

# A Robust Inverse Scattering Transform for the Focusing Nonlinear Schrödinger Equation

DENIZ BILMAN  
*University of Michigan*

PETER D. MILLER  
*University of Michigan*

## Abstract

We propose a modification of the standard inverse scattering transform for the focusing nonlinear Schrödinger equation (also other equations by natural generalization) formulated with nonzero boundary conditions at infinity. The purpose is to deal with arbitrary-order poles and potentially severe spectral singularities in a simple and unified way. As an application, we use the modified transform to place the Peregrine solution and related higher-order “rogue wave” solutions in an inverse-scattering context for the first time. This allows one to directly study properties of these solutions such as their dynamical or structural stability, or their asymptotic behavior in the limit of high order. The modified transform method also allows rogue waves to be generated on top of other structures by elementary Darboux transformations rather than the generalized Darboux transformations in the literature or other related limit processes. © 2019 Wiley Periodicals, Inc.

## 1 Introduction

Take the Cauchy initial-value problem for the focusing cubic nonlinear Schrödinger (NLS) equation in one space dimension written in the form

$$(1.1) \quad i \frac{\partial \psi}{\partial t} + \frac{1}{2} \frac{\partial^2 \psi}{\partial x^2} + (|\psi|^2 - 1)\psi = 0, \quad x \in \mathbb{R}, \quad t > 0.$$

Here we have written the equation in a “rotating frame” by replacing the traditional nonlinear term  $|\psi|^2\psi$  with  $(|\psi|^2 - 1)\psi$ , which ensures that  $\psi(x, t) \equiv 1$  is an exact solution, to which we demand solutions decay for large  $|x|$  by imposing the boundary conditions

$$(1.2) \quad \lim_{x \rightarrow \pm\infty} \psi(x, t) = 1, \quad t > 0.$$

The formulation of the Cauchy problem is complete (modulo technical details of the sense in which (1.1) is satisfied and in which the side conditions are achieved,

usually encoded in the choice of an appropriate space) by specifying an initial condition

$$(1.3) \quad \psi(x, 0) = \psi_0(x), \quad x \in \mathbb{R}.$$

The equation (1.1) frequently arises in an asymptotic limit as an amplitude equation for weakly nonlinear, monochromatic, strongly dispersive waves in conservative systems. In this setting, the exact solution  $\psi(x, t) \equiv 1$  corresponds in the physical system to a uniform traveling wavetrain, nearly sinusoidal due to weak nonlinearity. Such nearly sinusoidal waves were studied by G. G. Stokes in the setting of surface water waves and the corresponding solution  $\psi(x, t) \equiv 1$  of (1.1) is sometimes called a *Stokes wave*. The Cauchy problem (1.1)–(1.3) is therefore a mathematical formulation of the question of what becomes of a spatially localized perturbation of the Stokes wave. We will also follow convention and frequently refer to the solution  $\psi = \psi_{\text{bg}}(x, t) \equiv 1$  as the *background*.

One can check that the solution of the Cauchy problem (1.1)–(1.3) with initial data

$$(1.4) \quad \psi_0(x) := 1 - 4 \frac{1 - 2it_0}{1 + 4(x - x_0)^2 + 4t_0^2}$$

(here  $(x_0, t_0) \in \mathbb{R}^2$  is a pair of parameters) is the *Peregrine breather* [35]

$$(1.5) \quad \psi(x, t) = \psi_{\text{P}}(x, t) := 1 - 4 \frac{1 + 2i(t - t_0)}{1 + 4(x - x_0)^2 + 4(t - t_0)^2}.$$

This solution may be viewed as a model for a “rogue wave”: it converges uniformly to the background solution  $\psi_{\text{bg}}(x, t) \equiv 1$  as  $t \rightarrow \pm\infty$ , but has a peak at  $x = x_0$  that occurs at the time  $t = t_0$  in the sense that  $\max_{(x,t) \in \mathbb{R}^2} |\psi(x, t)| = |\psi(x_0, t_0)| = 3$ . See Figure 1.1. In studies of surface water waves, “rogue,” “extreme,” or “freak” waves are commonly defined as waves that are spatially and temporally localized (hence coming out of nowhere and disappearing again without a trace) disturbances of amplitudes exceeding that of the background by an order of magnitude. Rogue waves are viewed as rare events (and the background amplitude is usually defined in the sense of a statistical average) but they can cause damage to ships and fixed structures such as oil drilling platforms. For these reasons they have attracted substantial scientific attention in recent years; see [24, 28] for a review of these notions. Rogue waves have been produced in wave-tank experiments [11]. A recent application of the focusing NLS equation (1.1) to model ocean surface rogue waves is the statistical, data-driven computational tool described in [13] (see also [34]), which aims to give a warning of a few minutes for the appearance of a rogue wave.

Of course, as the focusing NLS equation (1.1) is a universal asymptotic model for weakly nonlinear waves, analogues of the Peregrine solution and of rogue waves more generally are expected to occur in a myriad of physical systems. One landmark in this direction was the experimental observation [25, 29] of extreme light waves well-modeled by the Peregrine solution in nonlinear optical

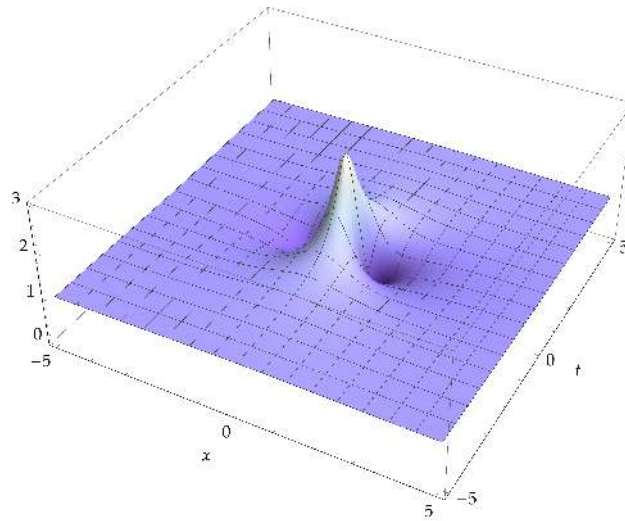


FIGURE 1.1. The modulus  $|\psi_P(x, t)|$  of the Peregrine breather solution  $\psi_P(x, t)$  (1.5) of the focusing NLS equation (1.1) with its peak located at  $(x_0, t_0) = (0, 0)$ .

fibers. Some more recent experiments involving the generation and observation of Peregrine-like rogue waves in nonlinear optics are described in [10, 36, 39]. The optical experiments in particular are very interesting because they suggest some sort of robustness (if not proper dynamical stability) of Peregrine rogue waves in the sense that the experiments are reliable and repeatable (and in some cases the reported results are actually averages over many propagating pulses).

Another context in which the Peregrine breather solution (1.5) reliably appears is in the examination of solutions of the focusing NLS equation (1.1) with *general* initial data but in a certain semiclassical limit. The semiclassical limit makes the characteristic length scale and duration of the Peregrine solution small compared to the length scale in the initial data and the length of time over which the solution is observed. Thus it becomes relatively unimportant exactly which kind of boundary conditions are imposed at infinity. In [6] the semiclassical scaling for (1.1) was studied under the assumption that the boundary condition (1.2) was replaced with

$$(1.6) \quad \lim_{x \rightarrow \pm\infty} \psi(x, t) = 0, \quad t > 0,$$

and it was shown that for a large class of (implicitly specified) initial data the solution of the Cauchy problem develops a universal wave pattern in which numerous Peregrine breathers appear at predictable (via the pole distribution of the *tritonquée* Painlevé-I transcendent) locations  $(x_0, t_0)$  within a wedge-shaped domain in the  $(x, t)$ -plane opening from its vertex in the positive  $t$ -direction. There are many Peregrine breathers centered at points  $(x_0, t_0)$  that, in the semiclassical scaling, are very far from one another compared to the characteristic scales of the

breathers themselves. All of these breathers rest upon a common nonzero background field that is asymptotically constant on length scales that contain arbitrarily many breathers, although on larger spatial scales this background has to decay to zero to match the mathematically imposed boundary conditions. This kind of pattern of waves might seem to be a strictly theoretical phenomenon, and yet it has recently also been observed in optical experiments [38].

There are many papers that describe generalizations of the elementary Peregrine solution (1.5), making use of the integrable structure behind the focusing NLS equation (1.1) to construct “higher-order” versions of  $\psi_P(x, t)$  by mostly algebraic means such as the Hirota method or Darboux-like transformations leading to determinantal or Wronskian formulae for exact solutions with interesting properties. An interesting point that was observed from the beginning is that it is not possible to generate  $\psi_P(x, t)$  or its higher-order analogues from the standard Darboux transformation method. Instead, one must take certain limits within the standard Darboux transformation method as was done in [2] and put into the framework of a generalized Darboux scheme in [23]. See also [4, 17, 21, 22, 26]. The fact that limits are required at some level mirrors what is also required to obtain these solutions via the inverse-scattering method as it is currently understood.

Despite all of this activity, and all of the knowledge available to date on the Peregrine solution and its generalizations mentioned above, *none of these solutions have been obtained directly as the solution of a Cauchy problem of the form (1.1)–(1.3) via an inverse-scattering transform.* This makes it challenging to address questions that would seem natural from the point of view of soliton theory:

- Are rogue wave solutions stable to localized perturbations? If not, what kinds of perturbations excite instabilities, and what is the long-term nonlinear saturation of the instabilities?
- Is it possible to make a prediction based on the computation of some relevant scattering data as to how many Peregrine-like peaks will be generated from a localized perturbation of the background? In other words, what kind of initial conditions generate rogue waves?
- How do rogue waves interact with other coherent structures, such as the time-periodic Kuznetsov-Ma soliton [30, 32] and its moving generalization sometimes called the Tajiri-Watanabe soliton [37]? How do they interact with more general waves that are not in the realm of “exact solutions”?
- Since it is possible to define mathematically a “rogue wave of order  $N$ ” for arbitrary  $N$ , what can be said quantitatively about the asymptotics for large  $N$ ?

In this paper, we aim to resolve this difficulty by generalizing the inverse-scattering solution for the Cauchy problem (1.1)–(1.3) currently in the literature to capture rogue waves. As will be seen, rogue waves correspond to a particular type of *spectral singularity*; the method we propose handles these and other types of spectral singularities, as well as poles of arbitrary order, in a universal framework.

Moreover, the method also makes taking limits [2] or related constructions such as the generalized form of the Darboux transformation [23] unnecessary in generating high-order rogue waves from the background field, and it yields an analytical setting in which the large- $N$  asymptotics of  $N^{\text{th}}$ -order rogue wave solutions can be effectively studied [7].

Before we proceed, we give a remark on notational conventions used in this paper.

*Remark 1.1* (Notational convention). We denote by  $a^*$  the complex conjugate of a complex number  $a$ . When used with a matrix,  $\mathbf{A}^*$  denotes the element-wise complex conjugate, without transposition. We use the “dagger” symbol  $\mathbf{A}^\dagger$  to denote the conjugate transpose (Hermitian) of a matrix  $\mathbf{A}$ . In this paper,  $\bar{a}$  is not used to denote the complex conjugate of a quantity  $a$ . With the exception of the Pauli matrices defined by

$$(1.7) \quad \sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and the identity matrix  $\mathbb{I}$ , we use boldface capital letters to denote matrices and boldface lowercase letters to denote column vectors.

For two square matrices  $\mathbf{A}$  and  $\mathbf{B}$ , we use  $[\mathbf{A}, \mathbf{B}]$  to denote the commutator  $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$ . For two  $n \times 1$  (column) vectors  $\mathbf{u}$  and  $\mathbf{v}$ , we denote the  $n \times 2$  matrix whose columns are  $\mathbf{u}$  and  $\mathbf{v}$  by  $[\mathbf{u}; \mathbf{v}]$ . Finally, we use  $[c_1 : c_2]^\top$  for  $c_1, c_2 \in \mathbb{C}$  to denote a column vector (hence the transpose) whose elements are the homogeneous coordinates for a point in the projective space  $\mathbb{C}\mathbb{P}^1$  (see Remark 3.4 for the occurrence of  $\mathbb{C}\mathbb{P}^1$ ).

### 1.1 Rogue Waves and the Known Inverse-Scattering Transform for the Cauchy Problem 1.1–1.3

To show what goes wrong when rogue waves are considered in the inverse-scattering transform, we begin by reviewing the method that is already in the literature. The existing method was introduced by Ma [32], further elucidated by Faddeev and Takhtajan [18], and substantially developed by Biondini and Kovačič [8], where the inverse problem was formulated as a matrix Riemann-Hilbert problem. This method has subsequently been used to prove a remarkable result [9], namely that under some conditions the solution of the Cauchy problem (1.1)–(1.3) takes on a universal form as  $t \rightarrow \infty$  that is independent of the initial condition (1.3).

The focusing NLS equation in the form (1.1) is the  $\lambda$ -independent compatibility condition for the simultaneous linear equations of a Lax pair [41]

$$(1.8) \quad \mathbf{u}_x = \mathbf{X}(\lambda; x, t)\mathbf{u}, \quad \mathbf{X}(\lambda; x, t) := \begin{bmatrix} -i\lambda & \psi \\ -\psi^* & i\lambda \end{bmatrix},$$

and

$$(1.9) \quad \mathbf{u}_t = \mathbf{T}(\lambda; x, t)\mathbf{u},$$

$$\mathbf{T}(\lambda; x, t) := \begin{bmatrix} -i\lambda^2 + i\frac{1}{2}(|\psi|^2 - 1) & \lambda\psi + i\frac{1}{2}\psi_x \\ -\lambda\psi^* + i\frac{1}{2}\psi_x^* & i\lambda^2 - i\frac{1}{2}(|\psi|^2 - 1) \end{bmatrix},$$

governing an auxiliary vector  $\mathbf{u}$  (or more generally, a column of a fundamental solution matrix) that depends on  $(x, t) \in \mathbb{R}^2$  and the complex spectral parameter  $\lambda \in \mathbb{C}$ . The existence of simultaneous solutions of this Lax pair constitutes the basis of the inverse-scattering transform (IST) method to solve the Cauchy initial-value problem (1.1)–(1.3).

**Review of the Inverse-Scattering Transform**

We now describe the method as it was formulated in [9], in which the inverse problem is set in the complex  $\lambda$ -plane rather than in terms of a uniformization map to the Riemann sphere [8].

DIRECT TRANSFORM. The Zakharov-Shabat problem (1.8) with the background field  $\psi = \psi_{\text{bg}}(x, t) \equiv 1$  is a constant-coefficient first-order system and hence can be solved by diagonalization (or the eigenvalue method) to obtain a fundamental matrix of solutions:

$$(1.10) \quad \mathbf{U}_{\text{bg}}(\lambda; x) := n(\lambda) \begin{bmatrix} 1 & i(\lambda - \rho(\lambda)) \\ i(\lambda - \rho(\lambda)) & 1 \end{bmatrix} e^{-i\rho(\lambda)x\sigma_3}$$

$$=: \mathbf{E}(\lambda) e^{-i\rho(\lambda)x\sigma_3},$$

where  $\rho(\lambda)$  is determined from the equation  $\rho^2 = \lambda^2 + 1$  (the characteristic equation for eigenvalues  $i\rho$  of the constant coefficient matrix in (1.8) with  $\psi \equiv 1$ ). To be concrete, we suppose that  $\rho(\lambda)$  is the function analytic for complex  $\lambda$  with the exception of a vertical branch cut  $\Sigma_c$  between the branch points  $\lambda = \pm i$ , whose square coincides with  $\lambda^2 + 1$  and that satisfies  $\rho(\lambda) = \lambda + O(\lambda^{-1})$  as  $\lambda \rightarrow \infty$ . The scalar complex-valued function  $n(\lambda)$  in (1.10) is well-defined as the function analytic for  $\lambda \in \mathbb{C} \setminus \Sigma_c$  satisfying

$$(1.11) \quad n(\lambda)^2 = \frac{\lambda + \rho(\lambda)}{2\rho(\lambda)}, \quad \lim_{\lambda \rightarrow \infty} n(\lambda) = 1.$$

The normalization factor  $n(\lambda)$  ensures that  $\det(\mathbf{U}_{\text{bg}}(\lambda; x)) = 1$ . Note that each matrix element of the unimodular matrix  $\mathbf{E}(\lambda)$  (and hence those of the fundamental matrix  $\mathbf{U}_{\text{bg}}(\lambda; x)$ ) has (mild) singularities at the (branch) points  $\lambda = \pm i$ , where all four matrix elements blow up as  $|\lambda \mp i|^{-1/4}$ . We also observe that  $\mathbf{U}_{\text{bg}}(\lambda; x)$  and  $\mathbf{E}(\lambda)$  are well-defined as analytic functions for  $\lambda \in \mathbb{C} \setminus \Sigma_c$ , but we may consider them as taking two different values on  $\Sigma_c \setminus \{-i, i\}$ , corresponding to evaluating  $\rho(\lambda)$  and  $n(\lambda)$  as boundary values taken on one side or the other.

Roughly speaking, the *continuous spectrum*  $\Gamma$  for the direct problem consists of those values of  $\lambda$  for which  $\rho(\lambda)$  is real. Since  $\rho(\lambda)$  takes distinct real boundary

values on the branch cut  $\Sigma_c$ , it will be more convenient to distinguish these boundary values by defining  $\Gamma$  as the union of the real intervals  $(-\infty, 0)$  and  $(0, +\infty)$  with the boundary of the slit domain  $\mathbb{C} \setminus \Sigma_c$ , which is topologically a circle. So, as a set of points in the  $\lambda$ -plane, we might write  $\Gamma := \mathbb{R} \cup \Sigma_c$ , but with each  $\lambda \in \Sigma_c \setminus \{-i, i\}$  we actually associate two distinct points of  $\Gamma$ , corresponding to the two boundary values of  $\rho(\lambda)$ . From here onwards, when we define or discuss quantities for  $\lambda \in \Gamma \cap (\Sigma_c \setminus \{-i, i\})$ , we are doing so with two distinct boundary values of  $\rho(\lambda)$  and  $n(\lambda)$  corresponding to distinct points on the boundary of the slit domain  $\mathbb{C} \setminus \Sigma_c$ .

Suppose that  $t$  is fixed and  $\Delta\psi(x, t) := \psi(x, t) - 1$  is an absolutely integrable complex-valued function of  $x \in \mathbb{R}$ . For  $\lambda \in \Gamma \setminus \{-i, i\}$ , the *Jost matrix solutions*  $\mathbf{U} = \mathbf{J}^\pm(\lambda; x, t)$  of (1.8) are defined uniquely (modulo the interpretation of  $\lambda \in \Gamma \cap (\Sigma_c \setminus \{-i, i\})$  discussed above) by the boundary conditions

$$(1.12) \quad \mathbf{J}^\pm(\lambda; x, t) e^{i\rho(\lambda)x\sigma_3} = \mathbf{E}(\lambda) + o(1), \quad x \rightarrow \pm\infty,$$

and through the renormalization  $\mathbf{K}^\pm(\lambda; x, t) := \mathbf{J}^\pm(\lambda; x, t) e^{i\rho(\lambda)x\sigma_3}$ , they can be obtained from the unique solutions of the Volterra integral equations

$$(1.13) \quad \begin{aligned} &\mathbf{K}^\pm(\lambda; x, t) \\ &= \mathbf{E}(\lambda) + \int_{\pm\infty}^x \mathbf{E}(\lambda) e^{-i\rho(\lambda)(x-y)\sigma_3} \mathbf{E}(\lambda)^{-1} \Delta\Psi(y, t) \\ &\quad \cdot \mathbf{K}^\pm(\lambda; y, t) e^{i\rho(\lambda)(x-y)\sigma_3} dy, \quad \rho(\lambda) \in \mathbb{R}, \end{aligned}$$

where

$$\Delta\Psi(x, t) := \begin{bmatrix} 0 & \Delta\psi(x, t) \\ -\Delta\psi(x, t)^* & 0 \end{bmatrix}.$$

Because  $e^{i\rho(\lambda)x\sigma_3}$  is a diagonal matrix, the following are consequences of standard analysis of the iterates that (see, for example, [8] and the references therein):

- The first column  $\mathbf{j}^{-,1}(\lambda; x, t)$  of the matrix  $\mathbf{J}^-(\lambda; x, t)$  and the second column  $\mathbf{j}^{+,2}(\lambda; x, t)$  of the matrix  $\mathbf{J}^+(\lambda; x, t)$  are the boundary values of vector-valued functions of  $\lambda$  analytic in the domain  $\mathbb{C}^+ \setminus \Sigma_c$  (corresponding to  $\text{Im}\{\rho(\lambda)\} > 0$ ). For  $\lambda \in \Sigma_c \setminus \{-i, i\}$  this is only relevant for  $\text{Im}\{\lambda\} \geq 0$ , in which case the statement refers to the columns of the Jost matrices associated with both boundary values of  $\rho(\lambda)$  and  $\mathbf{E}(\lambda)$  as are needed to traverse the boundary of the slit domain  $\mathbb{C}^+ \setminus \Sigma_c$ .
- The first column  $\mathbf{j}^{+,1}(\lambda; x, t)$  of the matrix  $\mathbf{J}^+(\lambda; x, t)$  and the second column  $\mathbf{j}^{-,2}(\lambda; x, t)$  of the matrix  $\mathbf{J}^-(\lambda; x, t)$  are the boundary values of vector-valued functions of  $\lambda$  analytic in the domain  $\mathbb{C}^- \setminus \Sigma_c$  (corresponding to  $\text{Im}\{\rho(\lambda)\} < 0$ ). For  $\lambda \in \Sigma_c \setminus \{-i, i\}$  this is only relevant for  $\text{Im}\{\lambda\} \leq 0$ , with a similar caveat.

Using an analogous argument in [16] for the defocusing problem, one can show that if also  $x^2\Delta\psi(x) \in L^1(\mathbb{R})$ , then  $\mathbf{j}^{-,1}$  and  $\mathbf{j}^{+,2}$  are  $O((\lambda - i)^{-1/4})$  as  $\lambda \rightarrow i$ , while  $\mathbf{j}^{+,1}$  and  $\mathbf{j}^{-,2}$  are  $O((\lambda + i)^{-1/4})$  as  $\lambda \rightarrow -i$ . Moreover,  $\det(\mathbf{J}^\pm(\lambda; x, t)) = 1$

for  $\lambda \in \Gamma \setminus \{-i, i\}$ , and  $\mathbf{J}^\pm(\lambda; x, t)$  are both fundamental matrices of solutions of (1.8) with  $\psi = \psi(x, t)$  for  $\lambda \in \Gamma \setminus \{-i, i\}$ . Thus, they satisfy the *scattering relation*  $\mathbf{J}^+(\lambda; x, t) = \mathbf{J}^-(\lambda; x, t)\mathbf{S}(\lambda; t)$ , where  $\mathbf{S}(\lambda; t)$  is called the *scattering matrix* defined for  $\lambda \in \Gamma \setminus \{-i, i\}$  (in the generalized sense of  $\Gamma$  described above), and  $\det(\mathbf{S}(\lambda; t)) = 1$ . We may write the scattering matrix in the following form:

$$(1.14) \quad \mathbf{S}(\lambda; t) =: \begin{bmatrix} \bar{a}(\lambda; t) & \bar{b}(\lambda; t) \\ -b(\lambda; t) & a(\lambda; t) \end{bmatrix}$$

with

$$(1.15) \quad \begin{aligned} a(\lambda; t) &:= \det([\mathbf{j}^{-,1}(\lambda; x, t); \mathbf{j}^{+,2}(\lambda; x, t)]), \\ \bar{a}(\lambda; t) &:= \det([\mathbf{j}^{+,1}(\lambda; x, t); \mathbf{j}^{-,2}(\lambda; x, t)]), \\ b(\lambda; t) &:= \det([\mathbf{j}^{+,1}(\lambda; x, t); \mathbf{j}^{-,1}(\lambda; x, t)]), \\ \bar{b}(\lambda; t) &:= \det([\mathbf{j}^{+,2}(\lambda; x, t); \mathbf{j}^{-,2}(\lambda; x, t)]), \end{aligned}$$

and the fact that these determinants are independent of  $x$  can be seen another way as a consequence of Abel’s theorem.

These Wronskian formulae show that  $a(\lambda; t)$  and  $\bar{a}(\lambda; t)$  extend analytically to the domains  $\text{Im}\{\rho(\lambda)\} > 0$  and  $\text{Im}\{\rho(\lambda)\} < 0$ , respectively, but  $b(\lambda; t)$  and  $\bar{b}(\lambda; t)$  do not necessarily enjoy analytic continuation from  $\Gamma$  in any direction. Moreover,  $a(\lambda; t), \bar{a}(\lambda; t) \rightarrow 1$  as  $\lambda \rightarrow \infty$  in  $\mathbb{C}^+, \mathbb{C}^-$  respectively [8]. In fact, we have the following estimate for the rate at which  $a(\lambda) \rightarrow 1$ .

LEMMA 1.2. *Suppose that  $\Delta\psi(\cdot, t) := \psi(\cdot, t) - 1 \in L^1(\mathbb{R})$ , and  $\Delta\psi_x(\cdot, t) \in L^1(\mathbb{R})$ . Then  $a(\lambda; t) - 1 = O(\lambda^{-1})$  uniformly in  $\mathbb{C}^+$ , and moreover,  $|a(\lambda; t)| > \frac{1}{2}$  provided that  $|\lambda| \geq r[\Delta\psi]$ , where  $r[\Delta\psi]$  is defined by (B.25).*

We give the proof in Appendix B. If  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$ , then it follows from the estimates of the Jost solutions for  $\lambda \approx i$  and the Wronskian formula (1.15) for  $a$  that  $a(\lambda) = O((\lambda - i)^{-1/2})$  as  $\lambda \rightarrow i$ . In the special case that  $\lambda = i$  is a “virtual level” [8, app. B], one can show under a slightly stronger decay assumption on  $\Delta\psi$  that  $a(\lambda) = O(1)$  as  $\lambda \rightarrow i$  (this is the case for the background potential, for instance, for which  $a(\lambda) \equiv 1$ ). The ratio

$$(1.16) \quad R(\lambda; t) := b(\lambda; t)/a(\lambda; t), \quad \lambda \in \Gamma \setminus \{-i, i\},$$

(for  $\Gamma$  interpreted in the generalized sense) is called the *reflection coefficient* and it does not necessarily have any analytic extension from  $\Gamma$ . We also define  $\bar{R}(\lambda; t) := \bar{b}(\lambda; t)/\bar{a}(\lambda; t)$  for  $\lambda \in \Gamma$ .

