) Auxiliary Spectrum : Next, consider the z for which the first element of the canonical eigenfunction $\tilde{\phi}_{x_0,t_0,z}$ vanishes at x_0 and $x_0 + \ell$:

$$\left[\tilde{\phi}_{x_0,t_0,z}(x_0)\right]_1 = \left[\tilde{\phi}_{x_0,t_0,z}(x_0 + \ell)\right]_1 = 0.$$

By definition of the monodromy matrix, this condition corresponds to the upper right element of the monodromy matrix being zero. The following lemma implies that these z constitute the auxiliary spectrum discussed in Sec. II-A up to degenerate parts which can be canceled as follows.

Lemma 17. *Let $q(x_0,t_0) \neq 0$, and assume (11) and that the roots of $1 - \Delta^2$ are at most double. Furthermore, set $n(\zeta) := 1$ if $\zeta \in \mathbb{C}$ is a double root of $1 - \Delta^2$ and $n(\zeta) := 0$ otherwise. Then, the auxiliary spectrum is*

$$\{\mu_j(x_0,t_0)\}_{j=1}^{N-1} = \left\{\zeta \in \mathbb{C} : \lim_{z \to \zeta} \frac{[\mathbf{M}_{x_0,t_0}(z)]_{1,2}}{(z - \zeta)^{n(\zeta)}} = 0\right\}.$$

Proof: In the proof of Lem. 12, it was shown that $C^2 = (1 - \Delta^2)/P$ and that the roots of P are exactly the simple roots of $1 - \Delta^2$. Thus, the roots of $C^2 = (1 - \Delta^2)/P$ are exactly the remaining double roots of $1 - \Delta^2$. The claim follows from (12) because $Cg_z = [\mathbf{M}_{x_0,t_0}(z)]_{1,2}$ by Thm. 11. ∎

D. Connection between Ma-Ablowitz Finite-Band Solutions and the Monodromy Matrix

In this subsection, the approach in Ma and Ablowitz [35] is reviewed and some expressions that will be convenient later are derived. Ma and Ablowitz assume that a solution $q(x,t)$ to the periodic NSE (1)–(2) is given such that $1 - \Delta^2$ has only finitely many simple roots. In the focusing case, it is furthermore assumed that:[8] [35, p. 129f]

1) all real roots of $1 - \Delta^2$ are double,
2) all non-real roots of $1 - \Delta^2$ are simple,
3) all roots of the terms defining the auxiliary spectra [i.e., (34) and (35) below] are simple, and
4) all real roots of the terms defining the auxiliary spectra coincide with a root of $1 - \Delta^2$.

Under these conditions, Ma and Ablowitz prove that the following finite-band solution coincides with $q(x,t)$ [35].

The (non-degenerate) main spectrum of the finite-band representation of $q(x,t)$ has been defined as the simple roots of $1 - \Delta^2$, i.e., again via (33) [35, p. 116f+129]. The corresponding auxiliary spectra (20)–(21) have been defined as the solutions (with respect to ϱ and ξ) of [35, p. 116f+129]

$$\mathfrak{I}[\phi_{x_0,t_0,\varrho,1}(x_0 + \ell)] - \mathrm{i}\sqrt{\kappa}\,\mathfrak{I}[\phi_{x_0,t_0,\varrho,2}(x_0 + \ell)] = 0, \quad (34)$$

$$\mathfrak{I}[\phi_{x_0,t_0,\xi,1}(x_0 + \ell)] - \mathrm{i}\sqrt{\kappa}\,\mathfrak{R}[\phi_{x_0,t_0,\xi,2}(x_0 + \ell)] = 0, \quad (35)$$

where $\mathfrak{R}[\phi(z)] := \frac{1}{2}(\phi(z) + \bar{\phi}(\bar{z}))$ and $\mathfrak{I}[\phi(z)] := \frac{1}{2\mathrm{i}}(\phi(z) - \bar{\phi}(\bar{z}))$. These conditions can be rewritten using the functions

$$\Psi^\pm(z) := \mathrm{i}\,[\mathbf{M}_{x_0,t_0}(z)]_{2,2} - \mathrm{i}\,[\mathbf{M}_{x_0,t_0}(z)]_{1,1} \quad (36)$$
$$- \sqrt{\pm 1}\sqrt{\kappa}([\mathbf{M}_{x_0,t_0}(z)]_{2,1} \pm \kappa\,[\mathbf{M}_{x_0,t_0}(z)]_{1,2}).$$

[8]These assumptions are quite strong. The function in Example 14, e.g., has non-real double roots and therefore violates them.

Lemma 18. *The auxiliary spectra (20)–(21) satisfy*

$$\{\varrho_j(x_0,t_0)\}_{j=1}^N = \{z \in \mathbb{C}\backslash\mathbb{R} : \Psi^+(z) = 0\}, \quad (37)$$

$$\{\xi_j(x_0,t_0)\}_{j=1}^N = \{z \in \mathbb{C}\backslash\mathbb{R} : \Psi^-(z) = 0\}. \quad (38)$$

We have $\Psi^\pm(z) \in \mathbb{R}$ whenever $\kappa = -1$ and $z \in \mathbb{R}$.

Proof: We only discuss Ψ^+. Lemma 10 implies $\frac{1}{2}\Psi^+ = \mathfrak{I}[\mathbf{M}_{x_0,t_0}]_{1,1} - \mathrm{i}\sqrt{\kappa}\,\mathfrak{I}[\mathbf{M}_{x_0,t_0}]_{2,1}$. Hence, $\Psi^+(z) \in \mathbb{R}$ if $\kappa = -1$ and $z \in \mathbb{R}$. Eq. (37) follows with (31). ∎

The next lemma will turn out to be essential at a later point.

Lemma 19 ([36], p. 113f). *The auxiliary spectra (20)–(21) are real in the defocusing case $\kappa = -1$.*

IV. Rational Approximations of the Monodromy Matrix

In the previous section, the scattering data has been expressed in terms of the monodromy matrix. The fast NFT algorithms that will be given later require a numerically tractable approximation of the monodromy matrix. Hence, in this section, rational approximations

$$\hat{\mathbf{M}}(z) = \frac{\mathbf{S}(w)}{d(w)} \approx \mathbf{M}_{x_0,t_0}(z), \qquad w = \varphi^{-1}(z), \quad (39)$$

of the monodromy matrix are derived given D equidistant samples of the signal q. That is, $\mathbf{S}(w)$ is a matrix-valued polynomial and $d(w)$ is a scalar-valued polynomial, respectively. The function φ denotes a coordinate transform. Unless specified otherwise, we shall use a Möbius transform

$$\varphi^{-1}(z) = \frac{dz - b}{a - cz}, \quad \varphi(w) = \frac{\varphi_1(w)}{\varphi_2(w)} = \frac{aw + b}{cw + d}. \quad (40)$$

Here, $a, b, c, d \in \mathbb{C}$ with $ad - bc \neq 0$. This transform has no influence on the results with respect to exact arithmetic operations, but it can be used to improve the numerical properties of the problem in finite precision. Specific choices will be discussed at the end of the section.

A. Ansatz

The monodromy matrix has been defined in terms of the two solutions to the differential equation in Eq. (31) that arise from the initial conditions in Eq. (30). Define the quantity

$$\boldsymbol{P}_z(x) := \begin{bmatrix} -\mathrm{i}\,z & -q(x,t_0) \\ \kappa\bar{q}(x,t_0) & \mathrm{i}\,z \end{bmatrix}. \quad (41)$$

The two solutions can be joined into the single equation

$$\frac{d}{dx}\mathbf{V}_z = \mathbf{P}_z\mathbf{V}_z, \quad \mathbf{V}_z(x_0) = \mathbf{I}. \quad (42)$$

The monodromy matrix can now be written as

$$\mathbf{M}(z) = \mathbf{V}_z(x_0 + \ell). \quad (43)$$

The general idea will be to replace the differential equation (42) with a difference equation which is then solved for an approximation of $\mathbf{V}_z(x_0 + \ell)$. The difference equation will be based on given samples of $q(\cdot, t_0)$ taken at the sample points

$$x_n := x_0 + n\varepsilon, \quad \varepsilon := \frac{\ell}{D}, \quad (44)$$

where $n \in \{0, \ldots, D-1\}$. Knowing that $\mathbf{P}_z(x_0 + \ell) = \mathbf{P}_z(x_0)$ because $q(\cdot, t_0)$ is periodic, also set $x_D := x_0$.

B. Forward Euler Method

This method is arguably the simplest way to solve Eq. (42) for $\mathbf{V}_z(x_D) = \mathbf{M}(z)$. Although it is rarely used in practice, it is a nice and simple means to illustrate the general rational approximation in Eq. (39). The discretized version of Eq. (42) in this scheme is

$$\frac{\hat{\mathbf{V}}_z[n+1] - \hat{\mathbf{V}}_z[n]}{\varepsilon} = \mathbf{P}_z(x_n)\hat{\mathbf{V}}_z[n], \quad \hat{\mathbf{V}}[0] = \mathbf{I}.$$

Solving for $\hat{\mathbf{V}}_z[n+1]$ results in

$$\hat{\mathbf{V}}_z[n+1] = (\mathbf{I} + \varepsilon\mathbf{P}_z(x_n))\,\hat{\mathbf{V}}_z[n]. \tag{45}$$

Eq. (43) suggests the approximation[9]

$$\begin{aligned}
\hat{\mathbf{M}}(z) :=&\, \hat{\mathbf{V}}_{\varphi(w)}[D] \\
=&\, \prod_{n=1}^{D} \left(\mathbf{I} + \varepsilon\mathbf{P}_{\varphi(w)}(x_n)\right) \\
=&\, \prod_{n=1}^{D} \begin{bmatrix} 1 - \mathrm{i}\,\varepsilon\varphi(w) & -\varepsilon q(x_n, t_0) \\ \varepsilon\kappa\bar{q}(x_n, t_0) & 1 + \mathrm{i}\,\varepsilon\varphi(w) \end{bmatrix} \\
=&\, \frac{\prod_{n=1}^{D} \begin{bmatrix} \varphi_2(w) - \mathrm{i}\,\varepsilon\varphi_1(w) & -\varepsilon\varphi_2(w)q(x_n, t_0) \\ \varepsilon\kappa\varphi_2(w)\bar{q}(x_n, t_0) & \varphi_2(w) + \mathrm{i}\,\varepsilon\varphi_1(w) \end{bmatrix}}{\varphi_2(w)} \\
=:&\, \frac{\mathbf{S}(w)}{d(w)}.
\end{aligned}$$

At this point, note that $\mathbf{S}(w)$ and $d(w)$ are indeed polynomials. The coordinate transform $\varphi(w)$ has been incorporated into them.