The coefficient matrix  $\mathbf{X}(\lambda; x, t)$  in (1.8) satisfies the Schwarz symmetry

$$(1.17) \quad \sigma_2 \mathbf{X}(\lambda^*; x, t)^* \sigma_2 = \mathbf{X}(\lambda; x, t).$$

This implies that if  $\mathbf{u}(\lambda; x, t)$  is a solution of (1.8) for  $\lambda \in \Gamma \setminus \{-i, i\}$ , then so is  $i\sigma_2 \mathbf{u}(\lambda^*, x, t)^*$ . Since the Jost matrices  $\mathbf{J}^\pm(\lambda; x, t)$  are uniquely determined by their boundary conditions (1.12) for  $\lambda \in \Gamma \setminus \{-i, i\}$ , it follows that

$$(1.18) \quad \mathbf{J}^\pm(\lambda; x, t) = \sigma_2 \mathbf{J}^\pm(\lambda^*; x, t)^* \sigma_2, \quad \lambda \in \Gamma \setminus \{-i, i\},$$



where we used the property that  $\rho(\lambda) = \rho(\lambda^*)^*$  along with  $\mathbf{E}(\lambda) = \sigma_2 \mathbf{E}(\lambda^*)^* \sigma_2$ . Thus the scattering matrix also enjoys the same Schwarz symmetry on its domain, which implies that  $\bar{a}(\lambda; t) = a(\lambda^*; t)^*$  and  $\bar{b}(\lambda; t) = b(\lambda^*; t)^*$  for  $\lambda \in \Gamma \setminus \{-i, i\}$ . Moreover, the analytic continuations admitted by the Jost column vector solutions of (1.8) also satisfy the following symmetries in the indicated domains:

$$(1.19) \quad \mathbf{j}^{-,1}(\lambda; x, t) = i\sigma_2 \mathbf{j}^{-,2}(\lambda^*; x, t)^*, \quad \lambda \in \mathbb{C}^+ \setminus \Sigma_c,$$

$$(1.20) \quad \mathbf{j}^{+,1}(\lambda; x, t) = i\sigma_2 \mathbf{j}^{+,2}(\lambda^*; x, t)^*, \quad \lambda \in \mathbb{C}^- \setminus \Sigma_c.$$

These relations extend the symmetry  $\bar{a}(\lambda; t) = a(\lambda^*; t)^*$  to the values of  $\lambda \in \mathbb{C}^- \setminus \Sigma_c$ .

Assuming for convenience of exposition that  $a(\lambda; t) \neq 0$  for all  $\lambda \in \Gamma \setminus \{-i, i\}$  and that its analytic extension to  $\mathbb{C}^+ \setminus \Sigma_c$  has only finitely many simple zeros  $\xi_1(t), \dots, \xi_N(t)$ ,<sup>1</sup> the *scattering data* associated with  $\psi(\cdot, t)$  is

$$(1.21) \quad \mathcal{S}(\psi(\cdot, t)) := \{R(\lambda; t) \text{ for } \lambda \in \Gamma \setminus \{-i, i\}, \{(\xi_j(t), \gamma_j(t))\}_{j=1}^N\},$$

where the nonzero proportionality constants  $\gamma_1(t), \dots, \gamma_N(t)$  are defined in terms of the analytic Jost columns by (cf. (1.15))

$$(1.22) \quad \mathbf{j}^{-,1}(\xi_j(t); x, t) = \gamma_j(t) \mathbf{j}^{+,2}(\xi_j(t); x, t).$$

**TIME DEPENDENCE.** Because  $\psi(x, t)$  is a solution of (1.1) by assumption, it is easy to verify that the unimodular matrices  $\mathbf{J}^\pm(\lambda; x, t) e^{-i\rho(\lambda)\lambda t \sigma_3}$  are both fundamental matrices of *simultaneous* solutions of the Lax pair (1.8)–(1.9) for  $\lambda \in \Gamma \setminus \{-i, i\}$ , and from this follows the time evolution of the scattering matrix:

$$(1.23) \quad \begin{aligned} \mathbf{S}(\lambda; t) &= e^{-i\rho(\lambda)\lambda t \sigma_3} \mathbf{S}(\lambda; 0) e^{i\rho(\lambda)\lambda t \sigma_3}, \\ \mathbf{S}(\lambda; 0) &:= \mathbf{J}^-(\lambda; x, 0)^{-1} \mathbf{J}^+(\lambda; x, 0). \end{aligned}$$

From this formula, it follows that  $a(\lambda; t) = a(\lambda; 0) =: a(\lambda)$ , and hence the number of zeros  $N$  is also constant as are the zeros themselves:  $\xi_j(t) = \xi_j(0) =: \xi_j$ ,  $j = 1, \dots, N$ . It also follows that  $R(\lambda; t) = R(\lambda; 0) e^{2i\rho(\lambda)\lambda t}$  holds for  $\lambda \in \Gamma \setminus \{-i, i\}$ . Finally, it can be shown that  $\gamma_j(t) = \gamma_j(0) e^{2i\rho(\xi_j)\xi_j t}$ ,  $j = 1, \dots, N$ . Therefore, the scattering data  $\mathcal{S}(\psi(\cdot, t))$  enjoys explicit and elementary dependence on  $t$  when  $\psi(\cdot, t)$  evolves in time according to the Cauchy problem (1.1)–(1.3).

**INVERSE TRANSFORM.** The *Beals-Coifman* simultaneous solution of (1.8)–(1.9) is the sectionally meromorphic matrix-valued function

$$(1.24) \quad \mathbf{U}^{\text{BC}}(\lambda; x, t) := \begin{cases} \left[ \begin{array}{cc} a(\lambda)^{-1} \mathbf{j}^{-,1}(\lambda; x, t) e^{-i\rho(\lambda)\lambda t} & ; & \mathbf{j}^{+,2}(\lambda; x, t) e^{i\rho(\lambda)\lambda t} \end{array} \right], & \lambda \in \mathbb{C}^+ \setminus \Sigma_c, \\ \left[ \begin{array}{cc} \mathbf{j}^{+,1}(\lambda; x, t) e^{-i\rho(\lambda)\lambda t} & ; & \bar{a}(\lambda)^{-1} \mathbf{j}^{-,2}(\lambda; x, t) e^{i\rho(\lambda)\lambda t} \end{array} \right], & \lambda \in \mathbb{C}^- \setminus \Sigma_c, \end{cases}$$

<sup>1</sup> This assumption can be removed; indeed, the robust transform we shall introduce later renders it obsolete.

and it is convenient to introduce the related function<sup>2</sup>

$$(1.26) \quad \mathbf{M}^{\text{BC}}(\lambda; x, t) := \mathbf{U}^{\text{BC}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \mathbb{C} \setminus \Gamma.$$

$\mathbf{M}^{\text{BC}}(\lambda; x, t)$  has simple poles at the points  $\{\xi_1, \dots, \xi_N\}$  and  $\{\xi_1^*, \dots, \xi_N^*\}$ , the zeros of  $a(\lambda)$  and  $\bar{a}(\lambda) := a(\lambda^*)^*$ , respectively, and no other singularities in its domain of definition. It satisfies the normalization condition

$$(1.27) \quad \lim_{\lambda \rightarrow \infty} \mathbf{M}^{\text{BC}}(\lambda; x, t) = \mathbb{I},$$

and  $\det(\mathbf{M}(\lambda; x, t)) = 1$  holds identically. It also satisfies the Schwarz symmetry condition

$$(1.28) \quad \mathbf{M}^{\text{BC}}(\lambda^*; x, t) = \sigma_2 \mathbf{M}^{\text{BC}}(\lambda; x, t)^* \sigma_2, \quad \lambda \in \mathbb{C} \setminus \Gamma.$$

The boundary values taken by  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  on  $\Gamma = \mathbb{R} \cup \Sigma_c$  from either side are related by the following *jump conditions*. For  $\lambda \in \mathbb{R}, \lambda \neq 0$ ,  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  has well-defined nontangential boundary values

$$\mathbf{M}_{\pm}^{\text{BC}}(\lambda; x, t) := \lim_{\varepsilon \downarrow 0} \mathbf{M}^{\text{BC}}(\lambda \pm i\varepsilon; x, t);$$

from the scattering relation  $\mathbf{J}^+ = \mathbf{J}^- \mathbf{S}$  and the definition of  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  we have

$$(1.29) \quad \mathbf{M}_{+}^{\text{BC}}(\lambda; x, t) = \mathbf{M}_{-}^{\text{BC}}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^{\mathbb{R}}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3},$$

$$\lambda \in \mathbb{R} \setminus \{0\},$$

where

$$(1.30) \quad \mathbf{V}^{\mathbb{R}}(\lambda) := \begin{bmatrix} 1 + |R(\lambda; 0)|^2 & R(\lambda; 0)^* \\ R(\lambda; 0) & 1 \end{bmatrix}, \quad \lambda \in \mathbb{R} \setminus \{0\}.$$

For  $\lambda = i \operatorname{Im}\{\lambda\}$  with  $0 < \operatorname{Im}\{\lambda\} < 1$ ,  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  takes well-defined boundary values given by

$$\mathbf{M}_{\pm}^{\text{BC}}(\lambda; x, t) := \lim_{\varepsilon \downarrow 0} \mathbf{M}^{\text{BC}}(\lambda \pm \varepsilon; x, t)$$

related by

$$(1.31) \quad \mathbf{M}_{+}^{\text{BC}}(\lambda; x, t) = \mathbf{M}_{-}^{\text{BC}}(\lambda; x, t) e^{-i\rho-(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^{\downarrow}(\lambda) e^{i\rho+(\lambda)(x+\lambda t)\sigma_3},$$

$$\lambda = i \operatorname{Im}\{\lambda\}, \quad 0 < \operatorname{Im}\{\lambda\} < 1,$$

---

<sup>2</sup>By (1.19) and (1.20), for  $\lambda \in \mathbb{C}^+ \setminus \Sigma_c$  with  $a(\lambda) \neq 0$ , the matrix function  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  is uniquely determined by the fact that the columns of the associated matrix  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  satisfy the differential equation (1.8) and that boundary conditions

$$(1.25) \quad \mathbf{M}^{\text{BC}}(\lambda; x, t) \rightarrow \mathbb{I} \text{ as } x \rightarrow +\infty,$$

$$\mathbf{M}^{\text{BC}}(\lambda; x, t) \text{ is bounded as } x \rightarrow -\infty,$$

hold. These conditions can be combined into an integral equation for  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  of Fredholm type (see, for example, [5] for details).

where

$$(1.32) \quad \mathbf{V}^\downarrow(\lambda) := \begin{bmatrix} i\bar{R}_-(\lambda; 0) & -i \\ -i(1 + R_-(\lambda; 0)\bar{R}_-(\lambda; 0)) & iR_-(\lambda; 0) \end{bmatrix}, \\ \lambda = i \operatorname{Im}\{\lambda\}, \quad 0 < \operatorname{Im}\{\lambda\} < 1.$$

Here,  $\rho_\pm(\lambda) := \lim_{\varepsilon \downarrow 0} \rho(\lambda \pm \varepsilon)$ , and  $R_-(\lambda; 0)$  and  $\bar{R}_-(\lambda; 0)$  refer to the elements of the scattering matrix constructed from Jost solutions defined by using the boundary value  $\rho_-(\lambda)$  as well as the corresponding boundary value  $\mathbf{E}_-(\lambda)$  taken by  $\mathbf{E}(\cdot)$ . Finally, for  $\lambda = i \operatorname{Im}\{\lambda\}$  with  $-1 < \operatorname{Im}\{\lambda\} < 0$ ,  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  takes boundary values

$$\mathbf{M}_\pm^{\text{BC}}(\lambda; x, t) := \lim_{\varepsilon \downarrow 0} \mathbf{M}^{\text{BC}}(\lambda \mp \varepsilon; x, t);$$

these are related by

$$(1.33) \quad \mathbf{M}_+^{\text{BC}}(\lambda; x, t) = \mathbf{M}_-^{\text{BC}}(\lambda; x, t) e^{-i\rho_-(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^\uparrow(\lambda) e^{i\rho_+(\lambda)(x+\lambda t)\sigma_3}, \\ \lambda = i \operatorname{Im}\{\lambda\}, \quad -1 < \operatorname{Im}\{\lambda\} < 0,$$

where

$$(1.34) \quad \mathbf{V}^\uparrow(\lambda) := \mathbf{V}^\downarrow(\lambda^*)^\dagger, \quad \lambda = i \operatorname{Im}\{\lambda\}, \quad -1 < \operatorname{Im}\{\lambda\} < 0.$$

The simple poles of  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  are characterized by the following *residue conditions*:

$$(1.35) \quad \operatorname{Res}_{\lambda=\xi_j} \mathbf{M}^{\text{BC}}(\lambda; x, t) = \lim_{\lambda \rightarrow \xi_j} (\mathbf{M}^{\text{BC}}(\lambda; x, t) e^{-i\rho(\xi_j)(x+\xi_j t)\sigma_3} \cdot \mathbf{N}_j e^{i\rho(\xi_j)(x+\xi_j t)\sigma_3}), \quad j = 1, \dots, N, \\ \operatorname{Res}_{\lambda=\xi_j^*} \mathbf{M}^{\text{BC}}(\lambda; x, t) = \lim_{\lambda \rightarrow \xi_j^*} (\mathbf{M}^{\text{BC}}(\lambda; x, t) e^{-i\rho(\xi_j^*)(x+\xi_j^* t)\sigma_3} \cdot \sigma_2 \mathbf{N}_j^* \sigma_2 e^{i\rho(\xi_j^*)(x+\xi_j^* t)\sigma_3}), \quad j = 1, \dots, N,$$

where

$$(1.36) \quad \mathbf{N}_j := \begin{bmatrix} 0 & 0 \\ a'(\xi_j)^{-1} \gamma_j(0) & 0 \end{bmatrix}, \quad j = 1, \dots, N.$$

The analyticity of  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  in the domain  $\mathbb{C} \setminus (\Gamma \cup \{\xi_1, \dots, \xi_N, \xi_1^*, \dots, \xi_N^*\})$ , the normalization condition  $\mathbf{M}^{\text{BC}}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$ , the jump conditions (1.29), (1.31), and (1.33), and residue conditions (1.35) are all conditions on  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  involving the scattering data at  $t = 0$  that can be computed from the initial condition (1.3) alone via the direct transform. Moreover, these Riemann-Hilbert conditions are nearly enough to determine  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$ . The only missing piece of information concerns the nature of the boundary values taken by  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  on  $\Gamma$ . The direct transform shows that these boundary values are continuous functions of  $\lambda$  with the exception of the points  $\{-i, i\}$ . These points are singularities of  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$ , as must be the case because  $\mathbf{V}^\downarrow(\lambda)$  does not tend to the identity as  $\lambda \rightarrow i$ . Moreover, in the simplest case of the background potential

$\psi_{\text{bg}}(x, 0) \equiv 1$ , the Beals-Coifman matrix is  $\mathbf{M}_{\text{bg}}^{\text{BC}}(\lambda; x, t) = \mathbf{E}(\lambda)$ , which exhibits  $-\frac{1}{4}$  power singularities at  $\lambda = \pm i$ . We therefore supplement the Riemann-Hilbert conditions with an additional growth condition to account for this singularity. Taking into account the behavior in the background case, we impose the following growth condition:

$$(1.37) \quad \mathbf{M}^{\text{BC}}(\lambda; x, t) = O((\lambda \mp i)^{-1/4}), \quad \lambda \rightarrow \pm i.$$

Indeed, under the assumption that  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$  and that  $a(i) \neq 0$  if  $\lambda = i$  is a “virtual level” [8, app. B], this condition correctly estimates the growth rate of  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  as  $\lambda \rightarrow \pm i$ .<sup>3</sup> Another way to formulate (1.37) is simply to require that the boundary values of  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  lie in  $L^2(\Gamma)$ . The growth condition is not explicitly stated in [9], but it is necessary to ensure uniqueness for the inverse problem.

**SUMMARY OF THE IST.** The IST method to solve the Cauchy initial-value problem (1.1)–(1.3) consists of the following steps. Given suitable initial data  $\psi_0(x) = \psi(x; 0)$  for  $\lambda \in \Gamma \setminus \{-i, i\}$ , one calculates the Jost matrices  $\mathbf{J}^\pm(\lambda; x, 0)$  and hence  $\mathbf{S}(\lambda; 0)$ . Assuming finitely many simple zeros  $\xi_1, \dots, \xi_N$  of  $a(\cdot) = a(\cdot; 0)$  in  $\mathbb{C}^+ \setminus \Sigma_c$ , one also computes the proportionality constants  $\gamma_1(0), \dots, \gamma_N(0)$  using (1.22). This completes the calculation of the scattering data at  $t = 0$  and constitutes the direct transform.

Then one allows the scattering data to evolve explicitly in time  $t > 0$  and uses it to formulate the inverse problem, namely seeking  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  that satisfies the Riemann-Hilbert conditions augmented with (1.37). Under some technical conditions that are not of concern here, this Riemann-Hilbert problem has a unique solution, and thus  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  is constructed for all  $(x, t) \in \mathbb{R}^2$  from the scattering data obtained from the initial condition (1.3). Finally, from  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  one obtains a function  $\psi(x, t)$  satisfying (1.1) by the limit

$$(1.38) \quad \psi(x, t) = 2i \lim_{\lambda \rightarrow \infty} \lambda M_{12}^{\text{BC}}(\lambda; x, t).$$

Under some additional assumptions that in particular exclude rogue waves,  $\psi(x, t)$  is then the solution of the Cauchy initial-value problem (1.1)–(1.3).

### Spectral Singularities

*Spectral singularities* are points  $\lambda$  in the continuous spectrum  $\Gamma$  at which the matrix  $\mathbf{M}^{\text{BC}}(\cdot; x, t)$  fails to have a well-defined nontangential limit from one side or the other. There are potentially two kinds of spectral singularities in this problem.

One type of spectral singularity is a zero of the scattering coefficient  $a(\cdot)$  ( $\bar{a}(\cdot)$  respectively) at a point  $\lambda \neq i$  ( $\lambda \neq -i$  respectively) on the boundary of the domain  $\mathbb{C}^+ \setminus \Sigma_c$  ( $\mathbb{C}^- \setminus \Sigma_c$  respectively). Since the Jost solutions  $\mathbf{j}^{\mp,1}(\cdot; x, t)$  and

<sup>3</sup> It seems to us that it might be possible for  $\lambda = i$  to be a virtual level and also to have  $a(i) = 0$  for some special potentials  $\Delta\psi$  with  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$ , in which case (1.37) is violated. The robust transform we introduce in Section 2 will sidestep this difficulty as well.

$\mathbf{j}^{\pm,2}(\cdot; x, t)$  are not only analytic in  $\mathbb{C}^{\pm} \setminus \Sigma_c$  but also continuous up to the boundary except at  $\lambda = \pm i$ , from (1.24) and (1.26) one sees that such zeros of  $a(\cdot)$  or  $\bar{a}(\cdot)$  are indeed spectral singularities and they are the only possible ones that can occur for  $\lambda \neq \pm i$ . From one point of view, zeros of  $a$  or  $\bar{a}$  in  $\Gamma \setminus \{-i, i\}$  are to be expected because two initial conditions without such spectral singularities but with different values of  $N$ , the number of zeros of  $a$  in  $\mathbb{C}^+ \setminus \Sigma_c$  weighted by multiplicity, can be connected by a suitable homotopy in which there must be at least one point at which a zero of finite multiplicity is “born” from  $\Gamma$ . An even more disturbing situation was described by Zhou [42] in the context of the simpler Cauchy problem for (1.1) with zero boundary conditions (1.6). For the latter problem (see [1, 5, 41] for details of the IST solution in this setting) it was shown by Beals and Coifman [5] that for an open dense subset of initial conditions  $\psi_0 \in L^1(\mathbb{R})$  there are no spectral singularities and the analogue of the number  $N$  is finite. On the other hand, as shown by Zhou [42, example 3.3.16], the complement of this dense open subset contains Schwartz-class functions  $\psi_0$  for which  $N = \infty$  and infinitely many zeros of  $a$  accumulate at certain points in the continuous spectrum  $\Gamma$  (for zero boundary conditions,  $\Gamma = \mathbb{R}$ ). Moreover, there exist Schwartz-class  $\psi_0$  for which there are infinitely many of these accumulation points, which themselves accumulate from within  $\Gamma$  at particularly severe spectral singularities.

In order to deal with spectral singularities in the zero boundary condition setting, Zhou [42] and Deift and Zhou [14] developed a method based on combining the standard Beals-Coifman matrix  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  for values of  $\lambda$  suitably large that the latter has no singularities with another simultaneous solution of the Lax pair (1.8)–(1.9) for smaller  $\lambda$ . This other solution matrix was constructed as the product of the Beals-Coifman matrix taken for a sufficiently “cutoff” version of  $\psi_0$  that it generates no singularities and evaluated for  $x$  equal to the cutoff point  $x = L$  with a transfer matrix solution of (1.8) to obtain the solution at a general  $x \in \mathbb{R}$  from that at  $x = L$ . This approach leads to a Riemann-Hilbert problem for a sectionally holomorphic matrix (no poles at all) that takes continuous boundary values on  $\Gamma = \mathbb{R}$  as well as on a circle of large radius where the two matrix solutions are related by a computable jump condition.

To complete this brief discussion of spectral singularities caused by zeros of  $a$  or  $\bar{a}$ , we simply make the point that while the sort of severe spectral singularities exhibited by Schwartz-class  $\psi_0$  in the zero boundary conditions case have not yet been observed in the presence of nonzero boundary conditions, there is no reason to assume that they cannot occur for certain  $\psi_0$  with  $\Delta\psi_0 := \psi_0 - 1$  Schwartz-class.

The other spectral singularities in the IST solution of the Cauchy problem (1.1)–(1.3) with nonzero boundary conditions are those at  $\lambda = \pm i$ . These are present due to the nonzero background  $\psi \equiv 1$ , of which the initial condition  $\psi_0$  is a localized perturbation. We have already pointed out that the matrix  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) := \mathbf{E}(\lambda) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$ , which is the Beals-Coifman simultaneous solution matrix of (1.8)–(1.9) for  $\psi_0(x) \equiv 1$ , blows up like  $(\lambda \mp i)^{-1/4}$  as  $\lambda \rightarrow \pm i$ . These singularities generally propagate into the Jost matrix solutions  $\mathbf{J}^{\pm}(\lambda; x, t)$  through

the integral equations (1.13), which involve  $\mathbf{E}(\lambda)$ , although from (1.13) alone it is generally (i.e., in absence of an extra condition like  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$ ) difficult to assess how the severity of the singularities induced at  $\lambda = \pm i$  relates to  $\Delta\psi(x, t)$ . We wish to point out that the cutoff potential approach of [14, 43] does not sidestep the issue, because *these spectral singularities are features of the background potential itself*, and the cutoff approach simply replaces one potential by another one that is sufficiently close in  $L^1(\mathbb{R})$  to the background.

The reason that the background potential  $\psi_0(x) \equiv 1$  produces spectral singularities at  $\lambda = \pm i$  in the Beals-Coifman solution  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t)$  is simple: the constant coefficient matrices in the Lax pair (1.8)–(1.9) fail to be diagonalizable at  $\lambda = \pm i$ . Therefore, the eigenvalue method, on which the construction of the matrix  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t)$  is based, cannot produce a fundamental solution matrix when  $\lambda = \pm i$ . The spectral singularities thus appear upon normalizing the eigenvectors so that  $\det(\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t)) = 1$ . Note, however, that the essence of the difficulty lies in the collapse in the dimension of the eigenvector span, and this difficulty remains if one chooses to work with nonunimodular solution matrices or attempts to regularize the branch points at  $\lambda = \pm i$  by mapping the slit domain  $\mathbb{C} \setminus \Sigma_c$  to the Riemann sphere via the Joukowski mapping [8].

### Attempting to Capture Rogue Waves in the IST

The spectral singularities at  $\lambda = \pm i$  are the reason that the IST as described in Section 1.1 does not capture rogue waves. Here the example of the Peregrine solution  $\psi_{\text{P}}(x, t)$  (see (1.5)) is particularly instructive. Since  $\psi_{\text{P}}(\cdot, t) - 1 \in L^1(\mathbb{R})$  for all  $t \in \mathbb{R}$ , the direct transform applies. Rather than compute the Jost solutions via the Volterra equations (1.13), it is more convenient to appeal to algebraic methods such as the generalized Darboux transformation [23] to construct a basis of simultaneous solutions of (1.8)–(1.9) for  $\psi = \psi_{\text{P}}(x, t)$  and  $\lambda \in \Gamma \setminus \{-i, i\}$ . The result (see Corollary 3.7) is that  $\mathbf{S} = \mathbf{S}_{\text{P}}(\lambda; 0) \equiv \mathbb{I}$  for all  $\lambda \in \Gamma \setminus \{-i, i\}$ . Therefore, there are no poles of  $\mathbf{M}_{\text{P}}^{\text{BC}}(\lambda; x, t)$  in  $\mathbb{C} \setminus \Gamma$ , and the reflection coefficient is  $R = R_{\text{P}}(\lambda; 0) \equiv 0$ . This implies that *the Peregrine solution  $\psi = \psi_{\text{P}}(x, t)$  is indistinguishable from the background solution  $\psi_{\text{bg}}(x, t) \equiv 1$  at the level of the scattering data* (1.21); i.e., the direct transform  $\psi \mapsto \mathcal{S}(\psi)$  is not injective. Indeed,  $\mathcal{S}(\psi_{\text{P}}) = \mathcal{S}(\psi_{\text{bg}})$ . Although they share the same scattering data, the matrices  $\mathbf{U}_{\text{P}}^{\text{BC}}(\lambda; x, t)$  and  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t)$  do not coincide; they are distinguished by the rate at which they blow up as  $\lambda \rightarrow \pm i$ . The difficulty lies in working out how to take this asymptotic behavior of the matrix  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  as  $\lambda \rightarrow \pm i$  into account in the formulation of the IST. If one begins with the initial condition  $\psi_0(x) = \psi_{\text{P}}(x, 0)$  and enforces the growth condition (1.37) in the inverse problem, the IST returns instead the background solution  $\psi_{\text{bg}}(x, t) \equiv 1$ . However, if one simply replaces the growth condition (1.37) with the estimate  $\mathbf{M}^{\text{BC}}(\lambda; x, t) = O((\lambda \mp i)^{-3/4})$  as  $\lambda \rightarrow \pm i$  known to be correct for the matrix  $\mathbf{M}_{\text{P}}^{\text{BC}}(\lambda; x, t)$  (see Remark 3.8), then one loses the uniqueness of the solution of the inverse problem as  $\mathbf{M}_{\text{P}}^{\text{BC}}(\lambda; x, t)$  and  $\mathbf{M}_{\text{bg}}^{\text{BC}}(\lambda; x, t)$  are no longer distinguished by the Riemann-Hilbert conditions

augmented with the updated (weakened, really) growth condition. All of these remarks apply to the higher-order rogue waves [4] (see Figures 3.5 and 3.8) as well, although as the order increases, so does the rate at which  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  blows up as  $\lambda \rightarrow \pm i$ . Note that neither the Peregrine solution nor its higher-order analogues satisfy the condition  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$  that, with some additional technical assumptions related to virtual levels at  $\pm i$ , guarantees the growth estimate on  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  in the form (1.37).

None of this is to say that exact rogue wave solutions cannot be obtained from the IST summarized in Section 1.1 by a limiting process. Indeed, it is well-known that they can be found by first writing down the reflectionless multisoliton solutions in which the parameters  $\{\xi_1, \dots, \xi_N\}$  and  $\{\gamma_1(0), \dots, \gamma_N(0)\}$  are retained as variables, and then subsequently taking a suitable limit in which certain  $\xi_j$  tend to  $i$  (and  $\gamma_j$  is suitably scaled). However, the above discussion makes clear the fact that these solutions cannot be obtained directly. Moreover, the limiting techniques are useful for exact (reflectionless) solutions, but they are not as useful if  $R(\lambda; 0) \not\equiv 0$ .

## 1.2 Brief Description of the Robust IST

The purpose of this paper is to introduce a new version of the IST for the Cauchy problem (1.1)–(1.3) that is sufficiently robust to capture in an elementary way the missing information about the rate of growth of the Beals-Coifman matrix as  $\lambda \rightarrow \pm i$ . In fact, as the reader will see, it is never necessary to consider this rate of growth at all, as the information is encoded instead in the analytic columns of the Jost solution matrices *at a finite distance from these spectral singularities*.

To draw a comparison with the standard version of the IST for the focusing nonlinear Schrödinger equation with nonzero boundary conditions at infinity as described in Section 1.1, we summarize here the key steps in the robust IST we propose in this paper, full details of which can be found in Section 2 below.

**DIRECT TRANSFORM.** Assuming an initial condition  $\psi(x, 0) = \psi_0(x)$  with  $\Delta\psi_0(\cdot) := \psi_0(\cdot) - 1 \in L^1(\mathbb{R})$  and  $\Delta\psi_{0,x}(\cdot) \in L^1(\mathbb{R})$ , a positive radius  $r[\Delta\psi_0] > 1$  is defined by (B.25). The *scattering data for the robust IST* consists of the following:

- For  $\lambda \in \mathbb{R}$  with  $|\lambda| > r[\Delta\psi_0] > 1$  (a proper subset of  $\Gamma$  bounded away from the origin), the reflection coefficient  $R(\lambda; 0)$  is defined exactly as before, by (1.16). Then a jump matrix  $\mathbf{V}^{\mathbb{R}}(\lambda)$  with unit determinant is defined on the indicated intervals of  $\mathbb{R}$  by (1.30).
- For  $|\lambda| = r[\Delta\psi_0] > 1$  with  $\text{Im}\{\lambda\} > 0$ , the Jost solutions  $\mathbf{j}^{-,1}(\lambda; x, 0)$  and  $\mathbf{j}^{+,2}(\lambda; x, t)$  are calculated from the relevant columns of the Volterra integral equation (1.13) (the iterates of the relevant columns of (1.13) converge uniformly for the complex  $\lambda$  under consideration). Letting  $a(\lambda; 0)$  be defined for such  $\lambda$  by (1.15) (not zero by Lemma 1.2), a jump matrix  $\mathbf{V}^+(\lambda)$  with unit determinant is defined on the indicated semicircle by (2.13) below.

Note that the scattering data defined for  $|\lambda| > r[\Delta\psi_0]$  coincides with that for the traditional IST, but the jump matrix defined on the upper semicircle is a new ingredient. Also, there is no discrete scattering data of any sort in the robust IST, nor are any assumptions made about the number, type, or severity of singularities of  $\mathbf{M}^{\text{BC}}(\lambda; x, 0)$  when  $\lambda$  lies within the disk of radius  $r[\Delta\psi_0]$ .

**TIME DEPENDENCE.** When  $\psi(x, t)$  evolves under the focusing nonlinear Schrödinger equation (1.1), the scattering data collected from the initial data in the direct transform step evolve in a common and elementary way:  $\mathbf{V}^{\mathbb{R}}(\lambda)$  and  $\mathbf{V}^+(\lambda)$  are replaced with  $e^{-i\rho(\lambda)\lambda t\sigma_3}\mathbf{V}^{\mathbb{R}}(\lambda)e^{i\rho(\lambda)\lambda t\sigma_3}$  and  $e^{-i\rho(\lambda)\lambda t\sigma_3}\mathbf{V}^+(\lambda)e^{i\rho(\lambda)\lambda t\sigma_3}$ , respectively.

**INVERSE TRANSFORM.** In this step, we reconstruct a certain matrix  $\mathbf{M}(\lambda; x, t)$  from the time-evolved scattering data. In fact,  $\mathbf{M}(\lambda; x, t) = \mathbf{M}^{\text{BC}}(\lambda; x, t)$  for  $|\lambda| > r[\Delta\psi_0]$ , so  $\mathbf{M}(\lambda; x, t)$  is analytic in  $\lambda$  for  $|\lambda| > r[\Delta\psi_0]$  except on the real line, where its boundary values from the upper and lower half-planes are required to be related by the same jump condition as in the traditional IST, namely (1.29). The robust IST takes a different definition for  $\mathbf{M}(\lambda; x, t)$  when  $|\lambda| < r[\Delta\psi_0]$ , and the reader will see that  $\mathbf{M}(\lambda; x, t)$  is analytic for such  $\lambda$  except on the branch cut  $\Sigma_c$  where  $\lambda = i\text{Im}\{\lambda\}$  with  $|\text{Im}\{\lambda\}| \leq 1$ . Here it is required that  $\mathbf{M}(\lambda; x, t)$  take continuous boundary values from each side,<sup>4</sup> and that the boundary values are related by the (data independent) jump condition  $\mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t)e^{2i\rho_+(\lambda)(x+\lambda t)\sigma_3}$ .

Finally, it is necessary to specify that the boundary values taken by  $\mathbf{M}(\lambda; x, t)$  on the upper and lower semicircles of radius  $r[\Delta\psi_0]$  are continuous (they are analytic, in fact) and are related by the following jump conditions:

$$\mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t)e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}\mathbf{V}^+(\lambda)e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$$

holds for  $|\lambda| = r[\Delta\psi_0]$  and  $\text{Im}\{\lambda\} > 0$ , while

$$\mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t)e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}\mathbf{V}^+(\lambda^*)^\dagger e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$$

holds for  $|\lambda| = r[\Delta\psi_0]$  and  $\text{Im}\{\lambda\} < 0$ , where both arcs are oriented from left to right to define the boundary value subscripts. These conditions constitute a well-posed Riemann-Hilbert problem to determine  $\mathbf{M}(\lambda; x, t)$  from its time-evolved scattering data. The Riemann-Hilbert problem does not include any conditions to build in poles or specify implicitly or otherwise rates of growth of  $\mathbf{M}$  at spectral singularities because these things are instead encoded in the jump across the circle  $|\lambda| = r[\Delta\psi_0]$ , which is the novel ingredient in the robust IST. From the solution of the Riemann-Hilbert problem, the corresponding solution  $\psi(x, t)$  of the Cauchy problem (1.1)–(1.3) is then given by (2.26) below, which is really the same formula as in the traditional IST, because as remarked above,  $\mathbf{M}(\lambda; x, t) = \mathbf{M}^{\text{BC}}(\lambda; x, t)$  holds for  $|\lambda|$  sufficiently large.

---

<sup>4</sup> Throughout the paper, we adhere to the convention that subscripts  $+$  and  $-$  refer to the boundary value taken by a function on a given oriented contour arc from the left and the right, respectively.



The most important property of the above modification of the IST is, as we mean to suggest with its name, its robustness with respect to details of the initial data  $\psi_0(\cdot)$ . Indeed, the calculation of the scattering data requires no foreknowledge of the possible singularities of the Beals-Coifman matrix  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  except that they all must occur within a disk centered at the origin, the radius  $r[\Delta\psi_0]$  of which can be computed explicitly from  $\psi_0$  via (B.25). Similarly, the inverse problem avoids all reference to these singularities, and instead the Beals-Coifman matrix  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  is determined from how it connects to a certain interior eigenfunction across the circle of radius  $r[\Delta\psi_0]$ . These features allow the robust IST to capture all solutions of the Cauchy problem (1.1)–(1.3) obtainable from the traditional IST but without the need to be concerned about what kinds of singularities are generated within the disk  $|\lambda| < r[\Delta\psi_0]$  for the Beals-Coifman solutions. Moreover, the strength of the spectral singularities at  $\lambda = \pm i$  in the Beals-Coifman solutions need not be determined or specified; it is instead indirectly encoded in the jump across the circle through the Jost solutions. This allows the robust IST to capture solutions of the Cauchy problem (1.1)–(1.3) that cannot be obtained from the traditional IST, e.g., rogue waves.

### 1.3 Outline of the Paper

In Section 2 we formulate the robust IST for the Cauchy problem (1.1)–(1.3). Then in Section 3 we show how the robust IST makes the application of iterated Darboux transformations easier than in the traditional approach, and in particular how the robust IST builds in automatically the missing generalized eigenvectors that have to be extracted by differentiation with respect to  $\lambda$  in the generalized Darboux transformation method [23]. Thus rogue waves of arbitrary order can be calculated from the robust IST by using only standard Darboux methods. As a more analytical application, in Section 4 we begin to consider the question of dynamical stability/instability of rogue waves by linearizing the robust IST about a solution of the Cauchy problem (1.1)–(1.3). We present the solution of the linearized problem as a kind of spectral transform involving squared eigenfunctions, which allows us to draw some preliminary conclusions about the nature of instabilities and how they essentially arise from the background itself. Some of the more technical calculations are relegated to the appendices of the paper.

## 2 The Robust IST

As pointed out in Section 1.1, the cutoff approach of Deift and Zhou [14, 42] does not apply in the case of nonzero boundary conditions. However, we may take from it the following key ideas:

- Simultaneous solutions of the Lax pair (1.8)–(1.9) adapted to the boundary conditions at hand (the Beals-Coifman solutions  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$ ) are essential to consider for large  $\lambda$  in order to build in the correct asymptotic normalization condition for the Riemann-Hilbert problem of inverse scattering.

- On the other hand, other simultaneous solutions of (1.8)–(1.9) can be useful for bounded  $\lambda$ . In particular, solutions of *initial-value problems* for (1.8)–(1.9) have fantastic analytic properties with respect to  $\lambda$  although they fail to capture any information about boundary conditions at infinity.

To define what we mean by a solution of an initial-value problem for (1.8)–(1.9) and demonstrate its analytic features, we begin with the following proposition.

PROPOSITION 2.1. *Fix  $L \in \mathbb{R}$ . Suppose that  $\psi(x, t)$  is a bounded classical solution of the focusing NLS equation (1.1) defined for  $(x, t)$  in a simply connected domain  $\Omega \subseteq \mathbb{R}^2$  that contains the point  $(L, 0)$ . Then for each  $\lambda \in \mathbb{C}$  there exists a unique simultaneous fundamental solution matrix  $\mathbf{U} = \mathbf{U}^{\text{in}}(\lambda; x, t)$ ,  $(x, t) \in \Omega$ , of the Lax pair equations (1.8)–(1.9) together with the initial condition  $\mathbf{U}^{\text{in}}(\lambda; L, 0) = \mathbb{I}$ . Moreover,  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  is an entire function of  $\lambda$  for each  $(x, t) \in \Omega$ , and  $\det(\mathbf{U}^{\text{in}}(\lambda; x, t)) \equiv 1$ .*

PROOF. The proof is based on a standard Picard iteration argument for the system of simultaneous linear differential equations (1.8)–(1.9) in  $\Omega \subset \mathbb{R}^2$  (see, for example, [12, chap. 1] for a more detailed construction). Let  $(x, t) \in \Omega$  be fixed. Because  $\psi(x, t)$  is a solution of the focusing NLS equation (1.1) in  $\Omega$ , the linear differential equations (1.8)–(1.9) are compatible; i.e., the coefficient matrices  $\mathbf{X}(\lambda; x, t)$  and  $\mathbf{T}(\lambda; x, t)$  satisfy the zero-curvature condition

$$(2.1) \quad \mathbf{X}_t - \mathbf{T}_x + [\mathbf{X}, \mathbf{T}] = 0$$

and therefore since  $\Omega$  is simply connected, we can set up a Picard iteration to integrate the simultaneous equations (1.8)–(1.9) along an arbitrary smooth path  $\Pi$  in  $\Omega$  from  $(L, 0)$  to  $(x, t)$  and be guaranteed that the result of the iteration will be independent of the choice of path. Let the path  $\Pi$  be parametrized by  $x = X(u)$ ,  $t = T(u)$ ,  $0 \leq u \leq 1$ , such that  $(X(0), T(0)) = (L, 0)$  and  $(X(1), T(1)) = (x, t)$ . The integral equation governing the restriction  $\mathbf{W}^\Pi(\lambda; u) := \mathbf{U}^{\text{in}}(\lambda; X(u), T(u))$  of  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  to the path  $\Pi$  is then

$$(2.2) \quad \begin{aligned} & \mathbf{W}^\Pi(\lambda; u) \\ &= \mathbb{I} + \int_0^u (\mathbf{X}(\lambda; X(v), T(v))X'(v) \\ & \quad + \mathbf{T}(\lambda; X(v), T(v))T'(v))\mathbf{W}^\Pi(\lambda; v)dv, \quad 0 < u \leq 1. \end{aligned}$$

By assumption on the classical nature of the solution  $\psi(x, t)$  of (1.1), the expression in parentheses in the integrand is continuous in  $v$ , and this guarantees uniform convergence of the Picard iterates (possibly first chopping the interval  $[0, 1]$  into finitely many subintervals and restarting the iteration after each). The expression in parentheses in the integrand is also a (quadratic) polynomial in  $\lambda$ , so by an argument based on Morera’s theorem and Fubini’s theorem, the iterates are all entire functions. As the iterates converge uniformly on compact subsets of  $\lambda \in \mathbb{C}$ , the unique solution of the initial-value problem is entire in  $\lambda$ . By compatibility,

$\mathbf{U}^{\text{in}}(\lambda; x, t) := \mathbf{W}^{\Pi}(\lambda; 1)$  is well-defined regardless of the choice of path  $\Pi$ . The fact that  $\det(\mathbf{U}^{\text{in}}(\lambda; x, t)) = 1$  then follows from  $\det(\mathbf{U}^{\text{in}}(L, 0)) = \det(\mathbb{I}) = 1$  using Abel’s theorem because  $\text{tr}(\mathbf{X}(\lambda; x, t)) = \text{tr}(\mathbf{T}(\lambda; x, t)) = 0$  holds for  $\lambda \in \mathbb{C}$  and  $(x, t) \in \Omega$ . □

Note that the Schwarz symmetry of the coefficients  $\mathbf{X}(\lambda; x, t)$  and  $\mathbf{T}(\lambda; x, t)$  (see (1.17) for  $\mathbf{X}$ ; the analogous result also holds for  $\mathbf{T}$ ) of the simultaneous differential equations (1.8) and (1.9) implies that  $\sigma_2 \mathbf{U}^{\text{in}}(\lambda^*; x, t)^* \sigma_2$  is also an entire simultaneous solution matrix for the system (1.8)–(1.9). Since this solution matrix coincides with  $\mathbf{U}^{\text{in}}(\lambda; L, 0) = \mathbb{I}$  for  $(x, t) = (L, 0)$ , by uniqueness we conclude that

$$(2.3) \quad \mathbf{U}^{\text{in}}(\lambda; x, t) = \sigma_2 \mathbf{U}^{\text{in}}(\lambda^*; x, t)^* \sigma_2.$$

*Remark 2.2.*  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  can also be obtained from *any* matrix  $\mathbf{U}(\lambda; x, t)$  of simultaneous solutions of (1.8)–(1.9) defined and fundamental for  $\lambda \in \mathbb{C}$  with the possible exception of isolated points or even curves and defining for some chosen  $L \in \mathbb{R}$ :

$$(2.4) \quad \mathbf{U}^{\text{in}}(\lambda; x, t) := \mathbf{U}(\lambda; x, t) \mathbf{U}(\lambda; L, 0)^{-1}.$$

Regardless of whether  $\mathbf{U}(\lambda; x, t)$  is analytic anywhere, by uniqueness of the solution of the initial-value problem solved in the proof of Proposition 2.1 this formula produces the desired object for all  $\lambda$  with the exception of removable singularities at the aforementioned isolated points or curves.

Recall from Lemma 1.2 that if  $\Delta\psi_0$  and  $\Delta\psi'_0$  are absolutely integrable on  $\mathbb{R}$ , then  $|a(\lambda)| > \frac{1}{2}$  holds for  $|\lambda| \geq r[\Delta\psi] > 1$  and  $\text{Im}\{\lambda\} \geq 0$ . Let  $\Sigma_0$  denote the circle of radius  $r = r[\Delta\psi]$  centered at the origin in the  $\lambda$ -plane, and let  $D_0$  denote the open disk whose boundary is  $\Sigma_0$ . Since  $r > 1$ , the branch cut  $\Sigma_c$  is contained in  $D_0$ . We also define the following related domains:

$$(2.5) \quad D_{\pm} := \{\lambda \in \mathbb{C} : |\lambda| > r[\Delta\psi] \text{ and } \text{Im}\{\lambda\} \gtrless 0\},$$

and the following contours, which divide the  $\lambda$ -plane into disjoint regions  $D_0, D_+$ , and  $D_-$ :

$$(2.6) \quad \begin{aligned} \Sigma_+ &:= \{\lambda \in \Sigma_0 : \text{Im}\{\lambda\} \geq 0\}, \\ \Sigma_- &:= \{\lambda \in \Sigma_0 : \text{Im}\{\lambda\} \leq 0\}, \\ \Sigma_L &:= \{\lambda \in \mathbb{R} : \lambda \in (-\infty, -r[\Delta\psi]]\}, \\ \Sigma_R &:= \{\lambda \in \mathbb{R} : \lambda \in [r[\Delta\psi], +\infty)\}. \end{aligned}$$

Note that  $\Sigma_0 = \Sigma_+ \cup \Sigma_-$ . We orient both semicircular arcs  $\Sigma_{\pm}$  in the direction from  $\lambda = -r[\Delta\psi]$  to  $\lambda = r[\Delta\psi]$ ,  $\Sigma_L$  from  $\lambda = -\infty$  to  $\lambda = -r[\Delta\psi]$ , and  $\Sigma_R$  from  $\lambda = r[\Delta\psi]$  to  $\lambda = +\infty$ . See Figure 2.1. We denote by  $\Sigma$  the union of all of these contours.

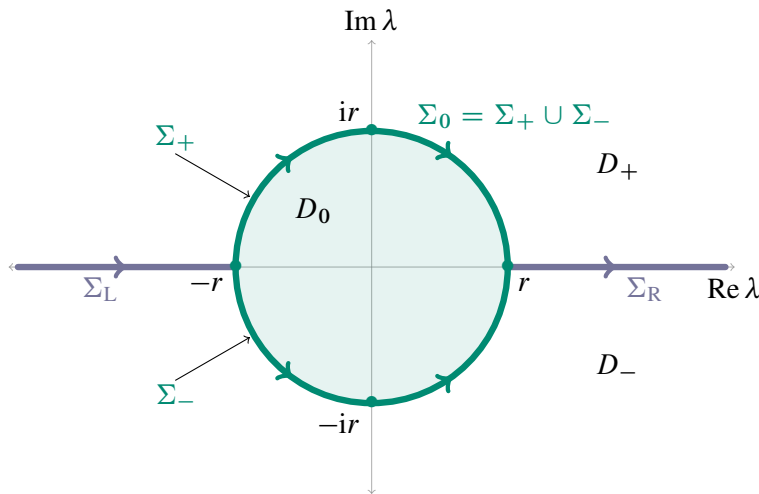


FIGURE 2.1. Definitions of the regions  $D_0$ ,  $D_{\pm}$ , and the contours  $\Sigma_0$ ,  $\Sigma_L$ , and  $\Sigma_R$ .

We now define for  $\lambda \in \mathbb{C} \setminus \Sigma$  a simultaneous fundamental solution matrix for the Lax pair (1.8)–(1.9) as follows:

$$(2.7) \quad \mathbf{U}(\lambda; x, t) := \begin{cases} \mathbf{U}^{\text{BC}}(\lambda; x, t), & \lambda \in D_+ \cup D_-, \\ \mathbf{U}^{\text{in}}(\lambda; x, t), & \lambda \in D_0, \end{cases}$$

where  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  is defined by (1.24) and  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  is defined in Proposition 2.1.

Assuming only that the initial condition  $\psi_0$  in (1.3) satisfies  $\Delta\psi_0 := \psi_0 - 1 \in L^1(\mathbb{R})$  and  $\Delta\psi'_0 \in L^1(\mathbb{R})$  and generates a classical solution  $\psi(x, t)$  of the Cauchy problem (1.1)–(1.3), Lemma 1.2 and the formula (1.24) show that  $\mathbf{U}(\lambda; x, t)$  is analytic in  $D_+$  and  $D_-$  and takes continuous boundary values from these domains on  $\Sigma$ . Similarly, Proposition 2.1 guarantees that  $\mathbf{U}(\lambda; x, t)$  is analytic in  $D_0$  and takes continuous boundary values on  $\Sigma_0 = \partial D_0$ . Furthermore,  $\det(\mathbf{U}(\lambda; x, t)) = 1$  holds for all  $\lambda \in \mathbb{C} \setminus \Sigma$ . Combining (1.26)–(1.28) and (2.3), we see that

$$(2.8) \quad \mathbf{U}(\lambda^*; x, t) = \sigma_2 \mathbf{U}(\lambda; x, t)^* \sigma_2, \quad \lambda \in \mathbb{C} \setminus \Sigma, (x, t) \in \mathbb{R}^2.$$

Now define the related matrix  $\mathbf{M}(\lambda; x, t)$  by

$$(2.9) \quad \mathbf{M}(\lambda; x, t) := \mathbf{U}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \mathbb{C} \setminus (\Sigma \cup \Sigma_c), (x, t) \in \mathbb{R}^2,$$

where  $\mathbf{U}(\lambda; x, t)$  is defined by (2.7). The exponential factor  $e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$  introduces a new jump discontinuity across  $\Sigma_c$ , but otherwise  $\mathbf{M}(\lambda; x, t)$  is analytic where it is defined. Furthermore, from (1.26) and (1.27) we see that for each fixed  $(x, t) \in \mathbb{R}^2$ ,  $\mathbf{M}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$  in  $D_{\pm}$ , uniformly with respect to direction. It is easy to see that  $\mathbf{M}(\lambda; x, t)$  takes continuous boundary values on  $\Sigma \cup \Sigma_c$ ,

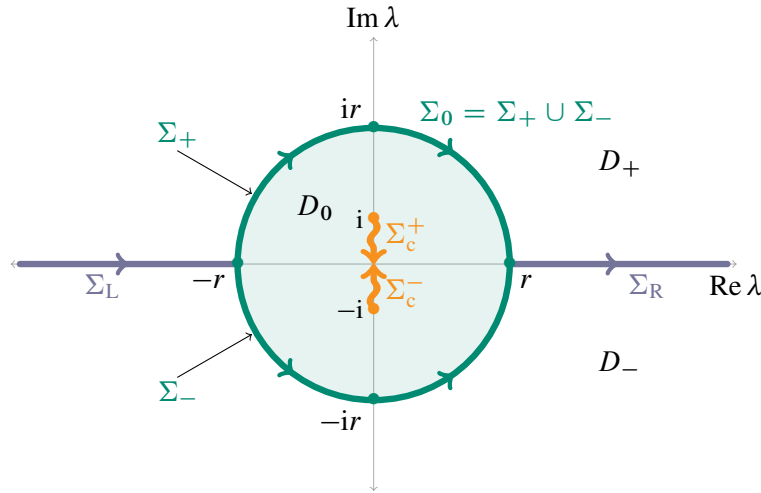


FIGURE 2.2. Definitions of the regions  $D_0$ ,  $D_{\pm}$ , and the contours  $\Sigma_0$ ,  $\Sigma_L$ , and  $\Sigma_R$ .  $\Sigma_c = \Sigma_c^+ \cup \Sigma_c^-$  is the branch cut of the function  $\rho(\lambda)$ .

including at the endpoints  $\lambda = \pm i$  of  $\Sigma_c$  (since  $\rho(\pm i) = 0$  unambiguously and  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  is analytic at  $\lambda = \pm i$ ).

We now compute the jump conditions relating these boundary values. It will be convenient to oppositely orient the parts of  $\Sigma_c$  in the upper and lower half-planes, so we define

$$(2.10) \quad \Sigma_c^+ := \{\lambda \in \Sigma_c: \text{Im}\{\lambda\} \geq 0\}, \quad \Sigma_c^- := \{\lambda \in \Sigma_c: \text{Im}\{\lambda\} \leq 0\},$$

and orient both contours  $\Sigma_c^{\pm}$  in the direction toward the origin  $\lambda = 0$ . This makes  $\Sigma \cup \Sigma_c$  a Schwarz-symmetric contour (taking orientation into account). See Figure 2.2. We adopt the usual convention that  $\mathbf{M}_+(\lambda; x, t)$  (respectively  $\mathbf{M}_-(\lambda; x, t)$ ) denotes the boundary value taken at a regular (non-self-intersecting) point of  $\Sigma \cup \Sigma_c$  from the left (respectively right) side by orientation.

PROPOSITION 2.3. *The continuous boundary values taken by  $\mathbf{M}(\lambda; x, t)$  satisfy the jump conditions*

$$(2.11) \quad \mathbf{M}_+(\lambda; x, t) = \begin{cases} \mathbf{M}_-(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}, & \lambda \in \Sigma, \\ \mathbf{M}_-(\lambda; x, t) e^{2i\rho_+(\lambda)(x+\lambda t)\sigma_3}, & \lambda \in \Sigma_c, \end{cases}$$

with

$$(2.12) \quad \mathbf{V}(\lambda) := \begin{cases} \mathbf{V}^+(\lambda), & \lambda \in \Sigma_+, \\ \mathbf{V}^-(\lambda), & \lambda \in \Sigma_-, \\ \mathbf{V}^{\mathbb{R}}(\lambda), & \lambda \in \Sigma_L \cup \Sigma_R, \end{cases}$$

where

$$(2.13) \quad \mathbf{V}^+(\lambda) := [a(\lambda)^{-1} \mathbf{j}^{-,1}(\lambda; L, 0); \mathbf{j}^{+,2}(\lambda; L, 0)],$$

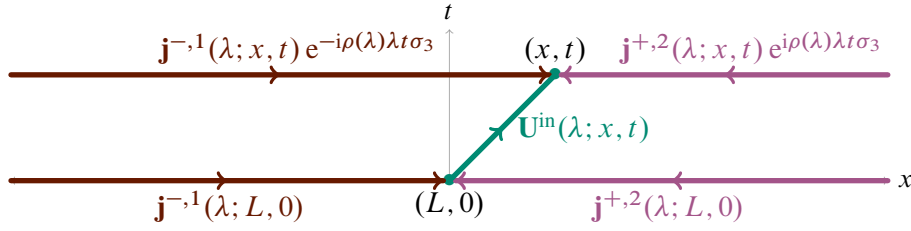


FIGURE 2.3. The computation of the jump condition for  $\lambda \in \Sigma_+$ . Thus  $\mathbf{j}^{-,1}(\lambda; x, t) e^{-i\rho(\lambda)\lambda t\sigma_3} = \mathbf{U}^{\text{in}}(\lambda; x, t)\mathbf{j}^{-,1}(\lambda; L, 0)$  and  $\mathbf{j}^{+,2}(\lambda; x, t) e^{i\rho(\lambda)\lambda t\sigma_3} = \mathbf{U}^{\text{in}}(\lambda; x, t)\mathbf{j}^{+,2}(\lambda; L, 0)$ .

$$(2.14) \quad \mathbf{V}^-(\lambda) := [\mathbf{j}^{+,1}(\lambda; L, 0); \bar{a}(\lambda)^{-1}\mathbf{j}^{-,2}(\lambda; L, 0)]^{-1},$$

and  $\mathbf{V}^{\mathbb{R}}(\lambda)$  is defined by (1.30) in terms of the reflection coefficient associated with the initial data  $\psi_0$ .

The main part of the proof is to establish the jump conditions across  $\Sigma_{\pm}$ , and for this purpose it is useful to think of  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  as a “transfer matrix” as illustrated in Figure 2.3. The sort of reasoning illustrated in this figure is similar to that used to study mixed initial-boundary value problems using the unified transform method; see [19, 20].