C. Crank-Nicolson

The Crank-Nicolson method is a quite popular finite difference scheme that is used in practice. The right side of Eq. (42) is now approximated with a central difference:

$$\frac{\hat{\mathbf{V}}_z[n+1] - \hat{\mathbf{V}}_z[n]}{\varepsilon} = \frac{\mathbf{P}_z(x_{n+1})\hat{\mathbf{V}}_z[n+1] + \mathbf{P}_z(x_n)\hat{\mathbf{V}}_z[n]}{2}.$$

Solving for $\hat{\mathbf{V}}_z[n+1]$ results in

$$\hat{\mathbf{V}}_z[n+1] = \left(\mathbf{I} - \frac{\varepsilon}{2}\mathbf{P}_z(x_{n+1})\right)^{-1}\left(\mathbf{I} + \frac{\varepsilon}{2}\mathbf{P}_z(x_n)\right)\hat{\mathbf{V}}_z[n]. \tag{46}$$

As before, the monodromy matrix will be approximated using the following ansatz suggested by Eq. (43):

$$\hat{\mathbf{M}}(z) := \hat{\mathbf{V}}_{\varphi(w)}[D] = \hat{\mathbf{V}}_z[D], \quad \hat{\mathbf{V}}_{\varphi(w)}[0] := \mathbf{I}.$$

The determinant of $\mathbf{I} - \frac{\varepsilon}{2}\mathbf{P}_{\varphi(w)}(x_{n+1})$ is

$$\frac{\varphi_2^2(w) + \frac{\varepsilon^2}{4}\left(\varphi_1^2(w) + \kappa\varphi_2^2(w)|q(x_{n+1}, t_0)|^2\right)}{\varphi_2^2(w)} =: \frac{d_n(w)}{\varphi_2^2(w)},$$

and the ansatz expands to

$$\begin{aligned}
\hat{\mathbf{M}}(z) =&\, \hat{\mathbf{V}}_{\varphi(w)}[D] \\
=&\, \prod_{n=1}^{D} \left(\mathbf{I} - \frac{\varepsilon}{2}\mathbf{P}_{\varphi(w)}(x_{n+1})\right)^{-1}\left(\mathbf{I} + \frac{\varepsilon}{2}\mathbf{P}_{\varphi(w)}(x_n)\right)
\end{aligned}$$

[9]Please note the order to the matrix product: $\prod_{k=1}^{K} \mathbf{A}_k = \mathbf{A}_K \times \cdots \times \mathbf{A}_1$.

$$\begin{aligned}
=&\, \prod_{n=1}^{D} \frac{\varphi_2^2(w)}{d_n(w)}\begin{bmatrix} 1 - \mathrm{i}\frac{\varepsilon}{2}\varphi(w) & -\frac{\varepsilon}{2}q(x_{n+1}, t_0) \\ \frac{\varepsilon}{2}\kappa\bar{q}(x_{n+1}, t_0) & 1 + \mathrm{i}\frac{\varepsilon}{2}\varphi(w) \end{bmatrix} \\
&\, \times \begin{bmatrix} 1 - \mathrm{i}\frac{\varepsilon}{2}\varphi(w) & -\frac{\varepsilon}{2}q(x_n, t_0) \\ \frac{\varepsilon}{2}\kappa\bar{q}(x_n, t_0) & 1 + \mathrm{i}\frac{\varepsilon}{2}\varphi(w) \end{bmatrix} = \frac{\mathbf{S}(w)}{d(w)},
\end{aligned}$$

where $d(w) := \prod_{n=1}^{D} d_n(w)$ and

$$\begin{aligned}
\mathbf{S}(w) :=&\, \prod_{n=1}^{D} \begin{bmatrix} \varphi_2(w) - \mathrm{i}\frac{\varepsilon}{2}\varphi_1(w) & -\frac{\varepsilon}{2}\varphi_2(w)q(x_{n+1}, t_0) \\ \frac{\varepsilon}{2}\kappa\varphi_2(w)\bar{q}(x_{n+1}, t_0) & \varphi_2 + \mathrm{i}\frac{\varepsilon}{2}\varphi_1(w) \end{bmatrix} \\
&\, \times \begin{bmatrix} \varphi_2(w) - \mathrm{i}\frac{\varepsilon}{2}\varphi_1(w) & -\frac{\varepsilon}{2}\varphi_2(w)q(x_n, t_0) \\ \frac{\varepsilon}{2}\kappa\varphi_2(w)\bar{q}(x_n, t_0) & \varphi_2(w) + \mathrm{i}\frac{\varepsilon}{2}\varphi_1(w) \end{bmatrix}.
\end{aligned}$$

D. Ablowitz-Ladik Scheme

The following scheme is known as the Ablowitz-Ladik scheme [52]:

$$\alpha_n := \sqrt{1 + \kappa\varepsilon^2|q(x_n, t_0)|^2},$$

$$\hat{\mathbf{V}}_z[n+1] = \alpha_n^{-1}\begin{bmatrix} \mathrm{e}^{-\mathrm{i}\varepsilon z} & -\varepsilon q(x_n, t_0) \\ \varepsilon\kappa\bar{q}(x_n, t_0) & \mathrm{e}^{\mathrm{i}\varepsilon z} \end{bmatrix}\hat{\mathbf{V}}_z[n].$$

Note that it is equivalent to the forward Euler method (45) up to an error of $O(\varepsilon^2)$ because $\mathrm{e}^{\mp\mathrm{i}\varepsilon z} = 1 \mp \mathrm{i}\varepsilon z + O(\varepsilon^2)$ and $\alpha_n = 1 + O(\varepsilon^2)$. The coordinate transform

$$w = \varphi^{-1}(z) := \mathrm{e}^{-\mathrm{i}\varepsilon z}, \quad z = \varphi(w) = \frac{\log w}{-\mathrm{i}\,\varepsilon}, \tag{47}$$

leads to the final form of the iteration:

$$\hat{\mathbf{V}}_w[n+1] = \alpha_n^{-1}\begin{bmatrix} w & -\varepsilon q(x_n, t_0) \\ \varepsilon\kappa\bar{q}(x_n, t_0) & w^{-1} \end{bmatrix}\hat{\mathbf{V}}_w[n]. \tag{48}$$

The monodromy matrix will be approximated using the ansatz

$$\hat{\mathbf{M}}(z) := \hat{\mathbf{V}}_w[D] \approx \hat{\mathbf{V}}_z^{(\text{Euler})}[D], \quad \hat{\mathbf{V}}_w[0] = \mathbf{I}, \tag{49}$$

in which $\hat{\mathbf{V}}_z^{(\text{Euler})}[D]$ is given by Eq. (45). This fits into the general framework of Eq. (39) if one chooses $d(w) := w^D$ and

$$\mathbf{S}(w) := \prod_{n=1}^{D} \alpha_n^{-1}\begin{bmatrix} w^2 & -\varepsilon w q(x_{n-1}, t_0) \\ \varepsilon\kappa w\bar{q}(x_{n-1}, t_0) & 1 \end{bmatrix}.$$

Remark 20. The normalization by α_n is not always given in the literature, but it has been reported to improve the numerical properties of the scheme in some cases [53].

Remark 21. The discretization (48) is amenable to a discrete version of the inverse scattering formalism [54].

E. Heuristic for Choosing The Coordinate Transform

Many polynomial operations such as root-finding or even simple evaluation are known to be problematic in finite precision arithmetic. Often problems arise if the coefficients of a polynomial cover a range that is too large for the commonly used IEEE 754 double precision floating point numbers. The problem can become even worse when a polynomial of a very high degree is evaluated at arguments x with absolute values $|x|$ that are not close enough to one. In this case, the powers $|x|^0, |x|^1, |x|^2, |x|^3, \ldots$ will cover a large range. Consider the following example, which illustrates the difficulties for the example of polynomial evaluation: $p_D(x) = \sum_{d=1}^{D} \frac{10^{-d}}{D}x^d$.

Degree D	128	256	512	1024	2048
Naive approach	1	1	NaN	NaN	NaN
Horner's method	1	1	0.60156	0.30078	0.15039
Reverse Horner	1	1	NaN	NaN	NaN

Table I
NUMERICAL EVALUATION OF $P_D(x)$ AT $x = 10$. THE EXACT RESULT IS $P_D(10) = 1$ FOR ANY D. NaN IS SHORT FOR "NOT A NUMBER" (E.G., $0 \times \infty =$ NaN IN IEEE 754). HORNER'S METHOD IS THE RECOMMENDED METHOD FOR THE NUMERICAL EVALUATION OF POLYNOMIALS IF $|x| \geq 1$. OTHERWISE, THE REVERSE HORNER'S METHOD SHOULD BE USED [55]. THAT IS, EVALUATE $P_D(u)/u$ AT $u = x^{-D}$.

It is $p_D(10) = 1$ for any D, but as is illustrated in Tab. I the numerical evaluation fails spectacularly for larger degrees.

The coordinate transform $z = \varphi(w)$ is a crucial factor in alleviating such problems when implementing fast NFTs, but finding good transforms remains a black art for now. Motivated by the issues just discussed, we wish to find transforms that will map the region of interest close to the unit circle. The coordinate transform (47) of the Ablowitz-Ladik scheme achieves this. For the other schemes, we use the Möbius transform (40) with $a = -M/\varepsilon$, $b = -a$, $c = i$, $d = i$ where $M = 1$ for the Euler scheme and $M = 2$ for Crank-Nicolson, respectively. These transforms map the real line to the unit circle. At the same time, they cancel several terms in the rational approximations.

V. FAST NUMERICAL NONLINEAR FOURIER TRANSFORM BASED ON FINITE-DIMENSIONAL EIGENPROBLEMS

In this section, a fast numerical NFT will be proposed that is based on approximating the main and auxiliary spectra through solutions of finite-dimensional eigenproblems. Its complexity is an order of magnitude lower than that of similar known algorithms based on matrix eigenproblems. This section consists of two parts. First, the algorithm is introduced. Then, it is compared with some other methods from the literature.

A. Description of the Algorithm

The input to this algorithm consists of the samples $q(x_0, t_0), \ldots, q(x_{D-1}, t_0)$, where the x_n are given in Eq. (44). The user has to decide for a rational approximation $\hat{\mathbf{M}}(z) = \mathbf{S}(w)/d(w)$, $z = \varphi(w)$, of the monodromy matrix that fits into Eq. (39). Several such schemes have been described in Sec. IV. The output of the algorithm will be the *numerical main spectrum* $\tilde{\lambda}_j$ and the *numerical auxiliary spectra* $\tilde{\mu}_j(x_0, t_0)$, $\tilde{\varrho}_j(x_0, t_0)$ and $\tilde{\xi}_j(x_0, t_0)$. The algorithm proceeds as follows.