PROOF. Since  $\mathbf{M}(\lambda; x, t) = \mathbf{M}^{\text{BC}}(\lambda; x, t)$  for  $\lambda \in D_+ \cup D_-$ , where the Beals-Coifman matrix function is defined in (1.26), the jump condition on  $\Sigma_L \cup \Sigma_R \subset \mathbb{R}$  is just a special case of the formula (1.29)–(1.30).

Next, fix  $\lambda \in \Sigma_+$ . Observe that since  $\mathbf{j}^{-,1}(\lambda; x, t) e^{-i\rho(\lambda)\lambda t}$  is a simultaneous solution of the Lax pair (1.8)–(1.9) and since  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  is a fundamental solution matrix for the same system, there exists a vector  $\mathbf{d}^{+,1}(\lambda)$  (independent of  $(x, t) \in \mathbb{R}^2$ ) such that the first column of  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  satisfies

$$(2.15) \quad a(\lambda)^{-1}\mathbf{j}^{-,1}(\lambda; x, t) e^{-i\rho(\lambda)\lambda t} = \mathbf{U}^{\text{in}}(\lambda; x, t)\mathbf{d}^{+,1}(\lambda), \quad \lambda \in \Sigma_+.$$

Taking  $(x, t) = (L, 0)$ , the identity  $\mathbf{U}^{\text{in}}(\lambda; L, 0) = \mathbb{I}$  then determines

$$\mathbf{d}^{+,1}(\lambda) = a(\lambda)^{-1}\mathbf{j}^{-,1}(\lambda; L, 0).$$

Similarly, since  $\mathbf{j}^{+,2}(\lambda; x, t) e^{i\rho(\lambda)\lambda t}$  is a simultaneous solution of the system (1.8)–(1.9), there exists a vector  $\mathbf{d}^{+,2}(\lambda)$  such that the second column of  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  satisfies

$$(2.16) \quad \mathbf{j}^{+,2}(\lambda; x, t) e^{i\rho(\lambda)\lambda t} = \mathbf{U}^{\text{in}}(\lambda; x, t)\mathbf{d}^{+,2}(\lambda), \quad \lambda \in \Sigma_+,$$

and again taking  $(x, t) = (L, 0)$  we get  $\mathbf{d}^{+,2}(\lambda) = \mathbf{j}^{+,2}(\lambda; L, 0)$ . Then (2.15) and (2.16) imply that for  $\lambda \in \Sigma_+$ , by the clockwise orientation of  $\Sigma_+$ , we have

$$(2.17) \quad \begin{aligned} \mathbf{U}_+(\lambda; x, t) &= \mathbf{U}_-(\lambda; x, t)[\mathbf{d}^{+,1}(\lambda); \mathbf{d}^{+,2}(\lambda)] \\ &= \mathbf{U}_-(\lambda; x, t)[a(\lambda)^{-1}\mathbf{j}^{-,1}(\lambda; L, 0); \mathbf{j}^{+,2}(\lambda; L, 0)], \end{aligned}$$

because  $\mathbf{U}_+(\lambda; x, t) = \mathbf{U}^{\text{in}}(\lambda; x, t)$  and  $\mathbf{U}_-(\lambda; x, t) = \mathbf{U}^{\text{BC}}(\lambda; x, t)$  for  $\lambda \in \Sigma_+$ . Because  $\rho(\lambda)$  is analytic on  $\Sigma_0$ , this proves that the jump condition satisfied by  $\mathbf{M}(\lambda; x, t)$  on  $\Sigma_+$  is

$$(2.18) \quad \mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^+(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3},$$

$$\lambda \in \Sigma_+,$$

where  $\mathbf{V}^+(\lambda)$  is defined by (2.13). A completely analogous calculation shows that

$$(2.19) \quad \mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^-(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3},$$

$$\lambda \in \Sigma_-,$$

where  $\mathbf{V}^-(\lambda)$  is defined by (2.14).

Finally, to prove the jump condition of  $\mathbf{M}(\lambda; x, t)$  across  $\Sigma_c = \Sigma_c^+ \cup \Sigma_c^-$ , observe that  $\Sigma_c$  is contained in  $D_0$ , where  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  is single-valued and analytic, and that  $\rho_+(\lambda) = -\rho_-(\lambda)$  for  $\lambda \in \Sigma_c$ . Thus,

$$(2.20) \quad \begin{aligned} \mathbf{M}_+(\lambda; x, t) &= \mathbf{U}^{\text{in}}(\lambda; x, t) e^{i\rho_+(\lambda)(x+\lambda t)\sigma_3} \\ &= \mathbf{U}^{\text{in}}(\lambda; x, t) e^{i\rho_-(\lambda)(x+\lambda t)\sigma_3} e^{i(\rho_+(\lambda)-\rho_-(\lambda))(x+\lambda t)\sigma_3} \\ &= \mathbf{M}_-(\lambda; x, t) e^{2i\rho_+(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \Sigma_c^\pm, \end{aligned}$$

and this completes the proof. □

Note that  $a(\lambda)$  and  $\bar{a}(\lambda)$  appearing in the jump matrices (2.13) and (2.14), respectively, are defined by (1.15), and as the relevant Wronskian determinants are independent of the variables  $(x, t)$ , we may choose  $(x, t) = (L, 0)$ :

$$(2.21) \quad \begin{aligned} a(\lambda) &= \det([\mathbf{j}^{-,1}(\lambda; L, 0); \mathbf{j}^{+,2}(\lambda; L, 0)]), \\ \bar{a}(\lambda) &= \det([\mathbf{j}^{+,1}(\lambda; L, 0); \mathbf{j}^{+,2}(\lambda; L, 0)]), \end{aligned}$$

and thus it is obvious that  $\det(\mathbf{V}^\pm(\lambda)) = 1$ .

One can now see that the matrix  $\mathbf{M}(\lambda; x, t)$  defined in (2.9) satisfies the conditions of the following Riemann-Hilbert problem, the data for which only involves the initial condition  $\psi_0(x)$ . Even more significantly, *we require no condition on  $\psi_0$  to guarantee any particular rate of growth of  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$  as  $\lambda \rightarrow \pm i$ .* Indeed, there is no need for any growth condition like (1.37) to make the following problem well-posed.

**Riemann-Hilbert Problem 1.** Seek a  $2 \times 2$  matrix function  $\mathbf{M}(\lambda; x, t)$  that has the following properties:

- **ANALYTICITY:**  $\mathbf{M}(\lambda; x, t)$  is analytic for  $\lambda \in D_+ \cup D_-$  and for  $\lambda \in D_0 \setminus \Sigma_c$ .
- **JUMP CONDITION:**  $\mathbf{M}(\lambda; x, t)$  takes continuous boundary values  $\mathbf{M}_\pm(\lambda; x, t)$  on  $\Sigma \cup \Sigma_c$ , and they are related by the jump conditions described in Proposition 2.3.
- **NORMALIZATION:**  $\lim_{\lambda \rightarrow \infty} \mathbf{M}(\lambda; x, t) = \mathbb{I}$ .

This Riemann-Hilbert problem has many of the features that are expected and desirable in the setting of inverse scattering. Indeed, the “core” jump matrix  $\mathbf{V}(\lambda)$  in (2.12) depends solely on scattering data (well-defined Jost solutions) associated with the initial data  $\psi_0$  for the Cauchy initial-value problem (1.1)–(1.3), and all  $(x, t)$ -dependence in the problem enters via conjugation of the core jump matrix by elementary diagonal exponential factors. Furthermore, we can show that this problem has a unique solution for all  $(x, t) \in \mathbb{R}^2$ .

**THEOREM 2.4.** *Riemann-Hilbert Problem 1 has a unique solution for all  $(x, t) \in \mathbb{R}^2$ .*

**PROOF.** We will show that the jump conditions and the jump matrices in Riemann-Hilbert Problem 1 satisfy the hypotheses of Zhou’s *vanishing lemma* [43, theorem 9.3]. Note that the jump contour  $\Sigma \cup \Sigma_c$  has the necessary invariance under Schwarz reflection with orientation. We first show that for all  $\lambda \in (\Sigma \cup \Sigma_c) \setminus \mathbb{R}$ , the jump matrix  $\mathbf{V}(\lambda; x, t) := \mathbf{M}_-(\lambda; x, t)^{-1} \mathbf{M}_+(\lambda; x, t)$  satisfies  $\mathbf{V}(\lambda^*; x, t) = \mathbf{V}(\lambda; x, t)^\dagger$ . Note that  $\mathbf{V}^+(\lambda) = \mathbf{U}^{\text{BC}}(\lambda; L, 0)$  for  $\lambda \in \Sigma_+$  while  $\mathbf{V}^-(\lambda) = \mathbf{U}^{\text{BC}}(\lambda; L, 0)^{-1}$  for  $\lambda \in \Sigma_-$ . Suppose that  $\lambda \in \Sigma_+$ . Since the Schwarz symmetry (1.28) implies the identity  $\mathbf{U}^{\text{BC}}(\lambda^*; x, t) = \sigma_2 \mathbf{U}^{\text{BC}}(\lambda; x, t)^* \sigma_2$ , which can be rewritten as  $\mathbf{U}^{\text{BC}}(\lambda^*; x, t)^\dagger = \mathbf{U}^{\text{BC}}(\lambda; x, t)^{-1}$  because  $\det(\mathbf{U}^{\text{BC}}(\lambda; x, t)) = 1$ ,

$$\begin{aligned}
 \mathbf{V}(\lambda^*; x, t) &= e^{-i\rho(\lambda^*)(x+\lambda^*t)\sigma_3} \mathbf{V}^-(\lambda^*) e^{i\rho(\lambda^*)(x+\lambda^*t)\sigma_3} \\
 &= e^{-i\rho(\lambda)^*(x+\lambda t)^*\sigma_3} \mathbf{U}^{\text{BC}}(\lambda^*; L, 0)^{-1} e^{i\rho(\lambda)^*(x+\lambda t)^*\sigma_3} \\
 (2.22) \quad &= (e^{i\rho(\lambda)(x+\lambda t)\sigma_3})^\dagger \mathbf{U}^{\text{BC}}(\lambda; L, 0)^\dagger (e^{-i\rho(\lambda)(x+\lambda t)\sigma_3})^\dagger \\
 &= (e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^+(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3})^\dagger = \mathbf{V}(\lambda; x, t)^\dagger.
 \end{aligned}$$

Supposing next that  $\lambda \in \Sigma_c^+$ , we have

$$\begin{aligned}
 \mathbf{V}(\lambda^*; x, t) &= e^{2i\rho_+(\lambda^*)(x+\lambda^*t)\sigma_3} \\
 (2.23) \quad &= e^{2i\rho_-(\lambda)^*(x+\lambda^*t)\sigma_3} = (e^{2i\rho_+(\lambda)(x+\lambda t)\sigma_3})^\dagger = \mathbf{V}_c(\lambda; x, t)^\dagger.
 \end{aligned}$$

It remains to show that if  $\lambda \in \Sigma_L \cup \Sigma_R$  (i.e., the part of the jump contour on the real axis), then  $\mathbf{V}(\lambda; x, t) + \mathbf{V}(\lambda; x, t)^\dagger$  is positive definite. For such  $\lambda$ ,  $\rho(\lambda)$  is real-valued and  $\mathbf{V}^{\mathbb{R}}(\lambda)$  defined by (1.30) is Hermitian, so we have

$$\begin{aligned}
 \mathbf{V}(\lambda; x, t) + \mathbf{V}(\lambda; x, t)^\dagger &= 2e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}^{\mathbb{R}}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3} \\
 (2.24) \quad &= \mathbf{H}(\lambda; x, t)^\dagger \mathbf{H}(\lambda; x, t),
 \end{aligned}$$

where

$$(2.25) \quad \mathbf{H}(\lambda; x, t) := \sqrt{2} \begin{bmatrix} 1 & 0 \\ R(\lambda) e^{i\rho(\lambda)(x+\lambda t)} & 1 \end{bmatrix}$$

is an invertible matrix ( $\det(\mathbf{H}(\lambda)) = 2$ ). Since every matrix of the form  $\mathbf{H}^\dagger \mathbf{H}$  with  $\mathbf{H}$  invertible is positive definite, the desired conclusion follows. With the normalization condition  $\mathbf{M}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$ , we have confirmed all the



hypotheses of Zhou's vanishing lemma. Consequently, Riemann-Hilbert Problem 1 is uniquely solvable.  $\square$

From the solution  $\mathbf{M}(\lambda; x, t)$  of Riemann-Hilbert Problem 1 we recover the solution  $\psi(x, t)$  of the Cauchy problem (1.1)–(1.3) via the limit

$$(2.26) \quad \psi(x, t) = 2i \lim_{\lambda \rightarrow \infty} \lambda M_{12}(\lambda; x, t) = 2i \lim_{\lambda \rightarrow \infty} \lambda M_{12}^{\text{BC}}(\lambda; x, t).$$

A reasonable question to ask is: what happened to all of the poles and spectral singularities that can be present for reasonable initial data in the traditional IST outlined in Section 1.1? The answer is that this information is now encoded instead in the jump matrices  $\mathbf{V}^{\pm}(\lambda)$  supported on the circle  $\Sigma_0$ . These jump matrices are locally analytic on  $\Sigma_0$  and upon carrying out analytic continuation into  $D_0$  all of these singularities will re-emerge. We wish to emphasize that in order to compute the jump matrices  $\mathbf{V}^{\pm}(\lambda)$  no analytic continuation away from the continuous spectrum  $\Gamma$  is necessary, since the Volterra integral equations characterizing the Jost solutions appearing in (2.13)–(2.14) can be solved by iteration for fixed  $\lambda$  on the corresponding jump contours. Finally, for the purposes of the robust IST it is sufficient to assume that  $\Delta\psi := \psi - 1$  and  $\Delta\psi'$  lie in  $L^1(\mathbb{R})$ ; there is no need for any stronger decay condition such as  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$ . Note that the Peregrine solution  $\psi_{\text{P}}$  given by (1.5) and its higher-order generalizations satisfy  $\Delta\psi, \Delta\psi' \in L^1(\mathbb{R})$ , but *not*  $(1 + x^2)\Delta\psi(x) \in L^1(\mathbb{R})$ .

*Remark 2.5.* The proof of Theorem 2.4 relies on only basic properties of the jump matrix  $\mathbf{V}(\lambda; x, t) = e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$  defined for  $\lambda \in \Sigma$ , namely that  $\mathbf{V}(\lambda^*) = \mathbf{V}(\lambda)^{\dagger}$  holds for  $\lambda \in \Sigma_0 \subset \Sigma$  and that  $\mathbf{V}(\lambda) + \mathbf{V}(\lambda)^{\dagger}$  is positive definite for  $\lambda \in \Sigma_L \cup \Sigma_R \subset \Sigma$ . Hence, whenever  $\mathbf{V}(\lambda)$  has such properties (in addition to standard technical properties characterizing a function space for  $\mathbf{V}$  and the sense in which  $\mathbf{V} \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$  from within  $\Sigma_L \cup \Sigma_R \subset \Sigma$ ), the conclusion of Theorem 2.4 holds. Moreover, a standard dressing argument based on the fact that the matrix  $\mathbf{U}(\lambda; x, t) = \mathbf{M}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$  satisfies Riemann-Hilbert conditions independent of  $(x, t) \in \mathbb{R}^2$  shows that the function  $\psi(x, t)$  extracted from the unique solution of Riemann-Hilbert Problem 1 via (2.26) is a global classical solution of the focusing NLS equation (1.1). Significantly, this conclusion holds even in the generalized setting in which  $\mathbf{V}$  need not originate via the (robust) direct scattering transform for any Cauchy data (1.3).

*Remark 2.6.* Note that  $L \in \mathbb{R}$  appears as an auxiliary parameter in the construction of  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  by Proposition 2.1 (and hence in  $\mathbf{M}(\lambda; x, t)$ ). The parameter  $L$  then appears in the core jump matrices  $\mathbf{V}^{\pm}(\lambda)$  on  $\Sigma^{\pm}$  in Riemann-Hilbert Problem 1. Now suppose that  $\mathbf{U}^{\text{in},j}(\lambda; x, t)$  is the matrix of Proposition 2.1 normalized at  $(x, t) = (L_j, 0)$  for  $j = 1, 2$ . Then

$$(2.27) \quad \mathbf{U}^{\text{in},2}(\lambda; x, t) = \mathbf{U}^{\text{in},1}(\lambda; x, t) \mathbf{U}^{\text{in},1}(\lambda; L_2, 0)^{-1}$$

since  $\mathbf{U}^{\text{in},1}(\lambda; x, t)$  is a matrix of fundamental simultaneous solutions of (1.8)–(1.9). Thus, changing the value of  $L$  manifests itself through the multiplication

of  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  on the right by a matrix that is constant in  $(x, t)$  and analytic in  $\lambda$ . This constant matrix can also be appropriately applied to the jump condition on the circle  $\Sigma_0$ , leaving  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  unchanged outside  $\Sigma_0$ . Doing so, for different values of  $L$ , we leave the residue of  $\mathbf{M}(\lambda; x, t)$  at  $\lambda = \infty$  (and hence the solution of the NLS equation recovered via (2.26)) unchanged. As in [14, 42], the IST method we have here provides a correspondence between a suitable solution of the NLS equation (1.1) and an equivalence class of Riemann-Hilbert problems, or an equivalence class of jump matrices supported on the circle  $\Sigma_0$  augmented by the jump matrix on the real line.

*Remark 2.7.* The basic method described here for the case of nonzero boundary conditions (1.2) can also be implemented in the case of zero boundary conditions (1.6) simply by replacing  $\rho(\lambda)$  with  $\lambda$ , and hence  $\mathbf{E}(\lambda)$  by the identity matrix  $\mathbb{I}$ . Furthermore, the branch cut  $\Sigma_c$  is not present in the latter case. Thus the robust IST provides an alternative approach in that case to the cutoff method advanced in [14, 42], and the robust IST avoids the jump on the real axis in  $D_0$  that is generally present in the cutoff method. The basic approach used to develop the robust IST described in this section can also be applied to other integrable problems like the derivative nonlinear Schrödinger equation with minimal modifications.

### 3 Darboux Transformations in the Context of the Robust IST

In this section, we develop a Darboux transformation scheme for Riemann-Hilbert Problem 1 and apply it to generate arbitrary-order rogue waves in such a way that they directly admit a Riemann-Hilbert representation with no limit process required. A good general reference for Darboux transformations in integrable systems is the book of Matveev and Salle [33].

#### 3.1 Basic Pole Insertion

Let  $\mathbf{M}(\lambda; x, t)$  be the matrix function characterized by the conditions of Riemann-Hilbert Problem 1, and let  $\mathbf{U}(\lambda; x, t) := \mathbf{M}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$ . Choose any point  $\xi \in D_0$ . For a 2-nilpotent matrix  $\mathbf{R}(x, t)$  to be determined, consider the gauge transformation

$$(3.1) \quad \dot{\mathbf{U}}(\lambda; x, t) := \left( \mathbb{I} + \frac{\mathbf{R}(x, t)}{\lambda - \xi} \right) \mathbf{U}(\lambda; x, t), \quad \mathbf{R}(x, t)^2 = \mathbf{0}.$$

This gauge transformation has the following effects:

- Because  $\mathbf{R}(x, t)$  is 2-nilpotent, the identity  $\det(\mathbf{U}(\lambda; x, t)) = 1$  implies that also  $\det(\dot{\mathbf{U}}(\lambda; x, t)) = 1$ .
- The normalization condition  $\mathbf{M}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$  implies that also  $\dot{\mathbf{M}}(\lambda; x, t) := \dot{\mathbf{U}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3} \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$ .
- At each non-self-intersection point of  $\Sigma$ , the jump condition  $\mathbf{U}_+(\lambda; x, t) = \mathbf{U}_-(\lambda; x, t)\mathbf{V}(\lambda)$  implies that also  $\dot{\mathbf{U}}_+(\lambda; x, t) = \dot{\mathbf{U}}_-(\lambda; x, t)\mathbf{V}(\lambda)$ .

For the gauge transformation to be nontrivial, although  $\mathbf{R}(x, t)^2 = \mathbf{0}$ , we assume  $\mathbf{R}(x, t) \neq \mathbf{0}$ , in which case  $\mathbf{R}(x, t)$  necessarily has the form

$$(3.2) \quad \mathbf{R}(x, t) = \mathbf{v}(x, t)\mathbf{v}(x, t)^\top \sigma_2$$

for a vector-valued function  $\mathbf{v}(x, t)$  not identically 0. The vector  $\mathbf{v}(x, t)$  is then to be determined so that

$$(3.3) \quad \text{Res}_{\lambda=\xi} \dot{\mathbf{U}}(\lambda; x, t) = \lim_{\lambda \rightarrow \xi} \dot{\mathbf{U}}(\lambda; x, t)\mathbf{C},$$

where  $\mathbf{C} \neq \mathbf{0}$  is a  $2 \times 2$  complex-valued constant 2-nilpotent matrix:  $\mathbf{C}^2 = \mathbf{0}$ . By analogy with (3.2), we may write  $\mathbf{C}$  in the form

$$(3.4) \quad \mathbf{C} = \mathbf{c}\mathbf{c}^\top \sigma_2 \quad \text{where } \mathbf{c} = [c_1 \quad c_2]^\top.$$

Here  $c_1$  and  $c_2$  are complex parameters, not both 0.

To see how condition (3.3) determines  $\mathbf{v}(x, t)$  (and hence  $\dot{\mathbf{U}}(\lambda; x, t)$ ) in terms of the parameters  $\xi \in D_0$  and  $\mathbf{c} \in \mathbb{C}^2 \setminus \{\mathbf{0}\}$ , one starts from the Laurent expansion about the pole  $\lambda = \xi$  of  $\dot{\mathbf{U}}(\lambda; x, t)$  defined by (3.1):

$$(3.5) \quad \begin{aligned} \dot{\mathbf{U}}(\lambda; x, t) &= \mathbf{R}(x, t)\mathbf{U}(\xi; x, t)(\lambda - \xi)^{-1} \\ &+ \mathbf{U}(\xi; x, t) + \mathbf{R}(x, t)\mathbf{U}'(\xi; x, t) + O(\lambda - \xi), \quad \lambda \rightarrow \xi, \end{aligned}$$

where the prime denotes differentiation with respect to  $\lambda$ . On this expansion we impose the condition (3.3); the existence of the limit on the right-hand side requires that

$$(3.6) \quad \mathbf{R}(x, t)\mathbf{U}(\xi; x, t)\mathbf{C} = \mathbf{0},$$

and then matching the residue on the left-hand side to the limit on the right,

$$(3.7) \quad \mathbf{R}(x, t)\mathbf{U}(\xi; x, t) = (\mathbf{U}(\xi; x, t) + \mathbf{R}(x, t)\mathbf{U}'(\xi; x, t))\mathbf{C}.$$

Using the representation (3.4) in (3.6), we see that (since  $\mathbf{c}^\top \sigma_2 \neq \mathbf{0}$ ), the vector  $\mathbf{U}(\xi; x, t)\mathbf{c}$  must lie in the kernel of  $\mathbf{R}(x, t)$ . Given the form (3.2) of  $\mathbf{R}(x, t)$ , the vector  $\mathbf{v}(x, t)$  must be proportional to  $\mathbf{U}(\xi; x, t)\mathbf{c}$ , so (3.2) can be rewritten in the form

$$(3.8) \quad \mathbf{R}(x, t) = \varphi(x, t)\mathbf{U}(\xi; x, t)\mathbf{c}\mathbf{c}^\top \mathbf{U}(\xi; x, t)^\top \sigma_2,$$

where the only ambiguity remaining is  $\varphi(x, t)$ , a nonzero complex scalar. To determine  $\varphi(x, t)$ , we substitute (3.8) into (3.7); using (3.4) and the fact that for any  $2 \times 2$  matrix  $\mathbf{A}$  with  $\det(\mathbf{A}) = 1$ ,  $\mathbf{A}^\top \sigma_2 \mathbf{A} = \det(\mathbf{A})\sigma_2 = \sigma_2$ , we find

$$(3.9) \quad \begin{aligned} \mathbf{U}(\xi; x, t)\mathbf{c} \cdot \varphi(x, t) \cdot \mathbf{c}^\top \sigma_2 &= \\ \mathbf{U}(\xi; x, t)\mathbf{c} \cdot (1 + \varphi(x, t)\mathbf{c}^\top \mathbf{U}(\xi; x, t)^\top \sigma_2 \mathbf{U}'(\xi; x, t)\mathbf{c}) \cdot \mathbf{c}^\top \sigma_2. \end{aligned}$$

Since  $\mathbf{c} \neq \mathbf{0}$  and  $\mathbf{U}(\xi; x, t)$  is invertible, neither the column vector  $\mathbf{U}(\xi; x, t)\mathbf{c}$  nor the row vector  $\mathbf{c}^\top \sigma_2$  can vanish, so (3.9) is just a scalar linear equation that can be

solved explicitly for  $\varphi(x, t)$ :

$$(3.10) \quad \varphi(x, t) = \varphi(\xi, \mathbf{c}; x, t) := \frac{1}{1 - \mathbf{c}^\top \mathbf{U}(\xi; x, t)^\top \sigma_2 \mathbf{U}'(\xi; x, t) \mathbf{c}}.$$

### 3.2 Darboux Transformation for the Robust IST

The insertion of a single pole breaks the Schwarz symmetry (2.8) initially present in the matrix  $\mathbf{U}(\lambda; x, t)$ . This is obviously true if  $\xi \notin \mathbb{R}$ , but even if one takes  $\xi \in D_0 \cap \mathbb{R}$  one finds that the condition  $\mathbf{R}(x, t) \neq 0$  is inconsistent with Schwarz symmetry of the Laurent expansion of  $\dot{\mathbf{U}}(\lambda; x, t)$  about  $\lambda = \xi$ . To maintain the Schwarz symmetry it is necessary to insert a complex-conjugate pair of poles at points  $\xi$  and  $\xi^*$  in  $D_0$  with  $\xi \neq \xi^*$ . The insertion of the conjugate pair of poles is done in two consecutive steps, after which the poles are removed in favor of a modified jump:

- A gauge transformation of the form (3.1) is applied with data  $(\xi, \mathbf{c})$  as described in Section 3.1. This step generates  $\dot{\mathbf{U}}(\lambda; x, t)$  from  $\mathbf{U}(\lambda; x, t)$ .
- A second gauge transformation of the form (3.1) is applied this time with data  $(\xi^*, \sigma_2 \mathbf{c}^*)$ . This step generates a matrix  $\ddot{\mathbf{U}}(\lambda; x, t)$  from  $\dot{\mathbf{U}}(\lambda; x, t)$ .
- The two new simple-pole singularities of  $\ddot{\mathbf{U}}(\lambda; x, t)$  are removed and instead transferred to the jump condition across the circle  $\Sigma_0$  by making an explicit renormalization for  $\lambda \in D_0$ :

$$(3.11) \quad \tilde{\mathbf{U}}(\lambda; x, t) := \begin{cases} \ddot{\mathbf{U}}(\lambda; x, t), & \lambda \in D_+ \cup D_-, \\ \ddot{\mathbf{U}}(\lambda; x, t) \ddot{\mathbf{U}}(\lambda; L, 0)^{-1}, & \lambda \in D_0. \end{cases}$$