1) Find the Monomial Basis Expansion: The polynomials $\mathbf{S}(w)$ and $d(w)$ have mostly been given in product forms $\mathbf{S}(w) = \prod_{n=1}^{D} \mathbf{S}_n(w)$ and $d(w) = \prod_{n=1}^{D} d_n(w)$, respectively, where the $\mathbf{S}_n(w)$ and $d_n(w)$ are polynomials of a low degree K. However, in the following, the coefficients of the polynomials $\mathbf{S}(w)$ and $d(w)$ with respect to the usual monomial basis $w^0, w^1, w^2, \ldots$ will be required. Algorithm 1 is a simple method that can find the coefficients of a polynomial in product form performing only $O(D \log^2 D)$ floating point operations (*flops*). It computes coefficients $\hat{\mathbf{S}}^{(k)} \in \mathbb{C}^{2 \times 2}$,

$\hat{d}^{(k)} \in \mathbb{C}$ and normalization constants $W_S, W_d \in \mathbb{Z}$ such that

$$\hat{\mathbf{M}}(w) = \frac{\mathbf{S}(w)}{d(w)} = \frac{2^{W_S} \sum_{k=0}^{\deg(\mathbf{S})} \hat{\mathbf{S}}^{(k)} w^k}{2^{W_d} \sum_{k=0}^{\deg(d)} \hat{d}^{(k)} w^k} =: 2^{W_S - W_d} \frac{\hat{\mathbf{S}}(w)}{\hat{d}(w)}. \quad (50)$$

The normalization constants arise from an effort to avoid overflows in Algorithm 1. They also ensure that the largest coefficients among the $[\hat{\mathbf{S}}^{(k)}]_{i,j}$ and $\hat{d}^{(k)}$ are of similar magnitude. The basis two was chosen for the normalization factor because multiplication and division by two can be carried out without loss of precision in IEEE 754 floating point numbers.

2) Find the Main Spectrum: Lemma 12 suggests to approximate the main spectrum by the roots of[10]

$$\hat{\Delta}(z) \pm 1 := \frac{1}{2} \left([\hat{\mathbf{M}}(w)]_{1,1} + [\hat{\mathbf{M}}(w)]_{2,2} \right) \pm 1$$
$$= 2^{W_S - 1} \frac{[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2} \pm 2^{W_d - W_S + 1} \hat{d}(w)}{\hat{d}(w)}. \quad (51)$$

The roots of these rational functions correspond to the roots of the two numerators $[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2} \pm 2^{W_d - W_S + 1} \hat{d}(w)$ that are not canceled by a root of the denominator $\hat{d}(w)$. It is a well-known fact that the roots of a polynomial correspond to the eigenvalues of an associated companion matrix that can be constructed from its monomial basis expansion. Companion matrices are highly structured, and recently several algorithms have been proposed that can find the eigenvalues of an $R \times R$ companion matrix with only $O(R^2)$ flops. See, e.g., [57], [58], [59], [60] and the references therein. We propose to find the roots of the two polynomials $[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2} \pm 2^{W_d - W_S + 1} \hat{d}(w)$ using this method. Therefore, one requires the monomial basis expansion of these two polynomials. Since the expansion (50) is already known, it can be computed in only $O(KD)$ flops. The roots of $\hat{d}(w)$ can often be found in closed form, or otherwise, if a product expansion $\hat{d}(w) = \prod_{n=1}^{D} \hat{d}_n(w)$ is known, numerically using at most $O(K^2 D)$ flops. Denote the roots of the two numerators that are not being canceled by a root of $\hat{d}(w)$ by $\hat{w}_j$. The worst-case complexity of finding them is $O(K^2 D^2)$ because for each root of the numerators one has to check whether that root is in the set of roots of the denominator. Often, this step can be simplified when $\hat{d}(w)$ has only a few distinct roots. We finally apply the coordinate transform in order to find the numerical main spectrum, $\hat{\lambda}_j := \varphi(\hat{w}_j)$. Adding up the complexities of the single steps, we see that the overall complexity of finding the numerical main spectrum is $O(K^2 D^2)$ flops.

3) Find the Kotlyarov-Its Auxiliary Spectrum: The replacement of the monodromy matrix in Lemma 17 with the rational approximation (50) results in[11]

$$[\hat{\mathbf{M}}(w)]_{1,2} = 2^{W_S - W_d} \frac{[\hat{\mathbf{S}}(w)]_{1,2}}{\hat{d}(w)} = 0. \quad (52)$$

---

[10]We do not check whether the roots are simple. Multiple roots will be detected by the root-finding algorithm and can be removed in Step 5, if desired.

[11]We do not check whether the found roots belong to the degenerated modes at this point. Such roots can be removed later in Step 5, if desired.

Algorithm 1 Fast product of N polynomials $\mathbf{p}_1, \ldots, \mathbf{p}_N$ (probably matrix-valued) with degree at most K.

Let us first illustrate the basic idea before the algorithm is given. The idea is to form the product in a tree-wise fashion, as is illustrated below for the case $N = 8$:
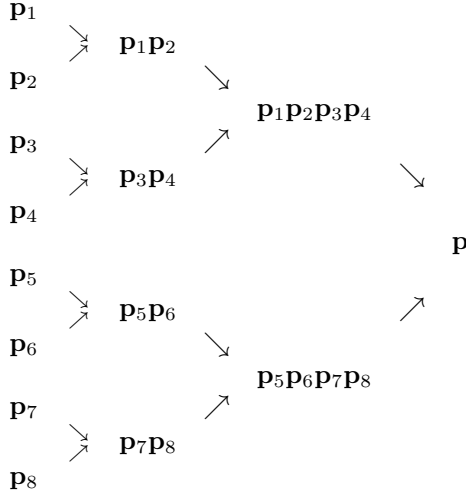
Assume for a moment that $N = 2^n$. Then, the algorithm will form $\frac{N}{2^i}$ products in the i-th level of the tree with degrees that are less than $2^i K$. If the FFT is used to multiply polynomials fast [56, p. 204ff], this leads to a complexity of $O(\frac{N}{2^i} 2^i K \log(2^i K))$ flops for level i [15]. The tree has $O(\log N)$ levels, thus the overall complexity is $O(NK \log^2(NK))$. A detailed implementation that also works for $N \neq 2^n$ is given below. Note that the intermediate products are normalized in order to avoid over-/underflows.

Input: N polynomials $\mathbf{p}_1, \ldots, \mathbf{p}_N$ of degree at most K
Output: W, $\mathbf{p} = 2^W \mathbf{p}_1 \mathbf{p}_2 \times \cdots \times \mathbf{p}_N$

- $W \leftarrow 0$
- while $N \geq 2$ do:
    - $N_m \leftarrow N \mod 2$
    - $N \leftarrow \frac{N - N_m}{2} + N_m$
    - for $n = 1, \ldots, N - N_m$ do:
        * $\mathbf{p}_n \leftarrow \mathbf{p}_{2n-1} \mathbf{p}_{2n}$
        * $a \leftarrow |\text{largest coefficient}(\mathbf{p}_n)|$
        * if $a > 0$ then $a \leftarrow \lfloor \log_2 a \rfloor$
        * $\mathbf{p}_n \leftarrow 2^{-a} \mathbf{p}_n$
        * $W \leftarrow W + a$
    - if $N_m \neq 0$ then: $\mathbf{p}_N \leftarrow \mathbf{p}_{2N-1}$
- $\mathbf{p} \leftarrow \mathbf{p}_1$

The roots in this equation are an approximation of the auxiliary spectrum. They correspond to the roots of the numerator $[\hat{\mathbf{S}}(w)]_{1,2}$ that are not canceled by the roots of $\hat{d}(w)$. Denote the remaining roots of $[\hat{\mathbf{S}}(w)]_{1,2}$ by $\hat{w}_j$. The numerical auxiliary spectrum (of Kotlyarov-Its type) is given by $\hat{\mu}_j(x_0, t_0) := \varphi(\hat{w}_j)$. If the same root finder as in Sec. V-A2 is used, it can be computed using $O(K^2 D^2)$ flops.

4) Find the Ma-Ablowitz Auxiliary Spectra: Now, the idea is to exploit Lemma 18. The replacement of the monodromy matrix in Eq. (36) with (50) results in

$$\hat{\Psi}^{\pm}(z) := 2^{W_S - W_d} \times \qquad (53)$$

$$\frac{\mathrm{i}[\hat{\mathbf{S}}(w)]_{2,2} - \mathrm{i}[\hat{\mathbf{S}}(z)]_{1,1} - \sqrt{\pm 1}\sqrt{\kappa}([\hat{\mathbf{S}}(w)]_{2,1} \pm \kappa[\hat{\mathbf{S}}(w)]_{1,2})}{\hat{d}(w)}.$$

The roots $\hat{w}_j^{\pm}$ of the numerators of $\hat{\Psi}^{\pm}$ that are not canceled by the denominators define the numerical auxiliary spectra (of Ma-Ablowitz type) $\hat{\varrho}_j(x_0, t_0) := \varphi(\hat{w}_j^+)$ and $\hat{\xi}_j(x_0, t_0) := \varphi(\hat{w}_j^-)$. The auxiliary spectra can be computed in $O(K^2 D^2)$ flops if the same root finder as in Sec. V-A2 is used.

5) Filter the Numerical Spectra: The numerical spectra will often contain artifacts that arise because of the discretization procedure. These artifacts will usually be well-separated from the real spectrum. Whenever there is some a-priori knowledge about the spectrum, it should be used to remove all other points in the numerical spectra that contradict this knowledge. The detected main spectrum may contain double roots indicating degenerate points in the main spectrum. These can be removed if desired. If a pair of degenerate points is removed from the main spectrum, then the same point should also be removed from the auxiliary spectra in order to remove the corresponding degenerate hyperelliptic modes as well.

Remark 22 (Root Cancellations). Depending on the discretization scheme, it is not always necessary to perform the root cancellations in the algorithm. The discretization schemes discussed in Sec. IV lead to denominators $\hat{d}(w)$ with roots that are usually well separated from the spectrum. The denominator in the Ablowitz-Ladik scheme has no finite roots at all. In the Euler and Crank-Nicolson schemes, when used with the Möbius transforms from Sec. IV-E, the roots of $\hat{d}(w)$ cluster around $w = -1$ as the step-size ε becomes small. In original coordinates, they will cluster around $z = \varphi(1) = \infty$. The root cancellation steps will therefore usually not be necessary with the discretization schemes from Sec. IV.

B. Comparison With Other Finite-Dimensional Eigenmethods

Another way to compute the scattering data is by direct discretization of the eigenrelation $\mathbf{L}_{t_0} \mathbf{v} = z \mathbf{v}$. Let us illustrate this approach with an example. The main spectrum consists of the eigenvalues of $\mathbf{L}_{t_0}$ that possess (anti-)periodic eigenvectors. The eigenrelation

$$\mathbf{L}_{t_0} \begin{bmatrix} u \\ v \end{bmatrix} = \mathrm{i} \begin{bmatrix} \frac{d}{dx} & q(\cdot, t_0) \\ \kappa \bar{q}(\cdot, t_0) & -\frac{d}{dx} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = z \begin{bmatrix} u \\ v \end{bmatrix} \tag{54}$$

can be discretized using Euler's method. With the sample points x_n and step size ε given in Eq. (44), this becomes

$$\mathrm{i}\frac{u(x_{n+1}) - u(x_n)}{\varepsilon} + \mathrm{i}\, q(x_n, t_0) v(x_n) = zu(x_n), \quad (55)$$

$$\mathrm{i}\, \kappa \bar{q}(x_n, t_0) u(x_n) - \mathrm{i}\frac{v(x_{n+1}) - v(x_n)}{\varepsilon} = zv(x_n), \quad (56)$$

where $n \in \{0, \ldots, D-1\}$. The signal q is periodic in x and the eigenfunctions are supposed to be (anti-)periodic. Hence, $q(x_D, t_0) = q(x_0, t_0)$ and $u(x_D) = \pm u(x_0)$, $v(x_D) = \pm v(x_0)$. The $2D$ linear equations in Eqs. (55)–(56) can be collected into a single (generalized) matrix eigenproblem that

can be solved with standard numerical methods such as the QZ algorithm. Each sign $\pm$ results in a separate eigenproblem, and the collected eigenvalues of both problems can be used as an approximation of the true main spectrum. The same approach can be used to approximate auxiliary spectra if the boundary conditions are modified appropriately. See, e.g., [61, Ch. 3.4]. Of course, methods other than Euler's can be used in order to discretize Eq. (54) including such where Eq. (54) is discretized in the Fourier domain. Several such methods have been proposed in the literature [9], [53], [62], [63], but so far no way of exploiting the structure of the resulting matrices seems to be known. Instead, standard eigenvalue solvers with $O(D^3)$ complexity had to be used. This is in contrast to our proposed algorithm, which requires only $O(D^2)$ flops.

Remark 23 (Floquet-Fourier-Hill Method). Deconinck and Kutz [64] proposed a novel approach to the numerical approximation of the spectrum of periodic differential operators like $\mathbf{L}_{t_0}$. Although their method is quite different from the ones discussed above, it still has a complexity of $O(D^3)$.

C. Comparison With Search Methods

Search methods are among the oldest methods for computing the nonlinear Fourier transform [65]. Let us illustrate the basic idea for the main spectrum. As before, the roots of the function $\hat{\Delta}(z) \pm 1$ will be used as the numerical approximations of the main spectrum. The difference is that this time, an iterative search method such as Newton's method will be used in order to locate the roots. Let $r[0]$ denote an estimate for a root of $\hat{\Delta}(z) \pm 1$. The derivative with respect to a complex variable $z = u + \mathrm{i} v$, $u, v \in \mathbb{R}$, is $\frac{d}{dz} := \frac{1}{2}(\partial_u - \mathrm{i}\,\partial_v)$. Therewith, the complex Newton's method can be written as [66]

$$
\begin{aligned}
r[i+1] =& r[i] + \left. \frac{\hat{\Delta} \pm 1}{\frac{d}{dz}(\hat{\Delta} \pm 1)} \right|_{z=r[n]} \\
=& r[i] + \left. \frac{[\hat{\mathbf{M}}]_{1,1} + [\hat{\mathbf{M}}]_{2,2} \pm 2}{[\frac{d}{dz}\hat{\mathbf{M}}]_{1,1} + [\frac{d}{dz}\hat{\mathbf{M}}]_{2,2}} \right|_{z=r[i]} .
\end{aligned}
\tag{57}
$$

The derivative of the approximated monodromy matrix $\hat{\mathbf{M}}(z)$ is required in order to perform this iteration. The exact formulas for this derivative depend on the discretization method. For Euler's method, e.g., it is $\hat{\mathbf{M}}(z) = \hat{\mathbf{V}}_z[D]$ where $\hat{\mathbf{V}}[D]$ is determined through the iteration (45). Consequently, $\frac{d}{dz}\hat{\mathbf{M}}(z) = \frac{d}{dz}\hat{\mathbf{V}}_z[D]$. Taking the derivative of (45) results in

$$
\begin{aligned}
\frac{d}{dz}\hat{\mathbf{V}}_z[n+1] =& (\mathbf{I} + \varepsilon \mathbf{P}_z(x_n)) \frac{d}{dz}\hat{\mathbf{V}}_z[n] \\
&+ \varepsilon \left(\frac{d}{dz}\mathbf{P}_z(x_n) \right) \hat{\mathbf{V}}_z[n].
\end{aligned}
\tag{58}
$$

Differentiation of the initial condition $\hat{\mathbf{V}}_z[0] = \mathbf{I}$ gives the missing initial condition for the derivative: $\frac{d}{dz}\hat{\mathbf{V}}_z[0] = \mathbf{0}$.

In summary, given an initial guess $r[0]$ of a root of $\hat{\Delta} \pm 1$, one iterates the Newton step (57) until some stopping criterion is fulfilled. For example, one may stop whenever the difference between two consecutive iterates falls below a certain threshold [9]. That is, $|r[i+1] - r[i]| \leq \beta$ for some $\beta > 0$.

The monodromy matrix and its derivative, which are required for each Newton step, can be found by using (45) and (58).

There are two remaining issues that have to be solved for any search method:

1) How does one choose the initial guesses for the roots?
2) How does one know that all roots of interest have been found? Newton's method does not return the multiplicities of the found roots.

These issues have been discussed in [9] for the NFT with non-periodic signals that vanish at infinity. In the communication scenario of [10], in which there are only finitely many possible locations for the roots, solving these problems is simple. It was proposed in [9] to use a few random perturbations of each potential root as initial guesses. Since Newton's method converges quickly for initial guesses that are close to a potential root, one can assume that all roots have been found after each potential root has been tested. The cost of performing one Newton step (57) using the iterations (45) and (58) is $O(D)$ flops. The number of iterations per root can be as high as $O(D)$ as well [67, p. 936]. Thus, in a communication scenario with P possible values for the spectral points, the worst-case complexity is at least $O(PD^2)$ flops. (The average complexity usually will be better, though.)

The situation becomes more involved in situations without precise a-priori knowledge. It was proposed in [9] to define a fixed region in which the spectrum is expected to be, and to choose initial guesses for the roots uniformly at random from this region. Regarding the stopping criterion, the authors of [9] proposed to check whether the spectrum found so far satisfies an energy conservation law. In the non-periodic case, the energy conservation law only involves the evaluation of an integral over the real line (see, e.g., [8, Apdx. B]). The corresponding law for the periodic case however involves other roots that also have to be found [50, Eq. (4.27)]. Hence, this stopping criterion seems to be less appropriate in the periodic case. Even very recent results on the initialization of Newton's method like [68] and [69] do not achieve $O(D^2)$ complexity.

In summary, search methods seem to be a good choice only if there is precise a-priori knowledge, and the number of roots P is low compared to the number of sample points D.

VI. FAST NUMERICAL NONLINEAR FOURIER TRANSFORM FOR THE DEFOCUSING NSE BASED ON SAMPLING

In this section, an especially fast numerical NFT that finds the main spectrum as well as the Ma-Ablowitz auxiliary spectra for the defocusing NSE (1) will be proposed. The idea is to exploit the fact that these spectra are always real in the defocusing case. The Kotlyarov-Its auxiliary spectrum does not have to be real and cannot be found with the method described in the following. The section is structured as follows. First, the algorithm is described. Then, its connection to a related approach from the literature is discussed.

A. Description of the Algorithm

The inputs and outputs are the same as in Sec. V-A with the exception that additionally a lower bound $A \in \mathbb{R}$ and an upper bound $B \in \mathbb{R}$, $A < B$, on the spectra have to be provided.

Algorithm 2 The root finding procedure for Sec.VI-A2. The evaluations of $\hat{\Delta}(z)$ have to implemented using the non-equidistant Fourier transform [70] in order to obtain an $O(LGD\log(GD))$ algorithm.

Input: Bounds $A, B \in \mathbb{R}$, oversampling factor $G \in \mathbb{N}$,
            number of bisection iterations $L \in \mathbb{N}$

Output: Numerical main spectrum $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_R$

- for $n = 0, \ldots, GD - 1$: $w_n = \varphi^{-1}(A + n\frac{B-A}{GD-1})$,

$$v_n \leftarrow \hat{\Delta}(\varphi(w_n)) = 2^{W_S - W_d - 1}\frac{[\hat{\mathbf{S}}(w_n)]_{1,1} + [\hat{\mathbf{S}}(w_n)]_{2,2}}{\hat{d}(w_n)}$$

- $R \leftarrow -1$
- for $n = 0, \ldots, GD - 2$:
    - if $\text{sign}(\text{Re}(v_{n+1}) - 1)) \neq \text{sign}(\text{Re}(v_n) - 1)$:
        * $R \leftarrow R + 1$, $a_R \leftarrow z_n$, $b_R \leftarrow z_{n+1}$
        * $s_R \leftarrow +1$, $u_R \leftarrow v_n$
    - if $\text{sign}(\text{Re}(v_{n+1}) + 1) \neq \text{sign}(\text{Re}(v_n) + 1)$:
        * $R \leftarrow R + 1$, $a_R \leftarrow z_n$, $b_R \leftarrow z_{n+1}$
        * $s_R \leftarrow -1$, $u_R \leftarrow v_n$
- for $l = 1, \ldots, L$:
    - for $i = 1, \ldots, R$: $u_r \leftarrow \hat{\Delta}(\varphi^{-1}(\frac{a_r + b_r}{2}))$
    - for $i = 1, \ldots, R$:
        * if $\text{sign}(\text{Re}(u_r) - s_r) = \text{sign}(\text{Re}(v_r) - s_r)$:
            $a_r \leftarrow \frac{a_r + b_r}{2}$, $u_r \leftarrow v_r$
        * else: $b_r \leftarrow \frac{a_r + b_r}{2}$
- for $i = 1, \ldots, R$: $\tilde{\lambda}_r \leftarrow \frac{a_r + b_r}{2}$

---

1) Find the Monomial Basis Expansion: One starts as in Sec. V-A. First, the monomial basis expansions in Eq. (50) are computed. All other expansions that will be required in this algorithm can be found from this information with $O(D)$ flops.