We first construct  $\ddot{\mathbf{U}}(\lambda; x, t)$  by explicitly composing the two pole insertions:

$$(3.12) \quad \begin{aligned} & \ddot{\mathbf{U}}(\lambda; x, t) \\ &= \left( \mathbb{I} + \frac{\dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) \dot{\mathbf{U}}(\xi^*; x, t) \sigma_2 \mathbf{c}^* \mathbf{c}^\dagger \sigma_2^\top \dot{\mathbf{U}}(\xi^*; x, t)^\top \sigma_2}{\lambda - \xi^*} \right) \dot{\mathbf{U}}(\lambda; x, t), \end{aligned}$$

where

$$(3.13) \quad \dot{\mathbf{U}}(\lambda; x, t) = \left( \mathbb{I} + \frac{\varphi(\xi, \mathbf{c}; x, t) \mathbf{U}(\xi; x, t) \mathbf{c} \mathbf{c}^\top \mathbf{U}(\xi; x, t)^\top \sigma_2}{\lambda - \xi} \right) \mathbf{U}(\lambda; x, t),$$

and in (3.12),  $\dot{\varphi}$  is defined from (3.10) using the matrix function  $\dot{\mathbf{U}}(\lambda; x, t)$  in place of  $\mathbf{U}(\lambda; x, t)$ . Inserting (3.13) into (3.12) and expanding the product of the parentheses in partial fractions gives  $\ddot{\mathbf{U}}(\lambda; x, t) = \mathbf{G}(\lambda; x, t) \mathbf{U}(\lambda; x, t)$ , where the composite gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  is

$$(3.14) \quad \mathbf{G}(\lambda; x, t) := \mathbb{I} + \frac{\mathbf{Y}(x, t)}{\lambda - \xi} + \frac{\mathbf{Z}(x, t)}{\lambda - \xi^*}$$

with coefficients  $\mathbf{Y}(x, t)$  and  $\mathbf{Z}(x, t)$  described as follows. Set

$$(3.15) \quad \begin{aligned} \beta &:= \text{Im}(\xi), & \mathbf{s}(x, t) &:= \mathbf{U}(\xi; x, t) \mathbf{c}, \\ N(x, t) &:= \|\mathbf{s}(x, t)\|^2 = \mathbf{s}(x, t)^\dagger \mathbf{s}(x, t). \end{aligned}$$

By using the given Schwarz symmetry condition  $\mathbf{U}(\lambda^*; x, t) = \sigma_2 \mathbf{U}(\lambda; x, t)^* \sigma_2$  (cf. (2.8)), along with the identity  $\mathbf{v}^\top \sigma_2 \mathbf{v} = 0$  for all vectors  $\mathbf{v}$ , we find

$$\begin{aligned}
 \mathbf{Y}(x, t) &= \varphi(\xi, \mathbf{c}; x, t) \left( 1 - \frac{\varphi(\xi, \mathbf{c}; x, t) \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) N(x, t)^2}{4\beta^2} \right) \mathbf{s}(x, t) \mathbf{s}(x, t)^\top \sigma_2 \\
 &\quad - \frac{\varphi(\xi, \mathbf{c}; x, t) \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) N(x, t)}{2i\beta} \sigma_2 \mathbf{s}(x, t)^* \mathbf{s}(x, t)^\top \sigma_2
 \end{aligned}
 \tag{3.16}$$

and

$$\begin{aligned}
 \mathbf{Z}(x, t) &= -\dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) \sigma_2 \mathbf{s}(x, t)^* \mathbf{s}(x, t)^\dagger \\
 &\quad + \frac{\dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) \varphi(\xi, \mathbf{c}; x, t) N(x, t)}{2i\beta} \mathbf{s}(x, t) \mathbf{s}(x, t)^\dagger.
 \end{aligned}
 \tag{3.17}$$

We simplify the scalars

$$\varphi(\xi, \mathbf{c}; x, t), \quad \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t), \quad \text{and} \quad \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) \varphi(\xi, \mathbf{c}; x, t)$$

as follows. First note that  $\varphi(\xi, \mathbf{c}; x, t)$  can be expressed as

$$\varphi(\xi, \mathbf{c}; x, t) = \frac{1}{1 - w(x, t)} = \frac{1 - w(x, t)^*}{|1 - w(x, t)|^2},
 \tag{3.18}$$

where

$$\begin{aligned}
 w(x, t) &:= \mathbf{c}^\top \mathbf{U}(\xi; x, t)^\top \sigma_2 \mathbf{U}'(\xi; x, t) \mathbf{c} = \mathbf{s}(x, t)^\top \sigma_2 \mathbf{s}'(x, t), \\
 \mathbf{s}'(x, t) &:= \mathbf{U}'(\xi; x, t) \mathbf{c}.
 \end{aligned}
 \tag{3.19}$$

Next, note that

$$\begin{aligned}
 &\mathbf{c}^\dagger \sigma_2^\top \dot{\mathbf{U}}(\xi^*; x, t)^\top \sigma_2 \dot{\mathbf{U}}'(\xi^*; x, t) \sigma_2 \mathbf{c}^* \\
 &= -\frac{\varphi(\xi, \mathbf{c}; x, t) N(x, t)^2}{4\beta^2} - \mathbf{s}(x, t)^\dagger \sigma_2 \mathbf{s}'(x, t)^* \\
 &= -\frac{\varphi(\xi, \mathbf{c}; x, t) N(x, t)^2}{4\beta^2} + w(x, t)^*,
 \end{aligned}
 \tag{3.20}$$

and hence  $\dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t)$  can be simplified to

$$\begin{aligned}
 \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) &= \frac{1}{1 - \mathbf{c}^\dagger \sigma_2^\top \dot{\mathbf{U}}(\xi^*; x, t)^\top \sigma_2 \dot{\mathbf{U}}'(\xi^*; x, t) \sigma_2 \mathbf{c}^*} \\
 &= 4\beta^2 \frac{|1 - w(x, t)|^2}{(4\beta^2 |1 - w(x, t)|^2 + N(x, t)^2)(1 - w(x, t)^*)}.
 \end{aligned}
 \tag{3.21}$$

Combining (3.18) and (3.21) then gives

$$\varphi(\xi, \mathbf{c}; x, t) \dot{\varphi}(\xi^*, \sigma_2 \mathbf{c}^*; x, t) = \frac{4\beta^2}{4\beta^2 |1 - w(x, t)|^2 + N(x, t)^2}.
 \tag{3.22}$$

Using (3.18), (3.21), and (3.22) in (3.16) and (3.17) gives

$$\begin{aligned}
 \mathbf{Y}(x, t) &= \frac{4\beta^2(1 - w(x, t)^*)}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2} \mathbf{s}(x, t)\mathbf{s}(x, t)^\top \sigma_2 \\
 &+ \frac{2i\beta N(x, t)}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2} \sigma_2 \mathbf{s}(x, t)^* \mathbf{s}(x, t)^\top \sigma_2
 \end{aligned}
 \tag{3.23}$$

and

$$\begin{aligned}
 \mathbf{Z}(x, t) &= \frac{-4\beta^2(1 - w(x, t))}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2} \sigma_2 \mathbf{s}(x, t)^* \mathbf{s}(x, t)^\dagger \\
 &+ \frac{-2i\beta N(x, t)}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2} \mathbf{s}(x, t)\mathbf{s}(x, t)^\dagger.
 \end{aligned}
 \tag{3.24}$$

Note that

$$4\beta^2|1 - w(x, t)|^2 + N(x, t)^2 \geq N(x, t)^2,
 \tag{3.25}$$

which can never vanish for any  $(x, t) \in \mathbb{R}^2$  because  $\mathbf{c} \neq \mathbf{0}$  and  $N(x, t)$  is the squared length of the vector  $\mathbf{s}(x, t)$ , a nontrivial linear combination of the columns of a matrix with unit determinant. It is now completely obvious that  $\mathbf{Z}(x, t) = \sigma_2 \mathbf{Y}(x, t)^* \sigma_2$ , which implies that  $\mathbf{G}(\lambda^*; x, t) = \sigma_2 \mathbf{G}(\lambda; x, t)^* \sigma_2$  and so the transformed matrix function  $\ddot{\mathbf{U}}(\lambda; x, t) = \mathbf{G}(\lambda; x, t)\mathbf{U}(\lambda; x, t)$  maintains the Schwarz symmetry of the original matrix function  $\mathbf{U}(\lambda; x, t)$ , namely

$$\ddot{\mathbf{U}}(\lambda^*; x, t) = \sigma_2 \ddot{\mathbf{U}}(\lambda; x, t)^* \sigma_2.$$

A ‘‘dressing’’ construction shows that for each  $\lambda \in \mathbb{C} \setminus (\Sigma \cup \{\xi, \xi^*\})$ ,  $\ddot{\mathbf{U}}(\lambda; x, t)$  simultaneously satisfies Lax pair equations of the form (1.8)–(1.9) in which  $\psi(x, t)$  is replaced by a modified potential  $\tilde{\psi}(x, t)$  given by

$$\begin{aligned}
 \tilde{\psi}(x, t) &= 2i \lim_{\lambda \rightarrow \infty} \lambda \ddot{M}_{12}(\lambda; x, t) \\
 &= 2iY_{12}(x, t) + 2iZ_{12}(x, t) + 2i \lim_{\lambda \rightarrow \infty} \lambda M_{12}(\lambda; x, t) \\
 &= \psi(x, t) + 2iY_{12}(x, t) + 2iZ_{12}(x, t) \\
 &= \psi(x, t) + 2i(Y_{12}(x, t) - Y_{21}(x, t)^*),
 \end{aligned}
 \tag{3.26}$$

where  $\ddot{\mathbf{M}}(\lambda; x, t) := \ddot{\mathbf{U}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$  satisfies  $\ddot{\mathbf{M}}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$ . To establish the formula (3.26), one notes that since each gauge transformation step  $\mathbf{U}(\lambda; x, t) \rightarrow \dot{\mathbf{U}}(\lambda; x, t) \rightarrow \ddot{\mathbf{U}}(\lambda; x, t)$  involves multiplication on the left by a matrix that is analytic everywhere except at the pole to be inserted, all pre-existent jump conditions and residue conditions of the form (3.3) are preserved. Thus  $\ddot{\mathbf{U}}(\lambda; x, t)$  is analytic for  $\lambda \in \mathbb{C} \setminus (\Sigma \cup \{\xi, \xi^*\})$ , and it satisfies the jump condition  $\ddot{\mathbf{U}}_+(\lambda; x, t) = \ddot{\mathbf{U}}_-(\lambda; x, t)\mathbf{V}(\lambda)$  for  $\lambda \in \Sigma$  exactly as does  $\mathbf{U}(\lambda; x, t)$  according to

Riemann-Hilbert Problem 1. Moreover, the simple poles at  $\xi$  and  $\xi^*$  are characterized by the related residue conditions

$$\begin{aligned}
 (3.27) \quad \operatorname{Res}_{\lambda=\xi} \ddot{\mathbf{U}}(\lambda; x, t) &= \lim_{\lambda \rightarrow \xi} \ddot{\mathbf{U}}(\lambda; x, t) \mathbf{C}, \\
 \operatorname{Res}_{\lambda=\xi^*} \ddot{\mathbf{U}}(\lambda; x, t) &= \lim_{\lambda \rightarrow \xi^*} \ddot{\mathbf{U}}(\lambda; x, t) \sigma_2 \mathbf{C}^* \sigma_2.
 \end{aligned}$$

Because the core jump matrix  $\mathbf{V}(\lambda)$  and the residue matrix  $\mathbf{C} = \mathbf{c}\mathbf{c}^\top \sigma_2$  are independent of  $(x, t) \in \mathbb{R}^2$ ,  $\ddot{\mathbf{U}}(\lambda; x, t)$  and its partial derivatives  $\ddot{\mathbf{U}}_x(\lambda; x, t)$  and  $\ddot{\mathbf{U}}_t(\lambda; x, t)$  are all analytic in the same domain and satisfy the same jump and residue conditions, and it then follows that the matrices

$$\begin{aligned}
 (3.28) \quad \ddot{\mathbf{X}}(\lambda; x, t) &:= \ddot{\mathbf{U}}_x(\lambda; x, t) \ddot{\mathbf{U}}(\lambda; x, t)^{-1}, \\
 \ddot{\mathbf{T}}(\lambda; x, t) &:= \ddot{\mathbf{U}}_t(\lambda; x, t) \ddot{\mathbf{U}}(\lambda; x, t)^{-1},
 \end{aligned}$$

have only removable singularities, and hence are essentially entire functions of  $\lambda$ .

Moreover, assuming that the asymptotic expansion

$$(3.29) \quad \ddot{\mathbf{M}}(\lambda; x, t) = \mathbb{I} + \frac{\ddot{\mathbf{M}}^{(1)}(x, t)}{\lambda} + \frac{\ddot{\mathbf{M}}^{(2)}(x, t)}{\lambda^2} + O(\lambda^{-3}), \quad \lambda \rightarrow \infty,$$

is differentiable term by term with respect to  $(x, t)$ , it follows that

$$\begin{aligned}
 (3.30) \quad \ddot{\mathbf{X}}(\lambda; x, t) &= \left( \ddot{\mathbf{M}}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \right)_x e^{i\rho(\lambda)(x+\lambda t)\sigma_3} \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &= -i\rho(\lambda) \ddot{\mathbf{M}}(\lambda; x, t) \sigma_3 \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &\quad + \ddot{\mathbf{M}}_x(\lambda; x, t) \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &= -i\lambda \sigma_3 + i \left[ \sigma_3, \ddot{\mathbf{M}}^{(1)}(x, t) \right] + O(\lambda^{-1}) \\
 &= -i\lambda \sigma_3 + i \left[ \sigma_3, \ddot{\mathbf{M}}^{(1)}(x, t) \right],
 \end{aligned}$$

and

$$\begin{aligned}
 (3.31) \quad \ddot{\mathbf{T}}(\lambda; x, t) &= \left( \ddot{\mathbf{M}}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \right)_t e^{i\rho(\lambda)(x+\lambda t)\sigma_3} \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &= -i\rho(\lambda) \lambda \ddot{\mathbf{M}}(\lambda; x, t) \sigma_3 \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &\quad + \ddot{\mathbf{M}}_t(\lambda; x, t) \ddot{\mathbf{M}}(\lambda; x, t)^{-1} \\
 &= -i\lambda^2 \sigma_3 + i\lambda \left[ \sigma_3, \ddot{\mathbf{M}}^{(1)}(x, t) \right] + i \left[ \sigma_3, \ddot{\mathbf{M}}^{(2)}(x, t) \right] \\
 &\quad + i \left[ \ddot{\mathbf{M}}^{(1)}(x, t), \sigma_3 \ddot{\mathbf{M}}^{(1)}(x, t) \right] - \frac{i}{2} \sigma_3 + O(\lambda^{-1}) \\
 &= -i\lambda^2 \sigma_3 + i\lambda \left[ \sigma_3, \ddot{\mathbf{M}}^{(1)}(x, t) \right] + i \left[ \sigma_3, \ddot{\mathbf{M}}^{(2)}(x, t) \right] \\
 &\quad + i \left[ \ddot{\mathbf{M}}^{(1)}(x, t), \sigma_3 \ddot{\mathbf{M}}^{(1)}(x, t) \right] - \frac{i}{2} \sigma_3,
 \end{aligned}$$

where the last equality in each case is a consequence of Liouville’s theorem. The dependence on  $\ddot{\mathbf{M}}^{(2)}(x, t)$  in (3.31) can be eliminated because the coefficient of  $\lambda^{-1}$  in the  $O(\lambda^{-1})$  error term on the third line of (3.30) is

$$(3.32) \quad i \left[ \sigma_3, \ddot{\mathbf{M}}^{(2)}(x, t) \right] + i \left[ \ddot{\mathbf{M}}^{(1)}(x, t), \sigma_3 \ddot{\mathbf{M}}^{(1)}(x, t) \right] + \ddot{\mathbf{M}}_x^{(1)}(x, t) - \frac{i}{2} \sigma_3$$

which must vanish (again by Liouville’s theorem). Therefore  $\ddot{\mathbf{T}}(\lambda; x, t)$  is the quadratic polynomial

$$(3.33) \quad \ddot{\mathbf{T}}(\lambda; x, t) = -i\lambda^2 \sigma_3 + i\lambda \left[ \sigma_3, \ddot{\mathbf{M}}^{(1)}(x, t) \right] - \ddot{\mathbf{M}}_x^{(1)}(x, t),$$

and from the diagonal part of (3.32) we see that

$$(3.34) \quad \begin{aligned} \ddot{M}_{11,x}^{(1)}(x, t) &= 2i\ddot{M}_{12}^{(1)}(x, t)\ddot{M}_{21}^{(1)}(x, t) + \frac{i}{2}, \\ \ddot{M}_{22,x}^{(1)}(x, t) &= -2i\ddot{M}_{12}^{(1)}(x, t)\ddot{M}_{21}^{(1)}(x, t) - \frac{i}{2}. \end{aligned}$$

It then follows from the fact that  $\ddot{\mathbf{M}}(\lambda^*; x, t) = \sigma_2 \ddot{\mathbf{M}}(\lambda; x, t)^* \sigma_2$  and definition (3.26) that  $\ddot{\mathbf{X}}(\lambda; x, t)$  and  $\ddot{\mathbf{T}}(\lambda; x, t)$  have exactly the forms (1.8) and (1.9), respectively, in which  $\psi(x, t)$  is merely replaced with  $\tilde{\psi}(x, t)$ . Rewriting (3.28) in the form  $\ddot{\mathbf{U}}_x = \ddot{\mathbf{X}}\ddot{\mathbf{U}}$  and  $\ddot{\mathbf{U}}_t = \ddot{\mathbf{T}}\ddot{\mathbf{U}}$  shows that  $\ddot{\mathbf{U}}(\lambda; x, t)$  is indeed a simultaneous solution matrix for the Lax pair (1.8)–(1.9) for the modified potential  $\tilde{\psi}(x, t)$ . This Lax pair is therefore also compatible, and it follows that  $\tilde{\psi}(x, t)$  is a solution of the focusing NLS equation (1.1). The bound (3.25) and the formula (3.26) together imply that since by hypothesis (cf. Theorem 2.4)  $\psi(x, t)$  is a global solution of the focusing NLS equation (1.1), so is the transformed potential  $\tilde{\psi}(x, t)$ . Therefore, (3.26) constitutes the *Bäcklund transformation* that corresponds to the twice-iterated gauge transformation  $\mathbf{U} \mapsto \ddot{\mathbf{U}}$ .

Having constructed  $\ddot{\mathbf{U}}(\lambda; x, t)$ , it remains only to implement the final step by defining  $\tilde{\mathbf{U}}(\lambda; x, t)$  from  $\ddot{\mathbf{U}}(\lambda; x, t)$  using (3.11). For this, we observe that for  $\lambda \in D_0$ , the original matrix  $\mathbf{U}(\lambda; x, t)$  is the solution  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  of an initial-value problem for the simultaneous equations of the Lax pair (1.8)–(1.9) (cf. Proposition 2.1) and therefore  $\mathbf{U}(\lambda; L, 0) = \mathbb{I}$  holds for all  $\lambda \in D_0$ . Consequently, (3.11) can be explicitly written in terms of the composite gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  as

$$(3.35) \quad \tilde{\mathbf{U}}(\lambda; x, t) := \begin{cases} \mathbf{G}(\lambda; x, t)\mathbf{U}(\lambda; x, t), & \lambda \in D_+ \cup D_-, \\ \mathbf{G}(\lambda; x, t)\mathbf{U}(\lambda; x, t)\mathbf{G}(\lambda; L, 0)^{-1}, & \lambda \in D_0. \end{cases}$$

Since the multiplication of  $\ddot{\mathbf{U}}(\lambda; x, t)$  on the right by  $\mathbf{G}(\lambda; L, 0)^{-1}$  for  $\lambda \in D_0$  preserves the Lax pair equations (1.8)–(1.9) for the modified potential  $\tilde{\psi}(x, t)$ , it follows that for all  $\lambda \in D_0$  with the possible exception of  $\xi, \xi^*$ ,  $\tilde{\mathbf{U}}(\lambda; x, t)$  is a simultaneous solution matrix for the modified Lax pair that also satisfies  $\tilde{\mathbf{U}}(\lambda; L, 0) = \mathbb{I}$ . Therefore by the uniqueness asserted in Proposition 2.1, for  $\lambda \in D_0$ ,  $\tilde{\mathbf{U}}(\lambda; L, 0)$



has removable singularities at  $\xi$  and  $\xi^*$  and therefore is analytic within  $D_0$  (actually it can be continued to  $\lambda \in \mathbb{C}$  as the entire function  $\tilde{\mathbf{U}}^{\text{in}}(\lambda; x, t)$  for the modified potential  $\tilde{\psi}(x, t)$ ). Since  $\xi, \xi^* \in D_0$ , it is also obvious that  $\tilde{\mathbf{U}}(\lambda; x, t)$  is analytic for  $\lambda \in D_+ \cup D_-$  and since  $\mathbf{G}(\lambda; x, t) \rightarrow \mathbb{I}$  as  $\lambda \rightarrow \infty$ , the related matrix  $\tilde{\mathbf{M}}(\lambda; x, t) := \tilde{\mathbf{U}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$  satisfies the conditions of Riemann-Hilbert Problem 1 in which the only change is that for  $\lambda \in \Sigma_0 \subset \Sigma$ , the core jump matrix  $\mathbf{V}(\lambda)$  is replaced by

$$(3.36) \quad \tilde{\mathbf{V}}(\lambda) := \begin{cases} \mathbf{G}(\lambda; L, 0)\mathbf{V}(\lambda), & \lambda \in \Sigma_+ \\ \mathbf{V}(\lambda)\mathbf{G}(\lambda; L, 0)^{-1}, & \lambda \in \Sigma_- \end{cases}$$

A crucial fact is that since  $\mathbf{G}(\lambda^*; x, t) = \sigma_2 \mathbf{G}(\lambda; x, t)^* \sigma_2$ , the modified jump matrix satisfies exactly the same Schwarz symmetry condition as does the original jump matrix, and hence the proof of Theorem 2.4 applies once again to guarantee unique solvability of the transformed Riemann-Hilbert problem for all  $(x, t) \in \mathbb{R}^2$ .

*Remark 3.1.* The use of nilpotent residue matrices  $\mathbf{C}$  in (3.27) that are not necessarily triangular matrices may seem unusual to some readers used to applying Darboux transformations to Beals-Coifman solutions in order to introduce poles specifically into one column or the other. Here the generalization is necessary because  $\xi \in D_0$  and in this domain  $\mathbf{U}(\lambda; x, t)$  differs from  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  by a right-multiplication by a matrix depending on  $\lambda$  only. This matrix factor conjugates triangular nilpotent residue matrices into general 2-nilpotent form.

*Remark 3.2.* Since after multiplying on the right by  $e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$  the result of the Darboux transformation  $\mathbf{U}(\lambda; x, t) \mapsto \tilde{\mathbf{U}}(\lambda; x, t)$  described above also satisfies Riemann-Hilbert Problem 1 with only a modified jump matrix on the circle  $\Sigma_0$ , it is obvious that the transformation may be iterated any number of times. In each iteration it makes no difference whether the points  $\xi, \xi^*$  at which the poles are first introduced and then transferred to a jump on  $\Sigma_0$  vary from iteration to iteration or whether they are fixed once and for all.

*Remark 3.3.* In each iteration, the poles may be placed at any nonreal conjugate pair of points within  $D_0$ . Although the matrix

$$\tilde{\mathbf{M}}(\lambda; x, t) = \tilde{\mathbf{U}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$$

satisfies the conditions of Riemann-Hilbert Problem 1 for a modified core jump matrix  $\tilde{\mathbf{V}}(\lambda)$  on  $\Sigma_0$ , it need not be the case that the resulting potential  $\tilde{\psi}(x, t)$  satisfies the boundary condition (1.2) initially satisfied by  $\psi(x, t)$ . For instance, if one starts with the Riemann-Hilbert problem for the background solution and applies a Darboux transformation in which  $\xi$  is taken to lie on the imaginary axis between the origin and  $\pm i$ , then for some choices of the auxiliary parameters  $\mathbf{c}$  the resulting solution  $\tilde{\psi}(x, t)$  is an *Akhmediev breather* [3], a solution that is periodic in  $x$  for fixed  $t$ , and hence does not decay to the background as  $x \rightarrow \pm\infty$ . The deciding factor in whether the boundary conditions are preserved is whether the

gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  decays to the identity as  $x \rightarrow +\infty$  and remains bounded as  $x \rightarrow -\infty$ . In many cases, this decay can be verified directly.

*Remark 3.4.* The composite gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  that takes  $\mathbf{U}(\lambda; x, t)$  into  $\check{\mathbf{U}}(\lambda; x, t)$  also makes sense in a limiting case where  $\xi \in D_0 \setminus \mathbb{R}$  is fixed and  $\mathbf{c}$  is taken in the form  $\varepsilon^{-1}\mathbf{c}_\infty$  for a fixed vector  $\mathbf{c}_\infty \in \mathbb{C}^2 \setminus \{\mathbf{0}\}$  and  $\varepsilon \in \mathbb{C}$ , after which the limit  $\varepsilon \rightarrow 0$  is taken in the matrices  $\mathbf{Y}(x, t)$  and  $\mathbf{Z}(x, t)$  given in (3.23) and (3.24), respectively. If we denote by  $\mathbf{s}_\infty(x, t)$ ,  $\mathbf{s}'_\infty(x, t)$ ,  $N_\infty(x, t)$ , and  $w_\infty(x, t)$  the quantities defined in (3.15) and (3.19) where  $\mathbf{c}$  is simply replaced with  $\mathbf{c}_\infty$ , then  $\mathbf{s}(x, t) = \varepsilon^{-1}\mathbf{s}_\infty(x, t)$  and  $\mathbf{s}'(x, t) = \varepsilon^{-1}\mathbf{s}'_\infty(x, t)$  while  $N(x, t) = |\varepsilon|^{-2}N_\infty(x, t)$  and  $w(x, t) = \varepsilon^{-2}w_\infty(x, t)$ , so it follows easily that in this situation

$$\begin{aligned} \mathbf{Y}_\infty(x, t) &:= \lim_{\varepsilon \rightarrow 0} \mathbf{Y}(x, t) \\ (3.37) \quad &= -\frac{4\beta^2 w_\infty(x, t)^*}{4\beta^2 |w_\infty(x, t)|^2 + N_\infty(x, t)^2} \mathbf{s}_\infty(x, t) \mathbf{s}_\infty(x, t)^\top \sigma_2 \\ &\quad + \frac{2i\beta N_\infty(x, t)}{4\beta^2 |w_\infty(x, t)|^2 + N_\infty(x, t)^2} \sigma_2 \mathbf{s}_\infty(x, t)^* \mathbf{s}_\infty(x, t)^\top \sigma_2 \end{aligned}$$

and

$$\begin{aligned} \mathbf{Z}_\infty(x, t) &:= \lim_{\varepsilon \rightarrow 0} \mathbf{Z}(x, t) \\ (3.38) \quad &= \frac{4\beta^2 w_\infty(x, t)}{4\beta^2 |w_\infty(x, t)|^2 + N_\infty(x, t)^2} \sigma_2 \mathbf{s}_\infty(x, t)^* \mathbf{s}_\infty(x, t)^\dagger \\ &\quad - \frac{2i\beta N_\infty(x, t)}{4\beta^2 |w_\infty(x, t)|^2 + N_\infty(x, t)^2} \mathbf{s}_\infty(x, t) \mathbf{s}_\infty(x, t)^\dagger. \end{aligned}$$

The corresponding composite gauge transformation

$$(3.39) \quad \mathbf{G}_\infty(\lambda; x, t) := \mathbb{I} + (\lambda - \xi)^{-1} \mathbf{Y}_\infty(x, t) + (\lambda - \xi^*)^{-1} \mathbf{Z}_\infty(x, t)$$

can then be used in place of  $\mathbf{G}(\lambda; x, t)$  whenever desired. This limiting case is especially useful for locating rogue wave solutions at the normalization point  $(x, t) = (L, 0)$  as will be seen shortly. It is also obvious from the formulae (3.37)–(3.38) that  $\mathbf{G}_\infty(\lambda; x, t)$  only depends on  $\mathbf{c}_\infty \in \mathbb{C}^2 \setminus \{\mathbf{0}\}$  up to a nonzero complex multiple; in other words, the parameter space of the limiting Darboux transformation for fixed  $\xi$  is a complex projective space:  $\mathbf{c}_\infty \in \mathbb{C}\mathbb{P}^1$ . Henceforth, we write the column vector  $\mathbf{c}_\infty$  in terms of its homogeneous coordinates as  $\mathbf{c}_\infty = [c_1 : c_2]^\top$ .