2) Find the Main Spectrum: As in Sec. V-A, the numerical main spectrum will be found from the roots of (51). However, this time Lemma 10 ensures that $\Delta(z) \pm 1$ is real-valued on the real axis. Additionally, all roots are real by Lemma 15. This makes the root finding problem much easier. Algorithm 2 shows a straight-forward three-step method to isolate the roots. First, $\hat{\Delta}(z)$ is sampled on an equidistant grid. Second, the adjacent sample points where any of the two functions $\hat{\Delta}(z) \pm 1$ changes its sign are located. The midpoint of any such pair of adjacent sample points forms an estimate of a root. In the third step, these estimates are refined using bisection.

The costs of Algorithm 2 are dominated by the evaluations of $\hat{\Delta}(z)$. With naive evaluations of (46), (48), etc., the overall costs of the algorithm are $O(LG^2D^2)$. However, note that the coordinate transforms $w = \varphi^{-1}(z)$ in Sec. IV-E as well as the transform (47) map the real axis to the complex unit circle. Therefore, the *non-equidistant fast Fourier transform (NFFT)* [70] can be used in order to evaluate $\hat{\Delta}(z)$ efficiently. For example, the $v_0, \ldots, v_{GD-1}$ can be found as follows. First, one uses the NFFT to evaluate $[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2}$ at the GD points $w_n := \varphi^{-1}(z_n)$ using only $O(GD\log(GD))$ flops because $|w_n| = 1$ for all n. Second, one uses the NFFT to

evaluate $\hat{d}(w)$ at the w_n. Finally, the values of $\hat{\Delta}(z)$ are given by

$$v_n = 2^{W_S - W_d - 1}\frac{[\hat{\mathbf{S}}(w_n)]_{1,1} + [\hat{\mathbf{S}}(w_n)]_{2,2}}{\hat{d}(w_n)}.$$

Both NFFTs require $O(GD\log(GD))$ flops. The last step requires another $O(GD)$ flops. In total, one can thus find $v_0, \ldots, v_{GD-1}$ using only $O(GD\log(GD))$ flops. Similarly, the $u_1, \ldots u_R$ can be found using only $O(GD\log(GD))$ flops because $R \leq GD$. The costs of the remaining operations are negligible, so that the overall costs of Algorithm 2 amount to $O(LGD\log(GD))$ flops if the NFFT is utilized.

3) Find the Ma-Ablowitz Auxiliary Spectra: The roots of the function $\Psi^{\pm}(z)$ defined in (36) constitute the auxiliary spectra. These roots are real by Lemma 19. Furthermore, Lemma 18 shows that $\Psi^{\pm}(z)$ is real for real z. The auxiliary spectra can therefore be found with the same method as the main spectrum, using only $O(LGD\log(GD))$ flops. One only has to replace $\hat{\Delta} \pm 1$ with $\hat{\Psi}^{\pm}$ as defined in Eq. (53).

4) Filter the Numerical Spectra: As in Sec. V-A. The spectra of the defocusing NSE have a special band structure [35, p. 117], which can help to identify numerical artifacts.

Remark 24 (Root Refinements). In Algorithm 2, bisection has been used to refine the roots. Of course, more advanced methods like the secant method, Muller's method [71], or Ridders' method [72] could be used as well. Since derivatives are easily found in the Fourier domain, the NFFT can also be used the compute the derivative of $\hat{\Delta}(z)$ efficiently. Thus, even the use of Newton's method should be possible.

Remark 25 (Root Cancellations). In contrast to the fast eigenmethod from Sec. V-A, the algorithm in this section did not contain any root cancellation procedures. Root-finding methods based on sampling do not indicate the multiplicities of the found roots. Hence, one cannot detect, e.g., whether a root of $[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2}$ will been canceled by a root of $\hat{d}(w)$ unless it is assumed that the roots of $[\hat{\mathbf{S}}(w)]_{1,1} + [\hat{\mathbf{S}}(w)]_{2,2}$ are simple. Fortunately, Remark 22 applies in this case as well.

B. Comparison With A Similar Approach

Algorithm 2 is quite similar to a numerical NFT proposed by Christov in the context of the Korteweg-de Vries equation [18], whose associated Lax operator has a real spectrum as well. There are some differences in terms of the root refinement, but the main difference to our approach, which actually is the key to obtaining a fast algorithm, however is that rational approximations of the monodromy matrix are used in this paper. In contrast, an irrational approximation of the monodromy matrix has been used in [18]. Since an evaluation of the monodromy matrix in [18] takes $O(GD)$ flops, the overall runtime of the algorithm in [18] is about $O(LD^2G^2)$ flops. This is about an order of magnitude higher than for our implementation, which has runtime of $O(LGD\log(GD) + KD\log^2(KD))$. Here, the second term is due to Algorithm 1.

VII. NUMERICAL EXAMPLES

In this section, two numerical examples are investigated in order to demonstrate that the new proposed algorithms

can improve significantly over the existing ones in terms of runtimes without sacrificing numerical reliability. The first example evaluates the finite eigenmethod from Sec. V-A, while the second example investigates the performance of the specialized method from Sec. VI-A for the defocusing NSE. The results presented here should be understood as proof-of-concept. A comprehensive numerical or analytical study of the new algorithms remains a topic for future research.

Remark 26. The numerical examples have been carried out using MATLAB R2012a, but for the most time-consuming parts of the algorithms external implementations have been used. In order to realize the root-finding algorithm in Sec. V-A2, a Fortran implementation of the algorithm in [59], which is available at http://www.unilim.fr/pages_perso/paola.boito/software.html, has been interfaced. For the non-equidistant fast Fourier transform required in Sec. VI-A2, version 3.2.3 of the NFFT3 library [70], which is available at http://www.tu-chemnitz.de/~potts/nfft, has been used. Specifically, the included example given in the file `test_nfft1d.m` has been adapted to our needs. Algorithm 1 and the direct evaluation of (48) have been implemented in C. The FFT in Algorithm 1 has been realized using version 1.3.0 of the KISS FFT routine, which is available at http://sourceforge.net/projects/kissfft/. Root cancellations have not been implemented.

A. The Focusing Case

The first example is the initial condition $q(x, t_0) = q_0 \, e^{i \mu x}$, which has been discussed e.g. in [53]. The main spectrum with respect to the focusing NSE is $\zeta_n^\pm = -\frac{\mu}{2} \pm i \sqrt{|q_0|^2 - \frac{n^2}{4}}$, $n \in \mathbb{N}$, where all eigenvalues except $\zeta_0^\pm$ are double. Similar to Example 14, there is thus one non-degenerate band.

First, the runtimes of the new fast numerical NFT from Sec. V-A will be compared with finite difference methods as described in Sec. V-B. Two discretizations will be considered: the Ablowitz-Ladik ("AL") discretization and the Crank-Nicolson ("CN") discretization. The signal parameters are $q_0 = \mu = 3$ and $\ell = 2\pi$. The first plot in Fig. 1 shows the minimum runtime *per sample point*, taken over three runs, versus the number of sample points D. The per-sample runtimes of the fast algorithms grow approximately linearly with D as expected, while the per-sample runtimes of the standard algorithms grow quadratically. This corresponds to quadratic and cubic overall runtimes, respectively.

Next, the numerical accuracy of these algorithm is compared. Since the signal has infinitely many degenerate modes, only the errors with respect to a finite subset of the spectrum will be considered. Denote the rectangle spanned by two complex numbers $X, Y \in \mathbb{C}$ by $\Omega(X, Y)$. The error between the true spectrum $\{\lambda_n^\pm\}_n$ and the numerical spectrum $\{\tilde{\lambda}_j\}_j$ with respect to $\Omega(X, Y)$ can be measured by

$$e := \max \left\{ \max_{\lambda_n^\pm \text{ s.t. } \lambda_n^\pm \in \Omega(X,Y)} \min_{\tilde{\lambda}_j \text{ s.t. } \tilde{\lambda}_j \in \Omega(X,Y)} |\lambda_n^\pm - \tilde{\lambda}_j|, \right.$$
$$\left. \max_{\tilde{\lambda}_j \text{ s.t. } \tilde{\lambda}_j \in \Omega(X,Y)} \min_{\lambda_n^\pm \text{ s.t. } \lambda_n^\pm \in \Omega(X,Y)} |\lambda_n^\pm - \tilde{\lambda}_j| \right\}.$$

Note that the first term in the outer maximum grows large if an algorithm fails to approximate a part of the true spectrum within $\Omega(X, Y)$, while the second term becomes large if an algorithm creates spurious terms within $\Omega(X, Y)$ that have no correspondence in the true spectrum. The second plot in Fig. 1 depicts the error for $\Omega(-5 + i, 5 + 5i)$. That is, only the non-real spectrum is considered. All four errors decrease approximately linearly (i.e., doubling D halves the error), but the errors of the fast algorithms interestingly are lower than those of the standard algorithms although the same discretizations are used. (This is only an apparent discrepancy as the standard algorithms approximate the Lax operator (27), while the fast algorithms approximate the monodromy matrix (31).) Fig. 2 illustrates the different accuracies by comparing the exact and the numerical spectra for $D = 32$. The errors for $\Omega(-5+i, 5+5i)$ are all much higher (not shown) because all algorithms have problems with the approximation of the eigenvalue at zero. However, the errors still decrease linearly in that case and the relative performance of the algorithms with respect to each other remains the same.

B. The Defocusing Case

1) First Example: One-band Solution: The next example is the $q(x, t_0) = \frac{3}{2} \, e^{i \, 3x}$, where the period is $\ell = 2\pi$. This initial condition corresponds to a one-band solution as derived in Example 3 with $\lambda_1 = -3$ and $\lambda_2 = 0$. The first plot in Fig. 3 depicts the per-sample runtime of the fast NFT from Sec. VI-A with that of a naive implementation where the monodromy matrix is evaluated directly through (48). In each case, $L = 5$ bisection steps have been carried out. The oversampling factor was $G = 1$. The per-sample runtime of the fast algorithm grows only very slowly with the number of samples, while it grows linearly for the standard implementation. The second plot in Fig. 4 shows the same error as in Sec. VII-A where now $\{\tilde{\lambda}_j\} = \{-3, 0\}$, $X = -10$ and $Y = 10$. We can see that the fast algorithm gives exactly the same errors as the naive implementation. The first plot in Fig. 4 finally shows the Floquet diagram (i.e., a plot of $\hat{\Delta}(z)$ where the scale is linear if $|\hat{\Delta}(z)| \le 1$ and logarithmic otherwise) computed by the fast algorithm.

2) Second Example: Gaussian Wavepacket: The last example is $q(x, t_0) = q_0 \, e^{i \mu x} \, e^{-\frac{x^2}{\sigma}}$, which has been discussed in [17]. The parameters considered here are $q_0 = 1.9$, $\mu = 1$, $\sigma = 2$, $\ell = 10$. We do not show the runtime plot because it is very similar to the one in Fig. 3. The second plot in Fig. 4 shows the Floquet diagram computed by the fast algorithm. While the exact error cannot be quantified in this case because the analyical NFT seems to be unknown, a comparison of the Floquet diagram to that in [17, Fig. 2b] confirms the result found by the fast algorithm.

VIII. RAPIDLY DECAYING SIGNALS

The fast transforms derived in this paper can also be carried over to rapidly decaying non-periodic signals on the line. Details on the necessary modifications can be found in [15], where a preliminary form of the eigenmethod in Sec. V has been presented for rapidly decaying signals. Alternatively, the new algorithms for the periodic case may be used directly if the decaying signal is truncated and extended periodically for
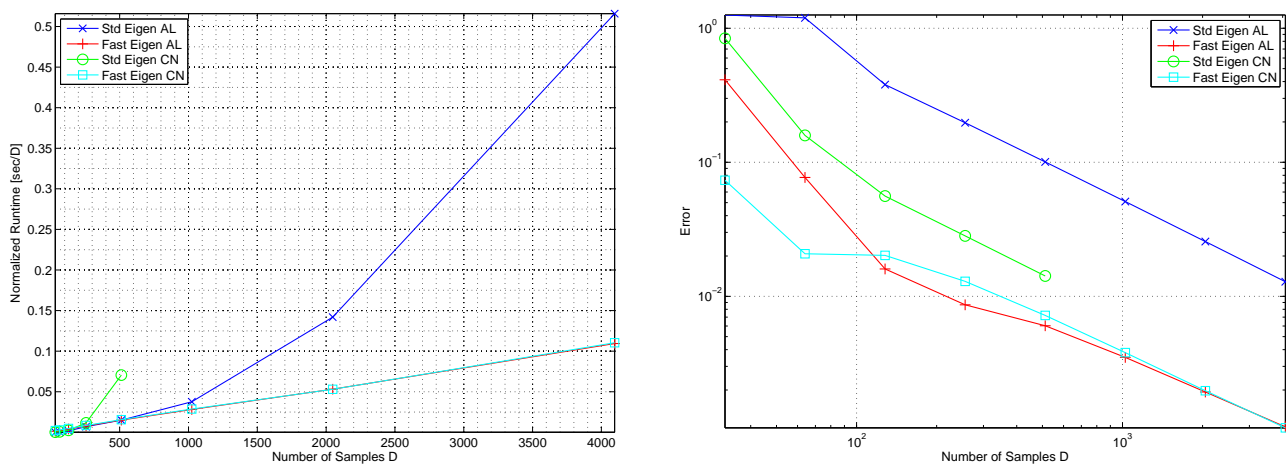
Figure 1. Focusing one-band solution. Left: Runtimes per sample, Right: Error. Please note that the Std-CN method was too slow for higher D.
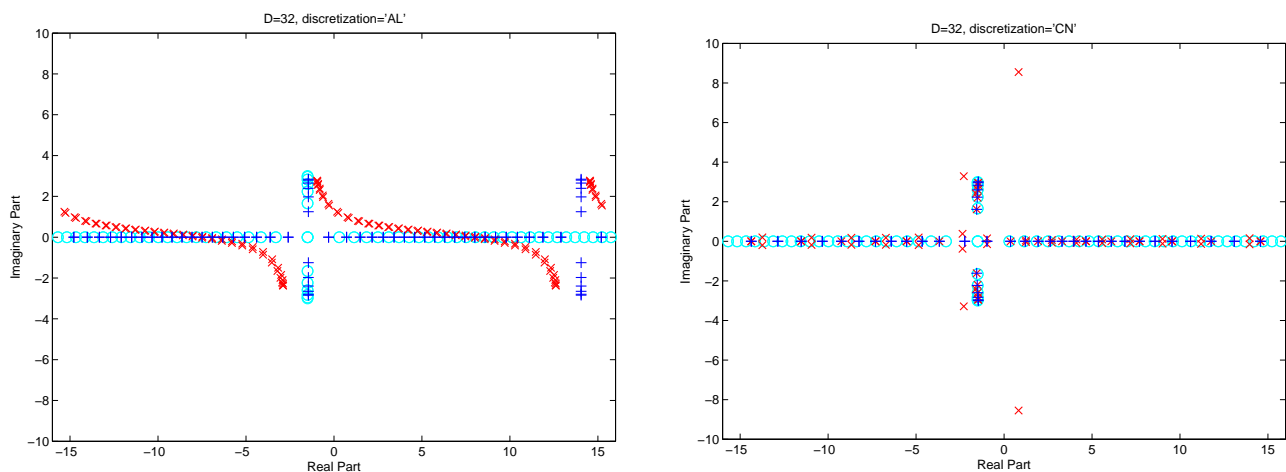

Figure 2. Exact main spectrum (o) vs. numerical main spectra found by the the new fast eigenmethod (+) and the conventional eigenmethod (x)

a large enough period $\ell > 0$ [73]. This however requires some additional transformations of the found spectra.

IX. OTHER INTEGRABLE EVOLUTION EQUATIONS

The new fast NFTs presented in this paper have been developed for signals that are governed by the NSE (1). However, the approach extends to signals governed by other evolution equations as well. In this section, the extension of our results will be discussed for the *Ablowitz-Kaup-Newell-Segur (AKNS)* Lax pair [7]. (The Lax pair formalism has been introduced in Sec. III-A.) The authors did not investigate extensions to other Lax pairs so far, although they feel that such extensions should be possible along lines similar to those outlined below. Finally, please note that the proposed extensions have not been investigated in numerical experiments so far. Their numerical accuracy remains to be examined.

A. The AKNS Lax Pair

In Sec. III-A it was mentioned that the NLS (1) arises from the compatibility condition (28) for certain Lax pairs. Ablowitz et al. [7] have established that many other important evolution

equations can be expressed through the condition (28) for Lax pairs with

$$\mathbf{L}_{t_0} = \mathrm{i} \begin{bmatrix} \frac{d}{dx} & q(\cdot, t_0) \\ r(\cdot, t_0) & -\frac{d}{dx} \end{bmatrix} \qquad (59)$$

for suitably chosen signals q, r and operator $\mathbf{B}$. The standard examples of evolution equations that fall into this framework (other than the NSE) are the following [7, p. 258]:

- The *Korteweg-de Vries equation*: $r = 1$,

$$\partial_t q + 6q\partial_x q + \partial_{xxx} q = 0.$$

- The *modified KdV equation*: $r = \pm q$,

$$\partial_t q \pm 6q^2 \partial_x q + \partial_{xxx} q = 0.$$

- The *sine-Gordon equation*: $r = q = \frac{1}{2}\partial_x u$,

$$\partial_{xt} u = \sin u.$$

- The *sinh-Gordon equation*: $r = -q = \frac{1}{2}\partial_x u$,

$$\partial_{xt} u = \sinh u.$$

Either under additional transformations, or by considering matrix-valued signals $\mathbf{q}$ and $\mathbf{r}$, many more equations can be fit into the AKNS framework [74], [75], [76], [77], [78], [79].
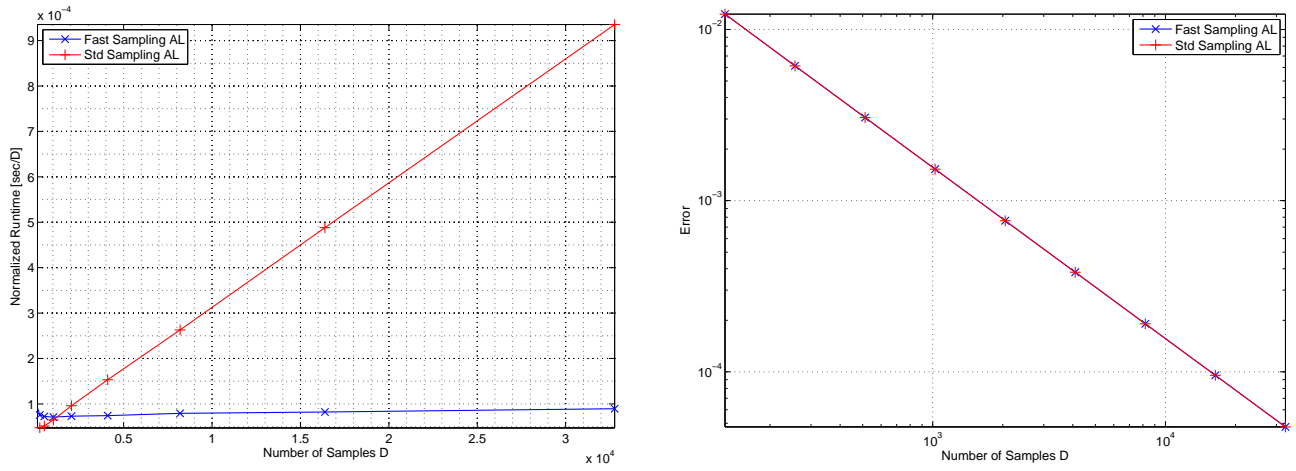
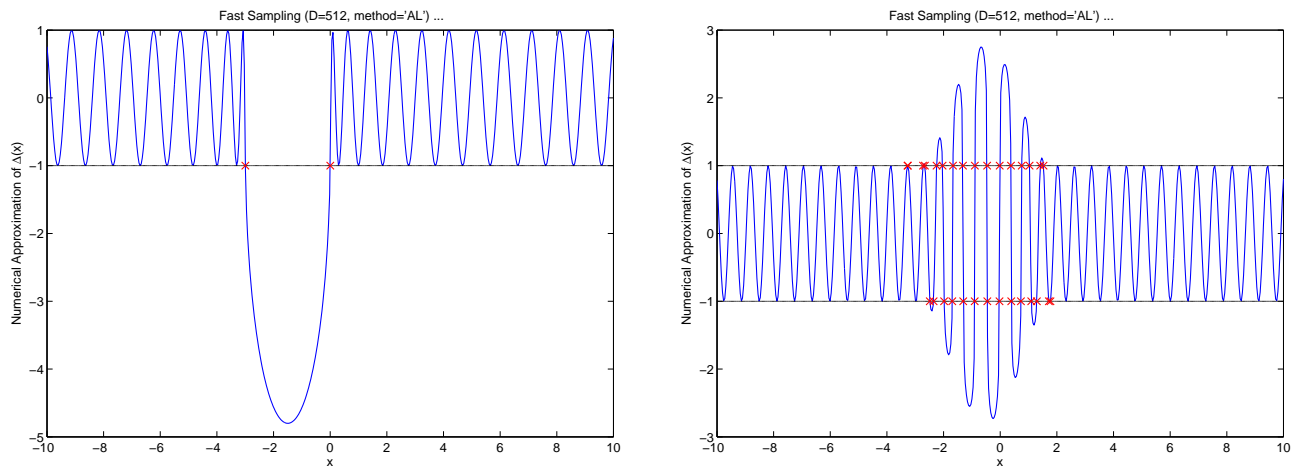Figure 3. Defocusing one-band solution. Left: Runtimes per sample, Right: Error.