We summarize this description of the Darboux transformation for the robust IST in the following theorem.

**THEOREM 3.5.** *Let  $\mathbf{U}(\lambda; x, t) := \mathbf{M}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$  correspond to the unique solution  $\mathbf{M}(\lambda; x, t)$  of Riemann-Hilbert Problem 1 formulated with core jump matrices  $\mathbf{V}^\mathbb{R}(\lambda)$  defined on  $\Sigma_L \cup \Sigma_R$  and  $\mathbf{V}^\pm(\lambda)$  defined on  $\Sigma_\pm$ , and generating a solution  $\psi(x, t)$  of the focusing nonlinear Schrödinger equation (1.1) via the formula (2.26). Given Darboux transformation data  $\xi \in D_0 \setminus \mathbb{R}$  and  $\mathbf{c} \in \mathbb{C}^2 \setminus \{\mathbf{0}\}$ ,*

define the gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  by (3.14) in terms of the coefficient matrices  $\mathbf{Y}(x, t)$  and  $\mathbf{Z}(x, t)$  given by (3.23) and (3.24), respectively, using (3.15) and (3.19). Then

- $\tilde{\mathbf{M}}(\lambda; x, t) := \tilde{\mathbf{U}}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$ , where  $\tilde{\mathbf{U}}(\lambda; x, t)$  is defined explicitly in terms of  $\mathbf{G}(\lambda; x, t)$  and  $\mathbf{U}(\lambda; x, t)$  by (3.35), is the unique solution of another Riemann-Hilbert problem of the form of Riemann-Hilbert Problem 1 with the same jump contour and core jump matrices  $\tilde{\mathbf{V}}^{\mathbb{R}}(\lambda) := \mathbf{V}^{\mathbb{R}}(\lambda)$  for  $\lambda \in \Sigma_L \cup \Sigma_R$  and  $\tilde{\mathbf{V}}(\lambda)$  is defined in terms of  $\mathbf{V}(\lambda)$  for  $\lambda \in \Sigma_0 = \Sigma_+ \cup \Sigma_-$  by (3.36).
- The transformed Riemann-Hilbert problem generates a new solution of (1.1),  $\tilde{\psi}(x, t)$ , by the formula (2.26) in which  $\mathbf{M}(\lambda; x, t)$  is replaced by  $\tilde{\mathbf{M}}(\lambda; x, t)$  on the right-hand side. Equivalently,  $\tilde{\psi}(x, t)$  is given in terms of  $\psi(x, t)$  and the Darboux transformation data by the Bäcklund transformation (3.26).

All of these statements also apply for transformation data  $(\xi, \mathbf{c})$  with  $\mathbf{c} = \infty$  in the sense of Remark 3.4; i.e.,  $\mathbf{c}$  is replaced by  $\mathbf{c}_\infty \in \mathbb{C}\mathbb{P}^1$  and  $\mathbf{Y}(x, t)$  and  $\mathbf{Z}(x, t)$  are replaced by  $\mathbf{Y}_\infty(x, t)$  and  $\mathbf{Z}_\infty(x, t)$  given by (3.37) and (3.38), respectively.

### 3.3 The Simplest Case: Darboux/Bäcklund Transformation of the Background Potential in the Setting of the Robust IST

As mentioned in Section 1.1, the background potential  $\psi = \psi_{\text{bg}}(x, t) \equiv 1$  has the simultaneous fundamental solution matrix  $\mathbf{U}_{\text{bg}}(\lambda; x, t) = \mathbf{E}(\lambda) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$  for the corresponding linear equations of the Lax pair (1.8)–(1.9). Here  $\mathbf{E}(\lambda)$  is the matrix function defined in (1.10). Taking  $t = 0$  and  $\lambda \in \Gamma$ , we may normalize this solution in the limits  $x \rightarrow \pm\infty$  to obtain the Jost solution matrices  $\mathbf{J}_{\text{bg}}^\pm(\lambda; x)$ . It is easy to confirm that in fact  $\mathbf{J}_{\text{bg}}^+(\lambda; x) = \mathbf{J}_{\text{bg}}^-(\lambda; x) = \mathbf{U}_{\text{bg}}(\lambda; x, 0)$ , so:

- The scattering matrix satisfies  $\mathbf{S}_{\text{bg}}(\lambda; t) = \mathbb{I}$  for all  $\lambda \in \Gamma$  and  $t \in \mathbb{R}$ .
- The Beals-Coifman matrix is analytic for  $\lambda \in \mathbb{C} \setminus \Sigma_c$  and is given explicitly by  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) = \mathbf{U}_{\text{bg}}(\lambda; x, t)$ .

Therefore, the scattering data for the background solution  $\psi = \psi_{\text{bg}}(x, t) \equiv 1$  in the setting of the robust transform described in Section 2 consists of:

- the core jump matrix  $\mathbf{V}_{\text{bg}}^{\mathbb{R}}(\lambda) \equiv \mathbb{I}$  for  $\lambda \in \Sigma_L \cup \Sigma_R$ ,
- the core jump matrix  $\mathbf{V}_{\text{bg}}^+(\lambda) = \mathbf{E}(\lambda)$  for  $\lambda \in \Sigma_+$ , and
- the core jump matrix  $\mathbf{V}_{\text{bg}}^-(\lambda) = \mathbf{E}(\lambda)^{-1}$  for  $\lambda \in \Sigma_-$ .

Here the radius  $r = r_{\text{bg}}$  of  $\Sigma_0$  can be taken to be any number larger than 1. It will also be useful to have an explicit expression for the matrix  $\mathbf{U}_{\text{bg}}^{\text{in}}(\lambda; x, t)$  described by Proposition 2.1 for the case of  $\psi = \psi_{\text{bg}}(x, t) \equiv 1$ . It can be found by normalizing<sup>5</sup>

---

<sup>5</sup>For convenience, for the rest of Section 3 we choose the normalization point of the robust IST to be  $L = 0$ .

the matrix  $\mathbf{U}_{\text{bg}}(\lambda; x, t)$  at  $(x, t) = (0, 0)$  to the identity:

$$(3.40) \quad \begin{aligned} \mathbf{U}_{\text{bg}}^{\text{in}}(\lambda; x, t) &= \mathbf{U}_{\text{bg}}(\lambda; x, t)\mathbf{U}_{\text{bg}}(\lambda; 0, 0)^{-1} \\ &= \mathbf{E}(\lambda) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{E}(\lambda)^{-1}. \end{aligned}$$

As pointed out in Remark 2.2, this formula can have only removable singularities for  $\lambda \in \mathbb{C}$ , and we can see it explicitly by using the definition of  $\mathbf{E}(\lambda)$  given in (1.10)–(1.11). Thus, (3.40) can be rewritten as

$$(3.41) \quad \mathbf{U}_{\text{bg}}^{\text{in}}(\lambda; x, t) = (x + \lambda t) \frac{\sin(\theta)}{\theta} \begin{bmatrix} -i\lambda & 1 \\ -1 & i\lambda \end{bmatrix} + \cos(\theta)\mathbb{I}, \quad \theta := \rho(\lambda)(x + \lambda t),$$

and since  $\sin(\theta)/\theta$  and  $\cos(\theta)$  are both even entire functions of  $\theta$ , the desired analyticity for  $\lambda \in \mathbb{C}$  is now obvious because  $\rho(\lambda)^2 = \lambda^2 + 1$ .

Now we choose an arbitrary point  $\xi \in \mathbb{C} \setminus \mathbb{R}$  and assume that the radius  $r$  of the circle  $\Sigma_0$  exceeds  $|\xi|$  so that  $\xi \in D_0$ . We also select a vector  $\mathbf{c} \in \mathbb{C}^2 \setminus \{\mathbf{0}\}$ . To apply the Darboux transformation we need to calculate the vectors  $\mathbf{s}(x, t) := \mathbf{U}_{\text{bg}}(\xi; x, t)\mathbf{c} = \mathbf{U}_{\text{bg}}^{\text{in}}(\xi; x, t)\mathbf{c}$  and  $\mathbf{s}'(x, t) := \mathbf{U}'_{\text{bg}}(\xi; x, t)\mathbf{c} = \mathbf{U}_{\text{bg}}^{\text{in}'}(\xi; x, t)\mathbf{c}$ . In general,

$$(3.42) \quad \mathbf{s}(x, t) = -i\xi \mathcal{S}(\xi; x, t)\sigma_3\mathbf{c} + i\mathcal{S}(\xi; x, t)\sigma_2\mathbf{c} + \mathcal{C}(\xi; x, t)\mathbf{c},$$

where

$$(3.43) \quad \mathcal{S}(\xi; x, t) := \frac{\sin(\rho(\xi)(x + \xi t))}{\rho(\xi)} \quad \text{and} \quad \mathcal{C}(\xi; x, t) := \cos(\rho(\xi)(x + \xi t))$$

are well-defined and analytic for  $\xi \in D_0$ . Differentiating with respect to  $\xi$  gives

$$(3.44) \quad \begin{aligned} \mathbf{s}'(x, t) &= -i\xi \mathcal{S}'(\xi; x, t)\sigma_3\mathbf{c} - i\mathcal{S}(\xi; x, t)\sigma_3\mathbf{c} \\ &\quad + i\mathcal{S}'(\xi; x, t)\sigma_2\mathbf{c} + \mathcal{C}'(\xi; x, t)\mathbf{c}. \end{aligned}$$

From (3.42) and (3.44) we calculate  $N(x, t) := \mathbf{s}(x, t)^\dagger \mathbf{s}(x, t)$  as

$$(3.45) \quad \begin{aligned} N(x, t) &= (\mathbf{c}^\dagger \mathbf{c})((1 + |\xi|^2)|\mathcal{S}(\xi; x, t)|^2 + |\mathcal{C}(\xi; x, t)|^2) \\ &\quad + 2\beta(\mathbf{c}^\dagger \sigma_1 \mathbf{c})|\mathcal{S}(\xi; x, t)|^2 \\ &\quad + 2(\mathbf{c}^\dagger \sigma_3 \mathbf{c}) \operatorname{Im}\{\xi \mathcal{S}(\xi; x, t)\mathcal{C}(\xi; x, t)^*\} \\ &\quad - 2(\mathbf{c}^\dagger \sigma_2 \mathbf{c}) \operatorname{Im}\{\mathcal{S}(\xi; x, t)\mathcal{C}(\xi; x, t)^*\} \end{aligned}$$

and  $w(x, t) := \mathbf{s}(x, t)^\top \sigma_2 \mathbf{s}'(x, t)$  as

$$(3.46) \quad \begin{aligned} w(x, t) &= -(\mathbf{c}^\top \sigma_3 \mathbf{c})\mathcal{S}(\xi; x, t)^2 + (\mathbf{c}^\top \sigma_1 \mathbf{c})\mathcal{S}(\xi; x, t)\mathcal{C}(\xi; x, t) \\ &\quad + (\mathbf{c}^\top (i\mathbb{I} + \xi\sigma_1)\mathbf{c})(\mathcal{C}(\xi; x, t)\mathcal{S}'(\xi; x, t) - \mathcal{S}(\xi; x, t)\mathcal{C}'(\xi; x, t)). \end{aligned}$$

In terms of these, the Bäcklund transformation (3.26) takes the form ( $\psi(x, t) = \psi_{\text{bg}}(x, t) \equiv 1$  here)

$$\begin{aligned}
 \tilde{\psi}(x, t) &= \psi(x, t) \\
 (3.47) \quad &+ 8\beta \frac{\beta s_1(x, t)^2(1 - w(x, t)^*) - \beta s_2(x, t)^*{}^2(1 - w(x, t))}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2} \\
 &+ 8\beta \frac{s_1(x, t)s_2(x, t)^*N(x, t)}{4\beta^2|1 - w(x, t)|^2 + N(x, t)^2}.
 \end{aligned}$$

### Properties of the Solutions Obtained for $\xi \neq \pm i$

Suppose that  $\xi \neq \pm i$ . Then if  $\mathbf{c}$  is chosen such that  $\mathbf{c}^\top(i\mathbb{I} + \xi\sigma_1)\mathbf{c} = 0$ , the solution obtained from the Bäcklund transformation exhibits quite a different character than for  $\mathbf{c}$  in general position. Indeed, this condition removes the terms on the second line of (3.46), with the result being that all dependence on  $(x, t)$  in  $\tilde{\psi}(x, t) - \psi(x, t)$  enters through the functions  $\mathcal{S}(\xi; x, t)$  and  $\mathcal{C}(\xi; x, t)$ . In the general case, however, derivatives of these functions with respect to  $\xi$  appear, and these produce additional dependence on  $(x, t)$  via polynomial factors. For convenience, we illustrate the difference between the special and general case for the Kuznetsov-Ma family of solutions [30, 32], corresponding to choosing  $\xi = (1 + \delta)i$  for some  $\delta > 0$ . For such  $\xi$ , the condition  $\mathbf{c}^\top(i\mathbb{I} + \xi\sigma_1)\mathbf{c} = 0$  reads  $c_1^2 + c_2^2 + 2(1 + \delta)c_1c_2 = 0$ , which implies  $c_2 = c_1(-(1 + \delta) \pm \sqrt{(1 + \delta)^2 - 1})$ . Figure 3.1 compares the solutions  $\tilde{\psi}(x, t)$  obtained for  $\xi = 2i$  with  $c_1 = 1 + i$  and  $c_2 = c_1(-2 + \sqrt{3})$  (the special case, on the left) and with  $c_1 = 1$  and  $c_2 = 1 + i$  (the general case, on the right). The reader may be familiar with the fact that in the context of the traditional IST for the Cauchy problem (1.1)–(1.3), the application of a single Darboux transformation inserting poles at a conjugate pair  $\lambda = \pm i(1 + \delta)$  for  $\delta > 0$  always produces a solution periodic in  $t$  and exponentially localized in  $x$  as in the special case shown in the plot in the left-hand panel of Figure 3.1.

The reason that the special case is selected in that setting is that the Darboux transformation is applied to the Beals-Coifman solution  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  rather than to  $\mathbf{U}^{\text{in}}(\lambda; x, t)$ , and then the parameters  $(c_1, c_2)$  are chosen to make sure that the pole is inserted into one column only of the simultaneous solution matrix (that is, the Darboux transformation is designed to introduce an additional Blaschke factor  $(\lambda - \xi)/(\lambda - \xi^*)$  into the function  $a(\lambda)$ ). Recalling that working with  $\mathbf{U}^{\text{in}}(\lambda; x, t)$  rather than  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  introduces a rotation in the parameter space, the condition  $\mathbf{c}^\top(i\mathbb{I} + \xi\sigma_1)\mathbf{c} = 0$  is seen as the analogue for the Darboux transformation of the robust IST of the latter choice. Moreover, we see that for  $\mathbf{c}$  in general position, the solution obtained (see the right-hand panel of Figure 3.1) is what one would expect in the context of the traditional IST from a solution in which  $\mathbf{U}^{\text{BC}}(\lambda; x, t)$  has a conjugate pair of *double poles*; the solution resembles a nonballistic collision of identical Kuznetsov-Ma solutions.

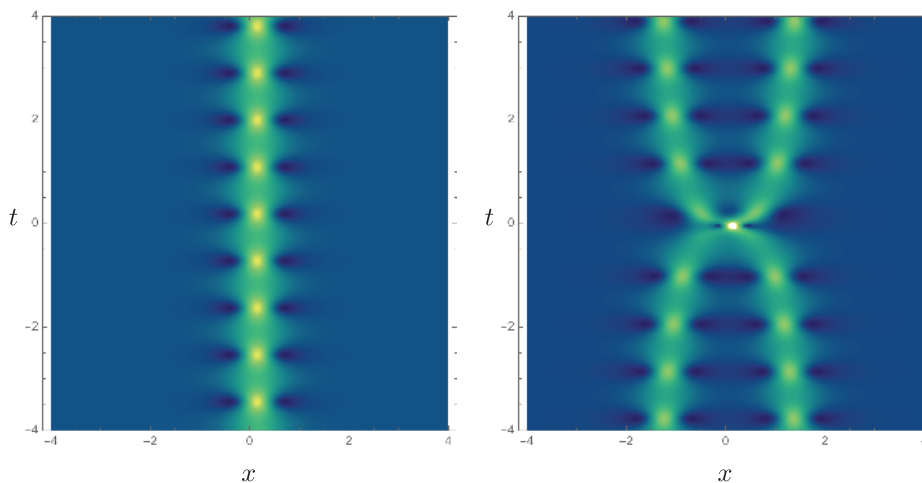


FIGURE 3.1. Left:  $|\tilde{\psi}(x, t)|$  for  $\xi = 2i, c_1 = 1 + i$ , and  $c_2 = c_1(-2 + \sqrt{3})$  for which the solution is periodic in  $t$  and exponential in  $x$ . Right:  $|\tilde{\psi}(x, t)|$  for  $\xi = 2i, c_1 = 1$ , and  $c_2 = 1 + i$  for which the solution formula includes additional terms polynomial in  $(x, t)$ .

Under the special condition that  $\mathbf{c}^T(i\mathbb{I} + \xi\sigma_1)\mathbf{c} = 0$ , all  $(x, t)$ -dependence in the solution enters via the functions  $\mathcal{S}(\xi; x, t)$  and  $\mathcal{C}(\xi; x, t)$ . These functions are periodic and bounded in the real part of  $\rho(\xi)(x + \xi t)$ , but they grow exponentially in all directions for which  $\text{Im}\{\rho(\xi)(x + \xi t)\}$  is unbounded. The only values of  $\xi \in D_0 \setminus \mathbb{R}$  for which  $\rho(\xi)$  is real and nonzero are the values  $\xi = i\delta$  with  $\delta \in (-1, 0) \cup (0, 1)$  and only for these values is the solution periodic rather than exhibiting exponential decay to the background in the  $x$ -direction. This case corresponds to the so-called *Akhmediev breather* solutions [3], and such solutions clearly do not satisfy the boundary condition (1.2). Nonetheless, they have a Riemann-Hilbert representation obtained by modifying the jump matrices  $\mathbf{V}_{\text{bg}}^\pm(\lambda) = \mathbf{E}(\lambda)^{\pm 1}$  for  $\lambda \in \Sigma_0$  by the gauge transformation matrix  $\mathbf{G}(\lambda; 0, 0)$  as indicated in (3.36). Although the solution is no longer periodic in the  $x$ -direction for  $\xi = i\delta$  with  $\delta \in (-1, 0) \cup (0, 1)$  if  $\mathbf{c}$  is in general position ( $\mathbf{c}^T(i\mathbb{I} + \xi\sigma_1)\mathbf{c} = i(c_1^2 + c_2^2 + 2\delta c_1 c_2) \neq 0$ ), still the boundary condition (1.2) is not satisfied. See Figure 3.2.

**Properties of the Solutions Obtained for  $\xi = i$**

Now we return to the previously excluded case that  $\xi = \pm i$  (we take  $\xi = i$  to be precise). The reason for considering this case separately is *not* that the Darboux transformation method described in Section 3.2 requires any modification,<sup>6</sup>

<sup>6</sup>This is a key property of the Darboux transformation in the setting of the robust IST that distinguishes it from other methods in the literature.

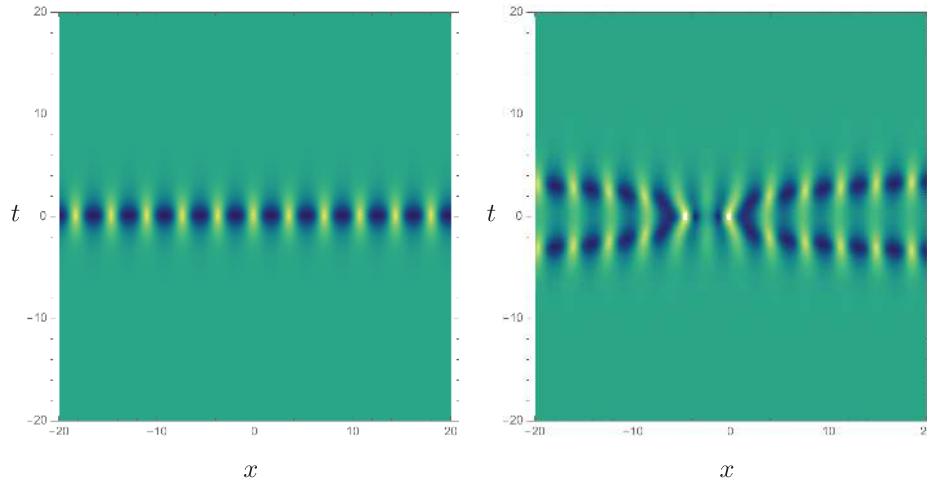


FIGURE 3.2. Left:  $|\tilde{\psi}(x, t)|$  for  $\xi = \frac{1}{2}i$ ,  $c_1 = 1$ , and  $c_2 = c_1(-\frac{1}{2} + \frac{1}{2}i\sqrt{3})$  for which the solution is periodic in  $x$  and exponential in  $t$ . Right:  $|\tilde{\psi}(x, t)|$  for  $\xi = \frac{1}{2}i$ ,  $c_1 = 1$ , and  $c_2 = 0$  for which the solution formula includes additional terms polynomial in  $(x, t)$ .

but rather that the nature of the resulting solution is quite different from the general case. The reason for the difference emerges upon evaluation of the analytic functions  $\mathcal{S}(\xi; x, t)$  and  $\mathcal{C}(\xi; x, t)$  and their derivatives at  $\xi = i$ , which requires some version of l’Hôpital’s rule. Equivalently, by Taylor expansion of  $\sin(\theta)$  and  $\cos(\theta)$  about  $\theta = 0$  one finds

$$(3.48) \quad \begin{aligned} \mathcal{S}(i; x, t) &= x + it, & \mathcal{S}'(i; x, t) &= t - \frac{1}{3}i(x + it)^3, \\ \mathcal{C}(i; x, t) &= 1, & \mathcal{C}'(i; x, t) &= -i(x + it)^2. \end{aligned}$$

Therefore, in the special case that  $\xi = i$ , the formulae (3.42) and (3.44) become

$$(3.49) \quad \mathbf{s}(x, t) = \begin{bmatrix} c_1 + (c_1 + c_2)(x + it) \\ c_2 - (c_1 + c_2)(x + it) \end{bmatrix}, \quad \xi = i,$$

and

$$(3.50) \quad \mathbf{s}'(x, t) = \begin{bmatrix} (t - \frac{1}{3}i(x + it)^3)(c_1 + c_2) - i(x + it)c_1 - i(x + it)^2c_1 \\ -(t - \frac{1}{3}i(x + it)^3)(c_1 + c_2) + i(x + it)c_2 - i(x + it)^2c_2 \end{bmatrix}, \quad \xi = i.$$

Therefore,  $N(x, t) = \mathbf{s}(x, t)^\dagger \mathbf{s}(x, t)$  is

$$(3.51) \quad \begin{aligned} N(x, t) &= |c_1|^2 + |c_2|^2 + 2\operatorname{Re}\{(c_1 - c_2)^*(c_1 + c_2)(x + it)\} \\ &\quad + 2|c_1 + c_2|^2|x + it|^2, \quad \xi = i, \end{aligned}$$

and  $w(x, t) = \mathbf{s}(x, t)^\top \sigma_2 \mathbf{s}'(x, t)$  is

$$(3.52) \quad w(x, t) = -\frac{2}{3}(c_1 + c_2)^2(x + it)^3 - (c_1^2 - c_2^2)(x + it)^2 + 2c_1c_2(x + it) + it(c_1 + c_2)^2, \quad \xi = i.$$

Noting that for large  $(x, t)$ , the dominant terms in  $\mathbf{s}(x, t)$ ,  $\mathbf{s}'(x, t)$ ,  $N(x, t)$ , and  $w(x, t)$  are all proportional to  $c_1 + c_2$ , it is clear that the solution  $\tilde{\psi}(x, t)$  obtained from the Bäcklund transformation (3.47) has a different character if  $c_1 + c_2 = 0$  than otherwise. Indeed, if  $c_1 = -c_2 = c$ , then

$$(3.53) \quad \mathbf{s}(x, t) = \begin{bmatrix} c \\ -c \end{bmatrix}, \quad N(x, t) = 2|c|^2, \quad \text{and} \quad w(x, t) = -2c^2(x + it),$$

$$\xi = i, \quad c_1 = -c_2 = c.$$

Using these formulae in (3.47) (also with  $\beta = \text{Im}\{\xi\} = 1$ ) then yields the Peregrine breather solution [35]  $\tilde{\psi}(x, t) = \psi_P(x, t)$  (given in (1.5); see also Figure 1.1) with parameters

$$(3.54) \quad x_0 = -\frac{\text{Re}\{c^2\}}{2|c|^4} \quad \text{and} \quad t_0 = \frac{\text{Im}\{c^2\}}{2|c|^4}.$$

For finite  $c$ , the peak of the breather can be placed anywhere except the origin (this coincides with the normalization point in Proposition 2.1). If it is desired to place the breather exactly at the origin, we simply use the limiting case of the Darboux transformation described in Remark 3.4 with homogeneous coordinates  $\mathbf{c}_\infty = [c : -c]^\top \in \mathbb{C}\mathbb{P}^1$ .