Figure 4. Floquet diagram (–) and the main spectral points (x) for the one-band solution (left) and the Gaussian wavepacket (right)

B. Finite-Band Solutions for the AKNS Lax Pair

Tracy has presented finite-band solutions for the AKNS Lax pair in [27, Ch. 2.2]. (An even more general case has been discussed in [80].) One starts with exactly the same form as in Sec. II. Then, one adds a second set of $N-1$ auxiliary variables $\eta_j(x,t)$ and Riemann sheet indices $\theta_j(x,t) \in \{\pm 1\}$, respectively. The η_j are governed by the differential equations

$$\partial_x \eta_j = \frac{2\,\mathrm{i}\,\theta_j \sqrt{\prod_{k=1}^{2N}(\eta_j - \lambda_k)}}{\prod_{\substack{m=1\\m\neq j}}^{N-1}(\eta_j - \eta_m)},$$

$$\partial_t \eta_j = -2\left(\sum_{\substack{m=1\\m\neq j}}^{N-1}\eta_m - \frac{1}{2}\sum_{k=1}^{2N}\lambda_k\right)\partial_x \eta_j.$$

The signal r evolves according to

$$\partial_x \ln r = -2\,\mathrm{i}\left(\sum_{j=1}^{N-1}\eta_j - \frac{1}{2}\sum_{k=1}^{2N}\lambda_k\right).$$

The squared eigenfunction (13) has to be replaced with

$$h_z(x,t) = \mathrm{i}\,r(x,t)\prod_{j=1}^{N-1}(z - \eta_j(x,t)).$$

Then, one has that q and r solve the system

$$\mathrm{i}\,\partial_t q + \partial_{xx} q + 2q^2 r = 0,$$
$$-\mathrm{i}\,\partial_t r + \partial_{xx} r + 2r^2 q = 0$$

if and only if (15) is a polynomial [27, Thm. 2.1].

C. Fast NFTs for the AKNS Lax Pair

The algorithms presented in this paper can easily be extended to general AKNS Lax pairs. All results except Lemma 10, Lemma 15, and Sec. III-D on the computation of the scattering data in Sec. III carry over to the general AKNS case if the operator $\mathbf{L}_{t_0}$ in (27) is replaced with (59) and Eq. (29) is updated accordingly. The scattering data has to be extended by the initial conditions $\eta_j(x_0, t_0)$, which turn out to be the roots of $[\mathbf{M}_{x_0,t_0}(z)]_{2,1}$.

Numerically, only minor changes are necessary in the development of Sec. IV in order to adapt the rational approximations of the monodromy matrix if the operator $\mathbf{L}_{t_0}$ is changed.

Basically, only the matrix $\mathbf{P}_z$ defined in (41) has to be changed such that the eigenproblem $\mathbf{L}_{t_0}\mathbf{v} = z\mathbf{v}$ is again equivalent to $\frac{d}{dx}\mathbf{v} = \mathbf{P}_z\mathbf{v}$. Using (59), one finds that $\mathbf{P}_z$ is given by

$$\frac{d}{dx}\mathbf{v} = \left[\begin{array}{cc} -\mathrm{i}\,z & -q(\cdot,t_0) \\ r(\cdot,t_0) & \mathrm{i}\,z \end{array} \right] \mathbf{v} =: \mathbf{P}_z\mathbf{v}. \qquad (60)$$

Of course, consecutive terms that involve $\mathbf{P}_z$ have to be reevaluated. Then, our proposed fast algorithms in Secs. V-A and VI-A can be run as before. The methods used to find the $\mu_j(x_0,t_0)$ can be used to find the $\eta_j(x_0,t_0)$ as well.

X. Conclusion

In this paper, two fast numerical nonlinear Fourier transforms for the periodic nonlinear Schrödinger equation have been proposed. The first algorithm has a complexity of $O(D^2)$ flops, where D denotes the number of sample points. The second algorithm applies only to the defocusing nonlinear Schrödinger equation, but its complexity is only $O(D\log^2 D)$ flops. In both cases, this is about an order of magnitude better than what other comparable algorithms achieve so far. The feasibility of the fast transforms has been demonstrated in several numerical examples. Extensions to other cases have been discussed as well.

References

[1] A. Hasegawa and M. Matsumoto, *Optical Solitons in Fibers*, 3rd ed. Springer, 2003.

[2] M. Nakazawa and K. Suzuki, "10 Gbit/s pseudorandom dark soliton data transmission over 1200 km," *Electron. Lett.*, vol. 31, no. 13, pp. 1076–1077, Jun. 1995.

[3] Y. S. Kivshar and B. Luther-Davies, "Dark optical solitons: physics and applications," *Phys. Rep.*, vol. 298, no. 2–3, pp. 81–197, May 1998.

[4] W. Craig, C. Sulem, and P. L. Sulem, "Nonlinear modulation of gravity waves: A rigorous approach," *Nonlinearity*, vol. 5, no. 2, pp. 497–522, 1992.

[5] V. F. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of wave in nonlinear media," *Sov. Phys. JETP*, vol. 34, no. 1, pp. 62–69, Jan. 1972.

[6] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Method for solving the Korteweg-deVries equation," *Phys. Rev. Lett.*, vol. 19, no. 19, pp. 1095–1097, 1967.

[7] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, "The inverse scattering transform – Fourier analysis for nonlinear problems," *Stud. Appl. Math.*, vol. 53, pp. 249–315, 1974.

[8] 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, Jul. 2014.

[9] ——, "Information transmission using the nonlinear Fourier transform, Part II: Numerical methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, Jul. 2014.

[10] ——, "Information transmission using the nonlinear Fourier transform, Part III: Spectrum modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.

[11] R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[12] P. J. Winzer, "Optical networking beyond WDM," *IEEE Photonics J.*, vol. 4, no. 2, pp. 647–651, 2012.

[13] A. Hasegawa and T. Nyu, "Eigenvalue communication," *J. Lightwave Technol.*, vol. 11, no. 3, pp. 395–399, Mar. 1993.

[14] J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, "Nonlinear spectral management: Linearization of the lossless fiber channel," *Optics Express*, vol. 21, no. 20, pp. 24 344–24 367, 2013.

[15] S. Wahls and H. V. Poor, "Introducing the fast nonlinear Fourier transform," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Vancouver, Canada, May 2013.

[16] M. I. Yousefi, "Information transmission using the nonlinear Fourier transform," Ph.D. thesis, Univ. Toronto, Toronto, Canada, Mar. 2013, http://hdl.handle.net/1807/35179.

[17] A. R. Osborne, "Numerical inverse scattering transform for the periodic, defocusing nonlinear Schrödinger equation," *Phys. Lett. A*, vol. 176, pp. 75–84, 1993.

[18] I. Christov, "Internal solitary waves in the ocean: Analysis using the periodic, inverse scattering transform," *Math. Comput. Simul.*, vol. 80, no. 1, pp. 192–201, Sep. 2009.

[19] A. R. Osborne, *Nonlinear Ocean Waves and the Inverse Scattering Transform*, 1st ed. Academic Press, 2010.

[20] M. Brühl and H. Oumeraci, "Nonlinear decomposition of transmitted wave trains behind submerged reef structures using "nonlinear Fourier transform": The nonlinear spectral basic components," in *Proc. ASME Int. Conf. Ocean Offshore Arctic Eng. (OMAE)*, Rio de Janeiro, Brazil, Jul. 2012.

[21] T. Dudok de Wit, "Data analysis techniques for resolving nonlinear processes in plasmas: a review," Preprint, Nov. 1996, arXiv:comp-gas/9611002v1.

[22] V. B. Matveev, "30 years of finite-gap integration theory," *Phil. Trans. R. Soc. A*, vol. 366, no. 1867, pp. 837–875, 2008.

[23] P. G. Grinevich, "Approximation theorem for the self-focusing nonlinear Schrödinger equation and for the periodic curves in R^3," *Physica D*, vol. 152–153, pp. 20–27, May 2001.

[24] V. P. Kotlyarov, "Periodic problem for the nonlinear Schrödinger equation," *Problems of Mathematical Physics and Functional Analysis (Naukova Dumka, Kiev)*, vol. 1, pp. 121–131, 1976, English translation: arXiv:1401.4445 [nlin.SI].

[25] A. R. Its and V. P. Kotlyarov, "Explicit formulas for solutions of the nonlinear Schrödinger equation," *Dokl. Akad. Nauk UkrSSR Ser. A*, vol. 11, pp. 965–968, 1976, English translation: arXiv:1401.4445 [nlin.SI] (p. 10ff).

[26] E. R. Tracy, H. H. Chen, and Y. C. Lee, "Study of quasiperiodic solutions of the nonlinear Schrödinger equation and the nonlinear modulation instability," *Phys. Rev. Lett.*, vol. 53, no. 3, pp. 218–221, Jul. 1984.

[27] E. R. Tracy, "Topics in nonlinear wave theory with applications," Ph.D. thesis, Univ. Maryland, College Park, 1984.

[28] E. R. Tracy and H. H. Chen, "Nonlinear self-modulation: An exactly solvable model," *Phys. Rev. A*, vol. 37, no. 3, pp. 815–839, Feb. 1988.

[29] M. G. Forest and J.-E. Lee, "Geometry and modulation theory for the periodic nonlinear Schrodinger equation," in *Oscillation Theory, Computation, and Methods of Compensated Compactness*, ser. IMA Volumes Math. Appl. Springer, 1986, vol. 2, pp. 35–69.

[30] A. R. Its, A. V. Rybin, and M. A. Sall, "Exact integration of nonlinear Schrödinger equation," *Theor. Math. Phys.*, vol. 74, no. 1, pp. 20–32, 1988.

[31] J.-E. J. Lee, "The inverse spectral solution, modulation theory and linearized stability analysis of n-phase, quasi-periodic solutions of the nonlinear Schrodinger equation," PhD thesis, Ohio State Univ., 1986.

[32] P. D. Miller, N. M. Ercolani, I. M. Krichever, and C. D. Levermore, "Finite genus solutions to the Ablowitz-Ladik equations," *Commun. Pure Appl. Math.*, vol. 48, no. 12, pp. 1369–1440, 1995.

[33] B. Deconinck, M. Heil, A. Bobenko, M. Van Hoeij, and M. Schmies, "Computing Riemann theta functions," *Math. Comput.*, vol. 73, no. 247, pp. 1417–1442, 2003.

[34] A. R. Osborne, "Rogue waves: Classification, measurement and data analysis, and hyperfast numerical modeling," *Eur. Phys. J. Special Topics*, vol. 185, pp. 225–245, 2010.

[35] Y.-C. Ma and M. J. Ablowitz, "The periodic cubic Schrödinger equation," *Stud. Appl. Math.*, vol. 65, pp. 113–158, 1981.

[36] Y.-C. Ma, "Studies of the cubic Schroedinger equation," PhD thesis, Princeton Univ., 1977.

[37] S. Hari, F. Kschischang, and M. Yousefi, "Multi-eigenvalue communication via the nonlinear Fourier transform," in *Proc. Bien. Symp. Commun.*, Kingston, ON, Canada, Jun. 2014, pp. 92–95.

[38] S. Wahls and H. V. Poor, "Inverse nonlinear Fourier transforms via interpolation: The Ablowitz-Ladik case," in *Proc. Int. Symp. Math. Theory Networks Systems (MTNS)*, Groningen. The Netherlands, Jul. 2014, pp. 1848–1855.

[39] E. G. Turitsyna and S. K. Turitsyn, "Digital signal processing based on inverse scattering transform," *Optics Lett.*, vol. 38, no. 20, pp. 4186–4188, 2013.

[40] J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," *Phys. Rev. Lett.*, vol. 113, Jul. 2014, article no. 013901.

[41] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," *Optics Express*, vol. 22, no. 22, pp. 26 720–26 741, Nov. 2014.

[42] Q. Zhang, T. H. Chan, and A. Grant, "Spatially periodic signals for fiber channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, Jun. 2014, pp. 2804–2808.

[43] H. Steudel and R. Meinel, "Periodic solutions generated by Bäcklund transformations," *Physica D*, vol. 21, no. 1, pp. 155–162, Aug. 1986.

[44] C. M. Schober, "Numerical and analytical studies of the discrete nonlinear Schroedinger equation," PhD thesis, Univ. Arizona, 1991, http://hdl.handle.net/10150/185595.

[45] Y. Li and D. W. McLaughlin, "Morse and Melnikov functions for NLS pde's," *Commun. Math. Phys.*, vol. 162, pp. 175–214, 1994.

[46] R. C. Cascaval, F. Gesztesy, H. Holden, and Y. Latushkin, "Spectral analysis of Darboux transformations for the focusing NLS hierarchy," *J. Anal. Math.*, vol. 93, no. 1, pp. 139–197, Dec. 2004.

[47] W. Eckhaus and A. Van Harten, *The Inverse Scattering Transformation and the Theory of Solitons: An Introduction*, ser. Math. Stud.   North-Holland, 1981, vol. 50.

[48] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," *Comm. Pure Appl. Math.*, vol. 21, no. 5, pp. 467–490, 1968.

[49] G. Teschl, *Ordinary Differential Equations and Dynamical Systems*, ser. Graduate Stud. Math.   Providence, RI: AMS, 2010, vol. 144.

[50] E. R. Tracy, J. W. Larson, A. R. Osborne, and L. Bergamasco, "The relationship between the spectral theories for the periodic Korteweg-deVries and nonlinear Schrödinger equations," in *Nonlinear Topics in Ocean Physics*, A. R. Osborne, Ed.   Amsterdam: Elsevier, 1991.

[51] D. R. Solli, C. Ropers, P. Koonath, and B. Jalali, "Optical rogue waves," *Nature*, vol. 450, pp. 1054–1057, Dec. 2007.

[52] M. J. Ablowitz and J. F. Ladik, "Nonlinear differential-difference equations and Fourier analysis," *J. Math. Phys.*, vol. 17, no. 6, pp. 1011–1018, Jun. 1976.

[53] J. A. C. Weideman and B. M. Herbst, "Finite difference methods for an AKNS eigenproblem," *Math. Comput. Simul.*, vol. 43, no. 1, pp. 77–88, Jan. 1997.

[54] M. J. Ablowitz, B. Prinari, and A. D. Trubatch, *Discrete and Continuous Nonlinear Schrödinger Systems*, ser. London Math. Soc. Lect. Notes Ser. Cambridge Univ. Press, 2004, vol. 302.

[55] C. S. Burrus. (2012, Mar.) Horner's method for evaluating and deflating polynomials. [Online]. Available: http://cnx.org/content/m15099/1.7/

[56] E. W. Kamen and B. S. Heck, *Fundamentals of Signals and Systems Using the Web and MATLAB*, 3rd ed.   Pearson, 2007.

[57] S. Chandrasekaran, M. Gu, J. Xia, and J. Zhu, "A fast QR algorithm for companion matrices," *Oper. Theory Adv. Appl.*, vol. 179, pp. 111–143, 2007.

[58] M. Van Barel, R. Vandebril, P. Van Dooren, and K. Frederix, "Implicit double shift QR-algorithm for companion matrices," *Numer. Math.*, vol. 116, no. 2, pp. 177–212, Aug. 2010.

[59] P. Boito, Y. Eidelman, L. Gemignani, and I. Gohberg, "Implicit QR with compression," *Indag. Math.*, vol. 23, no. 4, pp. 733–761, Dec. 2012.

[60] R. Bevilacqua, G. M. Del Corso, and L. Gemignani, "A CMV–based eigensolver for companion matrices," Preprint, Jun. 2014, arXiv:1406.2820v1.

[61] W. D. MacEvoy, "Techniques and instabilities for 1+1 and 2+1 dimensional integrable partial differential equations." PhD thesis, Univ. Arizona, 1994, http://hdl.handle.net/10150/187015.

[62] S. Burtsev, R. Camassa, and I. Timofeyev, "Numerical algorithms for the direct spectral transform with applications to nonlinear Schrödinger type systems," *J. Comput. Phys.*, vol. 147, no. 1, pp. 166–186, Nov. 1998.

[63] P. Chamorro-Posada, G. S. McDonald, G. H. C. New, and F. J. Fraile-Pelaez, "Fast algorithm for studying the evolution of optical solitons under perturbations," *IEEE Trans. Magn.*, vol. 25, no. 3, pp. 1558–1561, May 1999.

[64] B. Deconinck and J. N. Kutz, "Computing spectra of linear operators using the Floquet-Fourier-Hill method," *J. Comput. Phys.*, vol. 219, no. 1, pp. 296–321, Nov. 2006.

[65] G. Boffetta and A. R. Osborne, "Computation of the direct scattering transform for the nonlinear Schroedinger equation," *J. Comput. Phys.*, vol. 102, no. 2, pp. 252–264, Oct. 1992.

[66] L. Yau and A. Ben-Israel, "The Newton and Halley methods for complex roots," *Am. Math. Mon.*, vol. 105, no. 9, pp. 806–818, Nov. 1998.

[67] D. Schleicher, "On the number of iterations of Newton's method for complex polynomials," *Ergod. Th. & Dynam. Sys.*, vol. 22, no. 3, pp. 935–945, 2002.

[68] J. Hubbard, D. Schleicher, and S. Sutherland, "How to find all roots of complex polynomials by Newton's method," *Invent. Math.*, vol. 146, no. 1, pp. 1–33, 2001.

[69] T. Bilarev, M. Aspenberg, and D. Schleicher, "On the speed of convergence of Newton's method for complex polynomials," Preprint, Feb. 2012, arXiv:1202.2475v1.

[70] J. Keiner, S. Kunis, and D. Potts, "Using NFFT 3 – a software library for various nonequispaced fast Fourier transforms," *ACM Trans. Math. Software*, vol. 36, no. 4, Aug. 2009, article 19.

[71] D. E. Muller, "A method for solving algebraic equations using an automatic computer," *Math. Tables Other Aids Comput.*, vol. 10, no. 56, pp. 208–215, 1956.

[72] C. Ridders, "A new algorithm for computing a single root of a real continuous function," *IEEE Trans. Circuit Syst.*, vol. 26, no. 11, pp. 979–980, 1979.

[73] C. P. Olivier, B. M. Herbst, and M. A. Molchan, "A numerical study of the large-period limit of a Zakharov-Shabat eigenvalue problem with periodic potentials," *J. Phys. A: Math. Theor.*, vol. 45, no. 25, 2012.

[74] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Sov. Phys. JETP*, vol. 38, no. 2, pp. 248–253, Feb. 1974.

[75] D. J. Kaup, "A higher-order water-wave equation and the method for solving it," *Prog. Theor. Phys.*, vol. 54, no. 2, pp. 396–408, Aug. 1975.

[76] M. Jaulent and I. Miodek, "Connection between Zakharov-Shabat and Schrödinger-type inverse-scattering transforms," *Lettere Al Nuovo Cimento Ser. 2*, vol. 20, no. 18, pp. 655–660, 1977.

[77] M. R. Gupta, "Exact inverse scattering solution of a non-linear evolution equation in a non-uniform medium," *Phys. Lett. A*, vol. 72A, no. 6, pp. 420–422, Aug. 1979.

[78] Y. Ishimori, "A relationship between the Ablowitz-Kaup-Newell-Segur and Wadati-Konno-Ichikawa schemes of the inverse scattering method," *J. Phys. Soc. Jpn.*, vol. 51, no. 9, pp. 3036–3041, Sep. 1982.

[79] A. Sakovich and S. Sakovich, "The short pulse equation is integrable," *J. Phys. Soc. Jpn.*, vol. 74, no. 1, pp. 239–241, Jan. 2005.

[80] A. M. Kamchatnov, "On the Baker-Akhiezer function in the AKNS scheme," *J. Phys. A: Math. Gen.*, vol. 34, no. 33, pp. L441–L446, 2001.