If  $c_1 + c_2 \neq 0$ , the solution for  $\xi = i$  is more complicated. We may introduce complex parameters  $c$  and  $\delta$  such that  $c_1 = c + \frac{1}{2}\delta$  and  $c_2 = -c + \frac{1}{2}\delta$  so that  $\delta = c_1 + c_2$  measures the deviation from the Peregrine case. Taking  $c \in \mathbb{C} \setminus \{0\}$  fixed and  $\delta$  small, examination of the expressions (3.51)–(3.52) suggests that a natural scaling of  $(x, t) \in \mathbb{R}^2$  is to set  $x = \bar{x}/|\delta|$  and  $t = \bar{t}/|\delta|$ . If  $(\bar{x}, \bar{t}) \in \mathbb{R}^2$  is also fixed, then as  $\delta \rightarrow 0$ ,

$$(3.55) \quad N(x, t) = 2|c|^2 + 4 \text{Re}\{c^* e^{i \arg(\delta)}(\bar{x} + i\bar{t})\} + 2(\bar{x}^2 + \bar{t}^2) + O(\delta)$$

and

$$(3.56) \quad w(x, t) = |\delta|^{-1} \left[ -\frac{2}{3} e^{2i \arg(\delta)}(\bar{x} + i\bar{t})^3 - 2c e^{i \arg(\delta)}(\bar{x} + i\bar{t})^2 - 2c^2(\bar{x} + i\bar{t}) \right] + O(\delta).$$

Also,

$$(3.57) \quad \mathbf{s}(x, t) = \begin{bmatrix} c + e^{i \arg(\delta)}(\bar{x} + i\bar{t}) \\ -c - e^{i \arg(\delta)}(\bar{x} + i\bar{t}) \end{bmatrix} + O(\delta).$$

From the Bäcklund transformation formula (3.47) we then see that  $\tilde{\psi}(x, t) \approx 1$  unless the leading term in  $w(x, t)$  proportional to  $|\delta|^{-1}$  is canceled. These terms



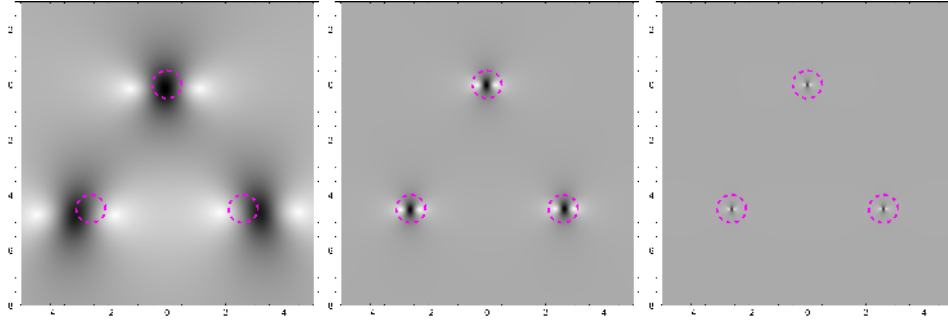


FIGURE 3.3.  $|\widetilde{\psi}(|\delta|^{-1}\bar{x}, |\delta|^{-1}\bar{t})|$  plotted in the  $\bar{x}$  (horizontal) and  $\bar{t}$  (vertical) plane for a solution obtained by a single application of the Darboux transformation with the parameters  $c_1 = c + \delta/2$  and  $c_2 = -c + \delta/2$ . Dark color indicates peaks and light color indicates troughs. Here  $c = 3$  is fixed, and  $\delta$  is taken to be  $\delta = 1.5 e^{-i\pi/2}/k$ ,  $k = 1, 4, 16$ , respectively, from left to right. The centers of the dashed magenta circles indicate asymptotically valid (in the limit of large separation or small  $\delta$ ) predictions for the locations of the three peaks. Note the shift in the vertical axis: the uppermost dashed circle is centered at the origin.

constitute a cubic equation for  $\bar{x} + i\bar{t}$ , one root of which is  $\bar{x} + i\bar{t} = 0$  and the other two of which are

$$(3.58) \quad \bar{x} + i\bar{t} = \frac{1}{2}c e^{-i \arg(\delta)}(-3 \pm i\sqrt{3}).$$

Therefore, when  $c \neq 0$  and  $\delta$  is small, the solution  $\widetilde{\psi}(x, t)$  is very close to the background solution unless  $(x, t)$  lies in  $O(1)$  neighborhoods of the three points  $x_0 + it_0 = 0, \frac{1}{2}c\delta^{-1}(-3 \pm i\sqrt{3})$ , which are easily seen to be the vertices of a large equilateral triangle of side length  $|c|\sqrt{3}/|\delta|$ . Near each of these three points, the solution resembles the Peregrine solution (1.5) located near  $(x_0, t_0)$ . See Figure 3.3 for convergence of peak locations to these points in the  $(\bar{x}, \bar{t})$ -plane as  $\delta \rightarrow 0$ .

The case when  $c = 0$  must be considered separately. Taking  $c_1 = c_2 = \frac{1}{2}\delta$  and  $|\delta| \ll 1$  again yields a configuration of three peaks each of which resembles the Peregrine solution; however, in this case, a dominant balance argument suggests scaling the variables  $(x, t) \in \mathbb{R}^2$  differently:  $x = \bar{x}/|\delta|^{2/3}$  and  $t = \bar{t}/|\delta|^{2/3}$ . If  $(\bar{x}, \bar{t}) \in \mathbb{R}^2$  is fixed, then as  $\delta \rightarrow 0$ ,

$$(3.59) \quad N(x, t) = 2e^{2i \arg(\delta)}|\delta|^{2/3} + O(|\delta|^2)$$

and

$$(3.60) \quad w(x, t) = -\frac{2}{3}e^{2i \arg(\delta)}(\bar{x} + i\bar{t})^3 + O(|\delta|^{4/3}).$$

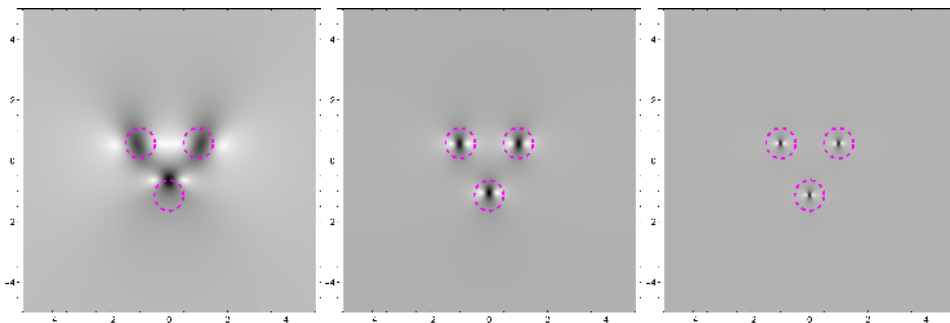


FIGURE 3.4.  $|\tilde{\psi}(|\delta|^{-2/3}\bar{x}, |\delta|^{-2/3}\bar{t})|$  plotted in the  $\bar{x}$  (horizontal) and  $\bar{t}$  (vertical) plane for a solution obtained by a single application of the Darboux transformation with the parameters  $c_1 = \delta/2$  and  $c_2 = -\delta/2$ . Dark color indicates peaks and light color indicates troughs. Here  $\delta$  is taken to be  $\delta = 0.75e^{3i\pi/4}/k$ ,  $k = 1, 4, 16$ , respectively from left to right. The centers of the dashed magenta circles indicate asymptotically valid (in the limit of large separation or small  $\delta$ ) predictions for the locations of the three peaks.

Also,

$$(3.61) \quad \mathbf{s}(x, t) = |\delta|^{1/3} e^{i \arg(\delta)} (\bar{x} + i\bar{t}) \begin{bmatrix} 1 \\ -1 \end{bmatrix} + O(|\delta|).$$

Again from the Bäcklund transform formula (3.47) we then see that  $\tilde{\psi}(x, t) \approx 1$  unless the leading term in  $|w(x, t) - 1|$ , which is proportional to 1, is canceled. This condition leaves us with the following cubic equation for  $(\bar{x} + i\bar{t})$ :

$$(3.62) \quad (\bar{x} + i\bar{t})^3 = -\frac{3}{2} e^{-2i \arg(\delta)}.$$

Therefore, when  $\delta$  is small, the solution  $\tilde{\psi}(x, t)$  is very close to the background solution unless  $(x, t)$  lies in  $O(1)$  neighborhoods of the three points  $(x_0, t_0)$  determined by  $(x_0 + it_0)^3 = -\frac{3}{2}\delta^{-2}$ , which are again vertices of a large equilateral triangle; however, now the triangle is centered at the origin and has an asymptotically smaller side length of  $2^{-1/3}3^{5/6}|\delta|^{-2/3}$  in the limit  $\delta \rightarrow 0$ . See Figure 3.4.

Another interesting limit corresponds to holding  $c = (c_1 - c_2)/2$  fixed and letting  $\delta = c_1 + c_2$  become large. This brings the three vertices of the equilateral triangle, where the (approximate) Peregrine breathers are placed, in toward the origin, and suggests that the limit may result in the fusion of the three peaks into a single structure. This limit is another application of the generalization described in Remark 3.4, in which the parameter is given by the homogeneous coordinates

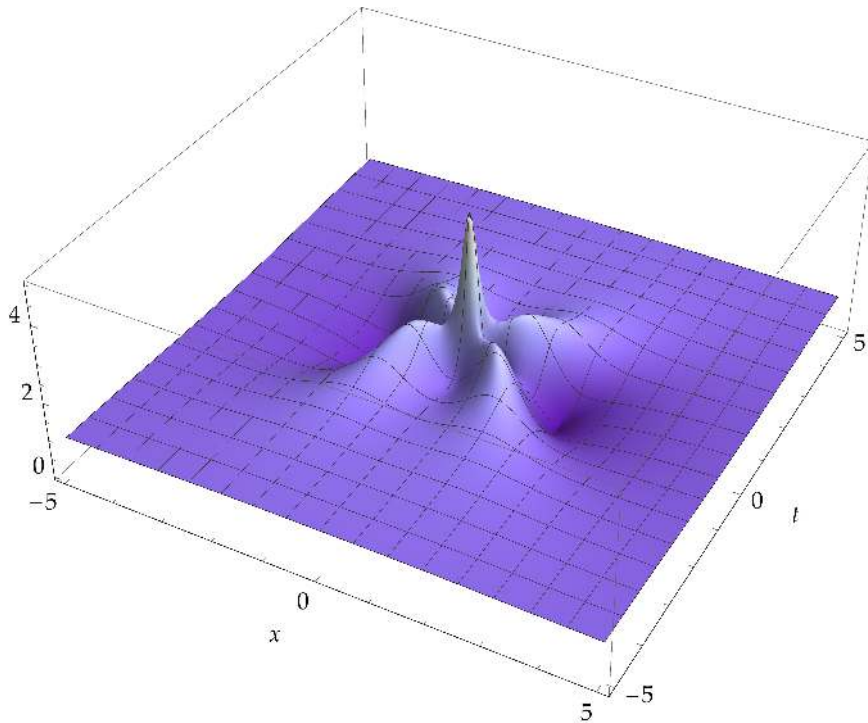


FIGURE 3.5.  $|\tilde{\psi}(x, t)|$  corresponding to a single application of the Darboux transformation to the background with parameters  $\xi = i$  and “ $\mathbf{c} = \infty$ ” in the sense described in Remark 3.4 with parameter  $\mathbf{c}_\infty = [1 : 1]^T \in \mathbb{CP}^1$ .

$\mathbf{c}_\infty = [1 : 1]^T \in \mathbb{CP}^1$ . The corresponding solution is given by

$$\begin{aligned}
 \tilde{\psi}(x, t) &= 1 + 12 \frac{P(x, t)}{Q(x, t)} \quad \text{with} \\
 (3.63) \quad P(x, t) &= -32it^5 - 80t^4 - 16it^3(4x^2 + 1) - 24t^2(4x^2 + 3) \\
 &\quad - 2it(16x^4 - 24x^2 - 15) - 16x^4 - 24x^2 + 3, \\
 Q(x, t) &= 64t^6 + 48t^4(4x^2 + 9) + 12t^2(16x^4 - 24x^2 + 33) \\
 &\quad + 64x^6 + 48x^4 + 108x^2 + 9,
 \end{aligned}$$

and it is plotted in Figure 3.5. It is an example of a “higher-order” rogue wave solution of (1.1).

A useful property of the matrix coefficient  $\mathbf{Y}(x, t)$  in the partial fraction expansion of the composite gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  constructed from the background solution  $\mathbf{U}_{\text{bg}}(\lambda; x, t)$  for  $\xi = i$  is the following.

LEMMA 3.6. Let  $\mathbf{Y}(x, t)$  be defined by (3.23) for  $\beta = 1$  with  $\mathbf{s}(x, t)$ ,  $N(x, t)$ , and  $w(x, t)$  being given by (3.49), (3.51), and (3.52), respectively. Then for each  $t \in \mathbb{R}$ ,

$$(3.64) \quad \mathbf{Y}(x, t) = \frac{i\tau}{x} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} + O(x^{-2}), \quad x \rightarrow \pm\infty,$$

where  $\tau = \frac{1}{4}$  if  $c_1 + c_2 = 0$  and  $\tau = \frac{3}{2}$  otherwise. The same conclusion holds for  $\mathbf{Y}_\infty(x, t)$  defined by (3.37) where  $\tau = \frac{1}{4}$  if  $\mathbf{c}_\infty = [c : -c]^T \in \mathbb{C}\mathbb{P}^1$  and  $\tau = \frac{3}{2}$  otherwise.

PROOF. If  $c_1 + c_2 \neq 0$ , then  $\mathbf{s}(x, t) = (c_1 + c_2)x[1 \ -1]^T + O(1)$ ,  $N(x, t) = 2|c_1 + c_2|^2x^2 + O(x)$ , and  $w(x, t) = -\frac{2}{3}(c_1 + c_2)^2x^3 + O(x^2)$  as  $x \rightarrow \pm\infty$ . On the other hand, if  $c_1 + c_2 = 0$ , then  $\mathbf{s}(x, t) = [c \ -c]^T$ ,  $N(x, t) = 2|c|^2$ , and  $w(x, t) = -2c^2x + O(1)$  as  $x \rightarrow \pm\infty$ . Using these in (3.23) and (3.37) completes the proof.  $\square$

This lemma has two related consequences. First, for any  $\mathbf{c} \in \mathbb{C}^2$ ,

$$(3.65) \quad Y_{12}(x, t) - Y_{21}(x, t)^* = O(x^{-2}), \quad x \rightarrow \pm\infty.$$

Thus, according to (3.26), the transformed potential  $\tilde{\psi}(x, t)$  is an  $L^1$ -perturbation of the background field  $\psi(x, t) = \psi_{\text{bg}}(x, t) \equiv 1$ . Second, for any  $\mathbf{c} \in \mathbb{C}^2$ , the gauge transformation matrix  $\mathbf{G}(\lambda; x, t)$  defined by (3.14) for  $\xi = i$  tends to the identity matrix as  $x \rightarrow \pm\infty$  and hence it preserves the leading-order behavior in the asymptotic expansion as  $x \rightarrow \pm\infty$  when applied as a prefactor to a matrix function of  $x$ . This implies that for all  $\lambda \in \Gamma$ ,  $\mathbf{G}(\lambda; x, 0)\mathbf{J}_{\text{bg}}^\pm(\lambda; x, 0)$  are precisely the Jost solution matrices associated with the transformed potential  $\tilde{\psi}(x, 0)$ . Since  $\mathbf{J}_{\text{bg}}^+(\lambda; x, 0) = \mathbf{J}_{\text{bg}}^-(\lambda; x, 0) = \mathbf{U}_{\text{bg}}(\lambda; x, 0)$ , we therefore arrive at the following result.

COROLLARY 3.7. All solutions obtained from the background from a single application of the Darboux transformation described in Section 3.2 for  $\xi = i$  have the same scattering matrix as the background potential:  $\mathbf{S}(\lambda; t) = \mathbb{I}$  for all  $\lambda \in \Gamma \setminus \{i, -i\}$ .

In particular, this holds for the Peregrine solution  $\psi_P(x, t)$  defined in (1.5). On the other hand, such a result is not true for  $\xi \neq \pm i$ , and the Akhmediev breather case shows that it may not even be possible to define the scattering matrix after the application of a Darboux transformation.

Remark 3.8. By definition (see (3.23)),  $\mathbf{Y}(x, t)$  always annihilates  $\mathbf{s}(x, t)$ . According to (3.53), in the case that  $c_1 + c_2 = 0$  giving rise to the Peregrine solution, the kernel of  $\mathbf{Y}(x, t)$  contains the span of  $[1 \ -1]^T$ . Now, the Beals-Coifman matrix for the background solution is defined by  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) = \mathbf{E}(\lambda) e^{-i\rho(x+\lambda t)\sigma_3}$  (cf.

(1.10)), and it satisfies

$$(3.66) \quad \mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) = n(\lambda) \left( \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} - i\rho(\lambda)(\sigma_1 + (x + it)\sigma_3) + O(\lambda - i) \right), \quad \lambda \rightarrow i.$$

Therefore, applying the gauge transformation  $\mathbf{G}(\lambda; x, t)$  to map  $\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t)$  into its Peregrine analogue  $\mathbf{U}_{\text{P}}^{\text{BC}}(\lambda; x, t)$ , the leading term is explicitly canceled:

$$(3.67) \quad \begin{aligned} \mathbf{U}_{\text{P}}^{\text{BC}}(\lambda; x, t) &= \mathbf{G}(\lambda; x, t)\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) \\ &= (\mathbf{Y}(x, t)(\lambda - i)^{-1} + O(1))\mathbf{U}_{\text{bg}}^{\text{BC}}(\lambda; x, t) \\ &= -in(\lambda)\rho(\lambda)(\lambda - i)^{-1}\mathbf{Y}(x, t)(\sigma_1 + (x + it)\sigma_3) + O(n(\lambda)), \quad \lambda \rightarrow i. \end{aligned}$$

Since  $n(\lambda)\rho(\lambda) = O((\lambda - i)^{1/4})$ , we see that for the Peregrine solution, the Beals-Coifman fundamental solution matrix exhibits a  $(\lambda - i)^{-3/4}$  singularity in all four matrix entries as  $\lambda \rightarrow i$ . Without the condition  $c_1 + c_2 = 0$ , the cancellation does not occur, and therefore more generally the Beals-Coifman matrix has a  $(\lambda - i)^{-5/4}$  singularity after one application of the Darboux transformation to the background solution. The same growth rate estimates apply to  $\mathbf{M}^{\text{BC}}(\lambda; x, t)$ ; in the above expansions one need only omit the term  $(x + it)\sigma_3$  that comes from expanding the exponential factor  $e^{-i\rho(\lambda)(x + it)\sigma_3}$ .

### 3.4 Iteration. Riemann-Hilbert Representations for High-Order Rogue Waves

Since according to Theorem 3.5 the result of applying a Darboux transformation in the setting of the robust IST is to transform one Riemann-Hilbert problem into another one of the same form and at the same time to transform the *solution* of the first problem into that of the second, the procedure can be repeated. The data  $(\xi, \mathbf{c})$  associated with each iteration can be related or unrelated to that of the previous step, and the basic procedure remains the same. In this section, we show how to iterate the Darboux transformation an arbitrary number of times for certain data chosen to produce interesting solutions of the focusing NLS equation (1.1) and also for which the resulting jump matrix after multiple iterations can be explicitly determined in closed form.

The solutions we wish to obtain are those commonly referred to as the higher-order rogue wave solutions. The simplest such solutions have already been obtained in Section 3.3 (and in particular the subsection “Properties of the Solutions Obtained for  $\xi = \pm i$ ”), namely the Peregrine solution (1.5) (a “first-order” rogue wave; see Figure 1.1) and another solution obtained by choosing parameters so as to fuse together three copies of the elementary Peregrine solution at a single point (a “second-order” rogue wave; see Figure 3.5). Let us denote these solutions respectively as  $\psi_{\text{P}_1}(x, t)$  and  $\psi_{\text{P}_2}(x, t)$ .

Both of these solutions were obtained by a single iteration of the Darboux transformation applied to the background solution  $U_{bg}(\lambda; x, t)$  with pole location  $\xi = i$  and different auxiliary data taken in the limiting sense of Remark 3.4. Namely,  $\psi_{P_1}(x, t)$  was generated from the data  $\mathbf{c}_\infty = [1 : -1]^T \in \mathbb{C}\mathbb{P}^1$ , while  $\psi_{P_2}(x, t)$  was generated from  $\mathbf{c}_\infty = [1 : 1]^T \in \mathbb{C}\mathbb{P}^1$ . To generate rogue wave solutions of arbitrary order, we simply apply these basic Darboux transformations iteratively.

**DEFINITION 3.9.** Let  $\mathcal{D}_o$  denote the Darboux transformation associated with the data ( $\xi = i, \mathbf{c}_\infty = [1 : -1]^T$ ), and let  $\mathcal{D}_e$  denote the Darboux transformation associated with the data ( $\xi = i, \mathbf{c}_\infty = [1 : 1]^T$ ). Then the rogue wave solutions of order  $2n - 1$  and  $2n$  are the solutions  $\psi = \psi_{P_{2n-1}}(x, t)$  and  $\psi = \psi_{P_{2n}}(x, t)$ , respectively, obtained from the iterated Darboux/Bäcklund transformations  $\mathcal{D}_o^n$  and  $\mathcal{D}_e^n$ , respectively, applied to the background.

Qualitatively speaking,  $\psi_{P_k}(x, t)$  represents a fusion or nonlinear superposition of  $2k - 1$  copies of the elementary Peregrine solution  $\psi_P(x, t)$  at the same point, namely  $(x, t) = (0, 0)$ . It has recently been proven [40] that for any value of  $k$ , the maximum amplitude of this solution is  $\max_{(x,t) \in \mathbb{R}^2} |\psi_{P_k}(x, t)| = |\psi_{P_k}(0, 0)| = 2k + 1$ .

According to Theorem 3.5, the effect of each application of  $\mathcal{D}_o$  or  $\mathcal{D}_e$  is to produce a new factor in the jump matrix on the upper/lower semicircles  $\Sigma_+/\Sigma_-$  of the jump contour, and at each iteration the factor is calculated by evaluating the gauge transformation  $\mathbf{G}(\lambda; x, t) = \mathbb{I} + (\lambda - i)^{-1} \mathbf{Y}_\infty(x, t) + (\lambda + i)^{-1} \mathbf{Z}_\infty(x, t)$  at  $(x, t) = (0, 0)$  after computing  $\mathbf{Y}_\infty(x, t)$  and  $\mathbf{Z}_\infty(x, t)$  from the fundamental solution matrix obtained from the previous iteration. Let  $\mathbf{M}_o^{[n]}(\lambda; x, t)$  and  $\mathbf{M}_e^{[n]}(\lambda; x, t)$  denote the solution matrices of Riemann-Hilbert Problem 1 after  $n$  applications of  $\mathcal{D}_o$  and  $\mathcal{D}_e$ , respectively. Thus  $\mathbf{M}_o^{[0]}(\lambda; x, t) = \mathbf{M}_e^{[0]}(\lambda; x, t)$  is the solution of Riemann-Hilbert Problem 1 formulated for the background solution  $\psi = \psi_{bg}(x, t) \equiv 1$ . Let  $\mathbf{G}_o^{[n]}(\lambda; x, t)$  and  $\mathbf{G}_e^{[n]}(\lambda; x, t)$  denote the gauge matrices calculated for the application of  $\mathcal{D}_o$  to  $\mathbf{U}_o^{[n]}(\lambda; x, t) := \mathbf{M}_o^{[n]}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda)t\sigma_3}$  and of  $\mathcal{D}_e$  to  $\mathbf{U}_e^{[n]}(\lambda; x, t) := \mathbf{M}_e^{[n]}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda)t\sigma_3}$ , respectively. Thus,

$$(3.68) \quad \mathbf{U}_{o/e}^{[n]}(\lambda; x, t) = \mathbf{G}_{o/e}^{[n-1]}(\lambda; x, t) \cdots \mathbf{G}_{o/e}^{[0]}(\lambda; x, t) \mathbf{U}_{bg}(\lambda; x, t) \begin{cases} \mathbb{I}, & \lambda \in D_+ \cup D_-, \\ \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)^{-1} \cdots \mathbf{G}_{o/e}^{[n-1]}(\lambda; 0, 0)^{-1}, & \lambda \in D_0, \end{cases}$$

and  $\mathbf{M}_{o/e}^{[n]}(\lambda; x, t) = \mathbf{U}_{o/e}^{[n]}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda)t\sigma_3}$ . The Riemann-Hilbert problem whose unique solution is  $\mathbf{M}_{o/e}^{[n]}(\lambda; x, t)$  is then the following.

*Riemann-Hilbert Problem 2.* Let  $n = 0, 1, 2, \dots$ , and seek a  $2 \times 2$  matrix function  $\mathbf{M}(\lambda; x, t) = \mathbf{M}_{o/e}^{[n]}(\lambda; x, t)$  that has the following properties:

- **ANALYTICITY:**  $\mathbf{M}(\lambda; x, t)$  is analytic for  $\lambda$  in the exterior of the disk  $D_0$  and for  $\lambda \in D_0 \setminus \Sigma_c$ .
- **JUMP CONDITION:**  $\mathbf{M}(\lambda; x, t)$  takes continuous boundary values on  $\Sigma_0 \cup \Sigma_c$  denoted by  $\mathbf{M}_\pm(\lambda; x, t)$ , and they are related by a jump condition of the form  $\mathbf{M}_+(\lambda; x, t) = \mathbf{M}_-(\lambda; x, t)\mathbf{V}_{o/e}^{[n]}(\lambda; x, t)$  for  $\lambda \in \Sigma \cup \Sigma_c$ , where

$$(3.69) \quad \mathbf{V}_{o/e}^{[n]}(\lambda; x, t) := e^{2i\rho+(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \Sigma_c,$$

$$(3.70) \quad \begin{aligned} \mathbf{V}_{o/e}^{[n]}(\lambda; x, t) &:= e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{G}_{o/e}^{[n-1]}(\lambda; 0, 0) \cdots \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0) \\ &\cdot \mathbf{E}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \Sigma_+, \end{aligned}$$

and

$$(3.71) \quad \begin{aligned} \mathbf{V}_{o/e}^{[n]}(\lambda; x, t) &:= e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{E}(\lambda)^{-1} \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)^{-1} \cdots \\ &\cdot \mathbf{G}_{o/e}^{[n-1]}(\lambda; 0, 0)^{-1} e^{i\rho(\lambda)(x+\lambda t)\sigma_3}, \quad \lambda \in \Sigma_-. \end{aligned}$$

- **NORMALIZATION:**  $\lim_{\lambda \rightarrow \infty} \mathbf{M}(\lambda; x, t) = \mathbb{I}$ .

Note that this Riemann-Hilbert problem, like that for the background solution  $\psi_{\text{bg}}(x, t) \equiv 1$ , has no jump across the real axis. The rogue waves themselves are obtained from the solution in the usual way:

$$(3.72) \quad \begin{aligned} \psi_{P_{2n-1}}(x, t) &= 2i \lim_{\lambda \rightarrow \infty} \lambda M_{o,12}^{[n]}(\lambda; x, t), \\ \psi_{P_{2n}}(x, t) &= 2i \lim_{\lambda \rightarrow \infty} \lambda M_{e,12}^{[n]}(\lambda; x, t), \quad n = 1, 2, 3, \dots \end{aligned}$$

Although the gauge transformation matrix  $\mathbf{G}_{o/e}^{[k]}(\lambda; x, t)$  is not generally equal to its predecessor  $\mathbf{G}_{o/e}^{[k-1]}(\lambda; x, t)$  because the former has to be calculated using the values of the latter near  $\xi = i$ , a remarkable simplification occurs for  $(x, t) = (0, 0)$ , and this is enough to make the evaluation of the jump matrix in Riemann-Hilbert Problem 2 completely explicit.

**PROPOSITION 3.10.** *For any integer  $n \geq 0$ ,  $\mathbf{G}_{o/e}^{[n]}(\lambda; 0, 0) = \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)$ .*

**PROOF.** Let  $\mathbf{Y}_{\infty, o/e}^{[n]}(x, t)$  denote the matrix defined by (3.37) for the data  $\xi = i$  and  $\mathbf{c}_\infty = [1 : -1]^T \in \mathbb{C}\mathbb{P}^1$  (for  $\mathbf{Y}_{\infty, o}^{[n]}(x, t)$ ) or  $\mathbf{c}_\infty = [1 : 1]^T \in \mathbb{C}\mathbb{P}^1$  (for  $\mathbf{Y}_{\infty, e}^{[n]}(x, t)$ ). For any integer  $n \geq 0$ , the  $(x, t)$ -dependence of  $\mathbf{Y}_{\infty, o/e}^{[n]}(\lambda; x, t)$  is encoded only via the quantities

$$(3.73) \quad \begin{aligned} \mathbf{s}_\infty(i; x, t) &= \mathbf{U}_{o/e}^{[n]}(i; x, t)\mathbf{c}_\infty, \\ N_\infty(x, t) &= \|\mathbf{s}(i; x, t)\|^2 = \mathbf{c}_\infty^\dagger \mathbf{U}_{o/e}^{[n]}(i; x, t)^\dagger \mathbf{U}_{o/e}^{[n]}(i; x, t)\mathbf{c}_\infty, \\ w_\infty(x, t) &= \mathbf{c}_\infty^\top \mathbf{U}_{o/e}^{[n]}(i; x, t)^\top \sigma_2 \mathbf{U}_{o/e}^{[n]'}(i; x, t)\mathbf{c}_\infty. \end{aligned}$$

Recall that by definition  $\mathbf{U}_{o/e}^{[n]}(\lambda; 0, 0) = \mathbb{I}$  for all  $\lambda \in D_0$ , and therefore we also have  $\mathbf{U}_{o/e}^{[n]'}(i; 0, 0) = \mathbf{0}$ . Thus, evaluating the quantities above at  $(x, t) = (0, 0)$  gives

$$(3.74) \quad \mathbf{s}_\infty(i; 0, 0) = \mathbf{c}_\infty, \quad N_\infty(0, 0) = \mathbf{c}_\infty^\dagger \mathbf{c}_\infty, \quad w_\infty(0, 0) = 0,$$

independent of  $n$ , and these coincide with the values of  $\mathbf{s}_\infty(i; 0, 0)$ ,  $N_\infty(0, 0)$ , and  $w_\infty(0, 0)$  for the gauge transformation matrix  $\mathbf{G}_{o/e}^{[0]}(\lambda; x, t)$  built from the Riemann-Hilbert matrix  $\mathbf{U}_{o/e}^{[0]}(\lambda; x, t) = \mathbf{U}_{\text{bg}}(\lambda; x, t)$  for the background field. This proves the claim.  $\square$

**COROLLARY 3.11.** *The product of factors that appear in the jump matrices in (3.70)–(3.71) is given by*

$$(3.75) \quad \mathbf{G}_{o/e}^{[n-1]}(\lambda; 0, 0) \cdots \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0) = \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)^n,$$

where

$$(3.76) \quad \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0) = \mathbb{I} + \frac{\mathbf{Y}_{\infty,o/e}^{[0]}(0, 0)}{\lambda - i} + \frac{\mathbf{Z}_{\infty,o/e}^{[0]}(0, 0)}{\lambda + i}$$

in which

$$(3.77) \quad \mathbf{Y}_{\infty,o}^{[0]}(0, 0) = -\mathbf{Z}_{\infty,e}^{[0]}(0, 0) = \mathbf{H} := i \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$(3.78) \quad \mathbf{Y}_{\infty,e}^{[0]}(0, 0) = -\mathbf{Z}_{\infty,o}^{[0]}(0, 0) = \mathbf{K} := i \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

**PROOF.** The identity (3.75) follows from Proposition 3.10. It only remains to prove the formulae (3.77)–(3.78); but these follow from the definitions (3.37)–(3.38) together with (3.74) and the values of  $\mathbf{c}_\infty = [1 \ : \ -1]^\top$  (for case o) and  $\mathbf{c}_\infty = [1 \ : \ 1]^\top$  (for case e).  $\square$

*Remark 3.12.* The rogue wave solutions of (1.1) of arbitrary order have thus been encoded in the solution of a simple Riemann-Hilbert problem with a jump matrix depending explicitly on  $(x, t) \in \mathbb{R}^2$  and the order  $k$  proportional to  $n$ , which appears as an exponent. This kind of problem is likely well-suited to asymptotic analysis by the Deift-Zhou steepest descent method [15] to determine asymptotic properties of rogue waves in the limit of large order. For instance, the plots in Figures 3.6, 3.7, and 3.8 (see [17, 21, 22, 26] for similar plots) suggest the following questions:

- Can one determine the asymptotic properties of the peaks and zeros of  $\psi_k(x, t)$  for large  $k$ ? In particular, what are the asymptotics of the extreme zeros?



- Can one find the spatiotemporal pattern of rapid oscillations of  $\psi_k(x, t)$  in the limit of large  $k$  by a proper multiple-scale formula accounting for a slowly varying envelope modulating a rapidly varying carrier wave? What is the nature of the carrier wave? Is it trigonometric, elliptic, or characterized by some other special function in the limit  $k \rightarrow \infty$ ?

This is work in progress [7].

### 3.5 From Analytic to Algebraic Representations of High-Order Rogue Waves

Although Riemann-Hilbert Problem 2 is an analytical characterization of rogue waves of arbitrary order, it yields an algebraic representation as well, which is more in line with what is in the literature [2, 4, 21–23], and which can lead to compact formulae for solutions in terms of determinants. While perhaps less useful for determining properties of high-order rogue waves due to combinatorial complexity, such formulae are effective for the calculation of rogue wave solutions of low order. Fix a value of  $n = 1, 2, \dots$  (everything below depends on  $n$  but we will not systematically indicate this dependence when making new definitions going forward). To convert Riemann-Hilbert Problem 2 into a finite-dimensional linear algebraic system, first observe that for  $\lambda \in \mathbb{C} \setminus D_0$ ,

$$\mathbf{M}_{o/e}^{[n]}(\lambda; x, t) = \mathbf{\Pi}_{o/e}(\lambda; x, t)\mathbf{M}^{[0]}(\lambda; x, t) = \mathbf{\Pi}_{o/e}(\lambda; x, t)\mathbf{E}(\lambda),$$

where  $\mathbf{\Pi}_{o/e}(\lambda; x, t)$  denotes the ordered product of the gauge transformation matrices:

$$(3.79) \quad \mathbf{\Pi}_{o/e}(\lambda; x, t) := \mathbf{G}_{o/e}^{[n-1]}(\lambda; x, t) \cdots \mathbf{G}_{o/e}^{[1]}(\lambda; x, t)\mathbf{G}_{o/e}^{[0]}(\lambda; x, t).$$

As a product of  $n$  matrix factors, each having simple poles at  $\pm i$  as its only singularities and decaying to  $\mathbb{I}$  as  $\lambda \rightarrow \infty$ ,  $\mathbf{\Pi}_{o/e}(\lambda; x, t)$  also decays to  $\mathbb{I}$  for large  $\lambda$  and has poles of order  $n$  at  $\pm i$  and is otherwise analytic. Hence, it necessarily has a finite partial fraction expansion of the form

$$(3.80) \quad \mathbf{\Pi}_{o/e}(\lambda; x, t) = \mathbb{I} + \sum_{k=1}^n \frac{\mathbf{A}_{o/e,k}^+(x, t)}{(\lambda - i)^k} + \frac{\mathbf{A}_{o/e,k}^-(x, t)}{(\lambda + i)^k}.$$

Next, taking into account Corollary 3.11, the jump conditions of Riemann-Hilbert Problem 2 together imply that

$$(3.81) \quad \mathbf{\Pi}_{o/e}(\lambda; x, t)\mathbf{E}(\lambda) = \mathbf{U}_{o/e}^{[n],in}(\lambda; x, t)\mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)^n\mathbf{E}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$$

holds for  $|\lambda| = r$ , where  $\mathbf{U}_{o/e}^{[n],in}(\lambda; x, t) = \mathbf{M}_{o/e}^{[n],in}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$  is a matrix function analytic for  $\lambda \in D_0$ . Therefore (3.81) can be recast as the identity (the exponent  $-n$  indicates the  $n^{\text{th}}$  power of the inverse matrix)

$$(3.82) \quad \mathbf{\Pi}_{o/e}(\lambda; x, t) \cdot \left( \mathbf{E}(\lambda) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{E}(\lambda)^{-1} \right) \cdot \mathbf{G}_{o/e}^{[0]}(\lambda; 0, 0)^{-n} = \mathbf{U}_{o/e}^{[n],in}(\lambda; x, t), \quad |\lambda| = r.$$

The expression in parentheses on the left-hand side is precisely the matrix function  $\mathbf{U}_{\text{bg}}^{\text{in}}(\lambda; x, t)$  defined in (3.40); it is an entire function of  $\lambda$  that is written in explicit form in (3.41). Since the right-hand side of (3.82) admits analytic continuation to the domain  $D_0$ , the same must be true of the left-hand side. But  $\mathbf{U}_{\text{bg}}^{\text{in}}(\lambda; x, t)$  is a known entire function, and  $\mathbf{G}_{\text{o/e}}^{[0]}(\lambda; 0, 0)^{-n}$  can be regarded as known using Corollary 3.11 and will be shortly shown to have poles of order  $n$  at  $\lambda = \pm i$ . Therefore, demanding analyticity of the left-hand side at  $\lambda = \pm i$  imposes conditions on the unknown coefficients in the partial fraction expansion (3.80). These conditions constitute an algebraic representation of the rogue wave solution.

To make these observations into an effective procedure, first note that according to (3.77)–(3.78), the following identities are obvious:

$$(3.83) \quad \mathbf{HK} = \mathbf{KH} = \mathbf{0}$$

and

$$(3.84) \quad \mathbf{H}^2 = 2i\mathbf{H} \quad \text{and} \quad \mathbf{K}^2 = 2i\mathbf{K}.$$

The  $n^{\text{th}}$  power on the right-hand side of (3.75) can now be explicitly computed from the representation (3.76). Indeed, (3.83) eliminates all mixed products and (3.84) reduces all matrix powers to scalar multiples:

$$(3.85) \quad \begin{aligned} \mathbf{G}_{\text{o}}^{[0]}(\lambda; 0, 0)^n &= \mathbb{I} + \sum_{k=1}^n \binom{n}{k} \left[ \frac{\mathbf{H}^k}{(\lambda - i)^k} + \frac{(-\mathbf{K})^k}{(\lambda + i)^k} \right] \\ &= \mathbb{I} + \sum_{k=1}^n \binom{n}{k} (2i)^{k-1} \left[ \frac{\mathbf{H}}{(\lambda - i)^k} + \frac{(-1)^k \mathbf{K}}{(\lambda + i)^k} \right], \end{aligned}$$

$$(3.86) \quad \begin{aligned} \mathbf{G}_{\text{e}}^{[0]}(\lambda; 0, 0)^n &= \mathbb{I} + \sum_{k=1}^n \binom{n}{k} \left[ \frac{\mathbf{K}^k}{(\lambda - i)^k} + \frac{(-\mathbf{H})^k}{(\lambda + i)^k} \right] \\ &= \mathbb{I} + \sum_{k=1}^n \binom{n}{k} (2i)^{k-1} \left[ \frac{\mathbf{K}}{(\lambda - i)^k} + \frac{(-1)^k \mathbf{H}}{(\lambda + i)^k} \right]. \end{aligned}$$

Moreover, (3.83) and (3.84) together imply the identity

$$(3.87) \quad \mathbf{G}_{\text{o}}^{[0]}(\lambda; 0, 0)^{-1} = \mathbf{G}_{\text{e}}^{[0]}(\lambda; 0, 0),$$

which further implies  $\mathbf{G}_{\text{o/e}}^{[0]}(\lambda; 0, 0)^{-n} = \mathbf{G}_{\text{e/o}}^{[0]}(\lambda; 0, 0)^n$ . Thus the final factor on the left-hand side of the relation (3.82) is given by either (3.85) or (3.86) and is therefore an explicit meromorphic function with poles of order  $n$  at  $\lambda = \pm i$ , written in partial fraction expansion form.

*Remark 3.13.* The relation (3.87) has an interesting consequence. If an application of  $\mathcal{D}_o$  is followed by an application of  $\mathcal{D}_e$  or vice versa, then the factors introduced in the jump matrix on  $\Sigma_0$  by these transformations are inverses of each other.

Therefore, the Riemann-Hilbert problem remains unchanged. By uniqueness, we conclude that  $\mathcal{D}_o \circ \mathcal{D}_e = \mathcal{I}$  and the corresponding Bäcklund transformations are inverses of each other.

We use this information to obtain the coefficients in the partial fraction expansion (3.80) in two steps:

- Denote by  $\mathbf{P}_{o/e}(\lambda; x, t)$  the product of the first two factors on the left-hand side of (3.82):  $\mathbf{P}_{o/e}(\lambda; x, t) := \mathbf{\Pi}_{o/e}(\lambda; x, t)\mathbf{U}_{bg}^{in}(\lambda; x, t)$ . This function has poles of order  $n$  at  $\lambda = \pm i$  and hence has convergent Laurent expansions of the form

$$(3.88) \quad \mathbf{P}_{o/e}(\lambda; x, t) = \sum_{j=-n}^{\infty} \mathbf{P}_{o/e,j}^{\pm}(x, t)(\lambda \mp i)^j, \quad |\lambda \mp i| < 2.$$

Then the condition that the left-hand side of (3.82) have removable singularities at  $\lambda = i$  and  $\lambda = -i$  implies the following  $4n$  necessary vector-valued conditions on the coefficient functions of the series (3.88):

$$(3.89) \quad \begin{aligned} \mathbf{P}_{o/e,j}^+(x, t)\mathbf{c}_{\infty,o/e} &= \mathbf{0}, \\ \mathbf{P}_{o/e,j}^-(x, t)\sigma_3\mathbf{c}_{\infty,o/e} &= \mathbf{0}, \end{aligned} \quad -n \leq j \leq n-1,$$

where  $\mathbf{c}_{\infty,o} = [1 : -1]^T \in \mathbb{C}\mathbb{P}^1$  and  $\mathbf{c}_{\infty,e} = [1 : 1]^T \in \mathbb{C}\mathbb{P}^1$ . In each case, the first  $n$  equations are obvious, but to derive the second  $n$  equations one makes repeated use of the identities (3.83) and (3.84).

- Now substitute into (3.89) the Laurent coefficients of the matrix function  $\mathbf{P}_{o/e}(\lambda; x, t) = \mathbf{\Pi}_{o/e}(\lambda; x, t)\mathbf{U}_{bg}^{in}(\lambda; x, t)$ , which can be explicitly written in terms of the unknown coefficients in the partial fraction expansion (3.80) and the Taylor coefficients of the known analytic function  $\mathbf{U}_{bg}^{in}(\lambda; x, t)$  about  $\lambda = \pm i$ :

$$(3.90) \quad \mathbf{U}_{bg}^{in}(\lambda; x, t) = \sum_{j=0}^{\infty} \mathbf{D}_j^{\pm}(x, t)(\lambda \mp i)^j.$$

Defining associated vectors by

$$(3.91) \quad \begin{aligned} \mathbf{w}_{o/e,k}^+(x, t) &:= \mathbf{D}_k^+(x, t)\mathbf{c}_{\infty,o/e}, \\ \mathbf{w}_{o/e,k}^-(x, t) &:= \mathbf{D}_k^-(x, t)\sigma_3\mathbf{c}_{\infty,o/e}, \end{aligned} \quad k = 0, 1, \dots,$$

the equations (3.89) for indices  $-n \leq j \leq -1$  imply the  $2n$  vector equations

$$(3.92) \quad \sum_{k=0}^j \mathbf{A}_{o/e,n-j+k}^+(x, t)\mathbf{w}_{o/e,k}^+(x, t) = 0, \quad 0 \leq j \leq n-1,$$

$$(3.93) \quad \sum_{k=0}^j \mathbf{A}_{o/e,n-j+k}^-(x, t)\mathbf{w}_{o/e,k}^-(x, t) = 0, \quad 0 \leq j \leq n-1.$$

To express the remaining equations in (3.89), first set

$$(3.94) \quad \gamma_{km} := \frac{(-1)^k}{(2i)^{m+k}} \binom{m+k-1}{k}, \quad m = 1, 2, \dots, k = 0, 1, \dots,$$

so that

$$(3.95) \quad \frac{1}{(\lambda + i)^m} = \sum_{k=0}^{\infty} \gamma_{km} (\lambda - i)^k, \quad \frac{1}{(\lambda - i)^m} = \sum_{k=0}^{\infty} (-1)^{m+k} \gamma_{km} (\lambda + i)^k.$$

Introducing the auxiliary unknown matrices

$$(3.96) \quad \begin{aligned} \Gamma_{o/e,k}^+(x, t) &:= \sum_{m=1}^n \gamma_{km} \mathbf{A}_{o/e,m}^-(x, t), \\ \Gamma_{o/e,k}^-(x, t) &:= \sum_{m=1}^n (-1)^{m+k} \gamma_{km} \mathbf{A}_{o/e,m}^+(x, t), \end{aligned}$$

the equations (3.89) for indices  $0 \leq j \leq n - 1$  take the form

$$(3.97) \quad \begin{aligned} &\sum_{k=0}^j \Gamma_{o/e,j-k}^+(x, t) \mathbf{w}_{o/e,k}^+(x, t) \\ &+ \sum_{k=1}^n \mathbf{A}_{o/e,k}^+(x, t) \mathbf{w}_{o/e,j+k}^+(x, t) = -\mathbf{w}_{o/e,j}^+(x, t), \quad 0 \leq j \leq n - 1, \end{aligned}$$

$$(3.98) \quad \begin{aligned} &\sum_{k=0}^j \Gamma_{o/e,j-k}^-(x, t) \mathbf{w}_{o/e,k}^-(x, t) \\ &+ \sum_{k=1}^n \mathbf{A}_{o/e,k}^-(x, t) \mathbf{w}_{o/e,j+k}^-(x, t) = -\mathbf{w}_{o/e,j}^-(x, t), \quad 0 \leq j \leq n - 1. \end{aligned}$$

Eliminating  $\Gamma_{o/e}^+(x, t)$  and  $\Gamma_{o/e}^-(x, t)$  using (3.96), it is clear that equations (3.92), (3.93), (3.97), and (3.98) constitute a square inhomogeneous linear system of dimension  $8n \times 8n$  governing the  $8n$  entries of the matrix coefficients in the partial fraction expansion (3.80).

According to (2.26) the solution of the NLS equation (1.1) stemming from Riemann-Hilbert Problem 2 is

$$(3.99) \quad \begin{aligned} \psi(x, t) &= 2i \lim_{\lambda \rightarrow \infty} \lambda M_{o/e,12}^{[n]}(\lambda; x, t) \\ &= 2i \lim_{\lambda \rightarrow \infty} \lambda (\mathbf{\Pi}_{o/e}(\lambda; x, t) \mathbf{M}^{[0]}(\lambda; x, t))_{12} \\ &= 1 + 2i (\mathbf{A}_{o/e,1}^+(x, t) + \mathbf{A}_{o/e,1}^-(x, t))_{12}. \end{aligned}$$

where we have used that the second column of  $\mathbf{M}^{[0]}(\lambda; x, t) = \mathbf{E}(\lambda)$  has the expansion  $[0 \ 1]^\top + \lambda^{-1}[2i \ 0]^\top + O(\lambda^{-2})$  as  $\lambda \rightarrow \infty$ . Therefore it is sufficient to solve for the first row of the matrices in (3.80), so we introduce row vectors

$$\begin{aligned} [r_{o/e,k}(x, t) \ u_{o/e,k}(x, t)] &:= [1 \ 0] \mathbf{A}_{o/e,k}^+(x, t), \\ [s_{o/e,k}(x, t) \ v_{o/e,k}(x, t)] &:= [1 \ 0] \mathbf{A}_{o/e,k}^-(x, t), \end{aligned}$$

for  $1 \leq k \leq n$ , and the system of equations (3.92), (3.93), (3.97), and (3.98) results in a square inhomogeneous linear system for the vector unknown

$$\mathbf{y} = [r_{o/e,n} \ s_{o/e,n} \ u_{o/e,n} \ v_{o/e,n} \ \cdots \ r_{o/e,1} \ s_{o/e,1} \ u_{o/e,1} \ v_{o/e,1}]^\top \in \mathbb{C}^{4n},$$

with a coefficient matrix that consists of  $(2n)^2 \ 2 \times 2$  blocks:

$$\mathbf{R}_{o/e} := \begin{bmatrix} \mathbf{F}_0^{[1]} & \mathbf{F}_0^{[2]} & \mathbf{0} & \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} \\ \mathbf{F}_1^{[1]} & \mathbf{F}_1^{[2]} & \mathbf{F}_0^{[1]} & \mathbf{F}_0^{[2]} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_{n-1}^{[1]} & \mathbf{F}_{n-1}^{[2]} & \mathbf{F}_{n-2}^{[1]} & \mathbf{F}_{n-2}^{[2]} & \cdots & \cdots & \mathbf{F}_0^{[1]} & \mathbf{F}_0^{[2]} \\ \mathbf{F}_n^{[1]} + \mathbf{H}_{0,n}^{[1]} & \mathbf{F}_n^{[2]} + \mathbf{H}_{0,n}^{[2]} & \mathbf{F}_{n-1}^{[1]} + \mathbf{H}_{0,n-1}^{[1]} & \mathbf{F}_{n-1}^{[2]} + \mathbf{H}_{0,n-1}^{[2]} & \cdots & \cdots & \mathbf{F}_1^{[1]} + \mathbf{H}_{0,1}^{[1]} & \mathbf{F}_1^{[2]} + \mathbf{H}_{0,1}^{[2]} \\ \mathbf{F}_{n+1}^{[1]} + \mathbf{H}_{1,n}^{[1]} & \mathbf{F}_{n+1}^{[2]} + \mathbf{H}_{1,n}^{[2]} & \mathbf{F}_n^{[1]} + \mathbf{H}_{1,n-1}^{[1]} & \mathbf{F}_n^{[2]} + \mathbf{H}_{1,n-1}^{[2]} & \cdots & \cdots & \mathbf{F}_2^{[1]} + \mathbf{H}_{1,1}^{[1]} & \mathbf{F}_2^{[2]} + \mathbf{H}_{1,1}^{[2]} \\ \mathbf{F}_{n+2}^{[1]} + \mathbf{H}_{2,n}^{[1]} & \mathbf{F}_{n+2}^{[2]} + \mathbf{H}_{2,n}^{[2]} & \mathbf{F}_{n+1}^{[1]} + \mathbf{H}_{2,n-1}^{[1]} & \mathbf{F}_{n+1}^{[2]} + \mathbf{H}_{2,n-1}^{[2]} & \cdots & \cdots & \mathbf{F}_{n-1}^{[1]} + \mathbf{H}_{2,1}^{[1]} & \mathbf{F}_{n-1}^{[2]} + \mathbf{H}_{2,1}^{[2]} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_{2n-1}^{[1]} + \mathbf{H}_{n-1,n}^{[1]} & \mathbf{F}_{2n-1}^{[2]} + \mathbf{H}_{n-1,n}^{[2]} & \mathbf{F}_{2n-2}^{[1]} + \mathbf{H}_{n-1,n-1}^{[1]} & \mathbf{F}_{2n-2}^{[2]} + \mathbf{H}_{n-1,n-1}^{[2]} & \cdots & \cdots & \mathbf{F}_n^{[1]} + \mathbf{H}_{n-1,1}^{[1]} & \mathbf{F}_n^{[2]} + \mathbf{H}_{n-1,1}^{[2]} \end{bmatrix},$$

where we momentarily suppressed the subscript o/e for the block elements. The blocks of  $\mathbf{R} = \mathbf{R}(x, t)$  are defined in terms of the following  $2 \times 2$  matrices:

$$\begin{aligned} \mathbf{F}_k^{[j]} = \mathbf{F}_k^{[j]}(x, t) &:= \begin{bmatrix} (\mathbf{w}_{e/o,k}^+(x, t))_j & 0 \\ 0 & (\mathbf{w}_{e/o,k}^-(x, t))_j \end{bmatrix} \\ (3.100) \qquad \qquad \qquad &=: \begin{bmatrix} f_{e/o,k}^{[j]+}(x, t) & 0 \\ 0 & f_{e/o,k}^{[j]-}(x, t) \end{bmatrix}, \quad j = 1, 2, \end{aligned}$$

$$\begin{aligned} \mathbf{H}_{m,k}^{[j]} = \mathbf{H}_{m,k}^{[j]}(x, t) \\ (3.101) \qquad \qquad \qquad &:= \begin{bmatrix} 0 & \sum_{\ell=0}^m \gamma_{\ell k} f_{e/o,m-\ell}^{[j]+}(x, t) \\ \sum_{\ell=0}^m (-1)^{\ell+k} \gamma_{\ell k} f_{e/o,m-\ell}^{[j]-}(x, t) & 0 \end{bmatrix}, \\ j = 1, 2. \end{aligned}$$

Thus, the first row of the solution of Riemann-Hilbert Problem 2 can be obtained by solving the linear system

$$(3.102) \quad \mathbf{R}_{e/o}(x, t)\mathbf{y}(x, t) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -f_{e/o,0}^{[1]+}(x, t) \\ -f_{o/e,0}^{[1]-}(x, t) \\ -f_{e/o,1}^{[1]+}(x, t) \\ -f_{e/o,1}^{[1]-}(x, t) \\ \vdots \\ -f_{e/o,n-1}^{[1]+}(x, t) \\ -f_{e/o,n-1}^{[1]-}(x, t) \end{bmatrix},$$

and consequently the rogue wave solutions of arbitrary order are recovered by Cramer’s rule:

$$(3.103) \quad \begin{aligned} \psi_{P_{2n-1}}(x, t) &= 1 + 2i(u_{o,1}(x, t) + v_{o,1}(x, t)) \\ &= 1 + 2i \frac{\det(\mathbf{R}_{o,2}(x, t)) + \det(\mathbf{R}_{o,1}(x, t))}{\det(\mathbf{R}_o(x, t))}, \end{aligned}$$

$$(3.104) \quad \begin{aligned} \psi_{P_{2n}}(x, t) &= 1 + 2i(u_{e,1}(x, t) + v_{e,1}(x, t)) \\ &= 1 + 2i \frac{\det(\mathbf{R}_{e,2}(x, t)) + \det(\mathbf{R}_{e,1}(x, t))}{\det(\mathbf{R}_e(x, t))}. \end{aligned}$$

Here  $\mathbf{R}_{o/e,k}(x, t)$  stands for the matrix  $\mathbf{R}_{o/e}(x, t)$  with its  $k^{\text{th}}$  column replaced by the right-hand side vector in (3.102). For  $n = 1$ , the system (3.102) is of size  $4 \times 4$ ,<sup>7</sup> and it can be solved by hand to obtain the Peregrine breather with its peak located at  $(0, 0)$ . As  $n$  grows, the complexity of the system grows rapidly and we solve both of the systems  $\mathbf{R}_{o/e,k}$  symbolically by computer for  $n = 1, 2, 3, 4$ . See Figure 3.6 for the “odd-order” rogue-wave solutions  $\psi_{P_{2n-1}}(x, t)$  of the NLS equation obtained by solving the linear system of (3.102) with  $\mathbf{R}(x, t) = \mathbf{R}_o(x, t)$ , and see Figure 3.7 for the “even-order” rogue wave solutions  $\psi_{P_{2n}}(x, t)$  of the NLS equation in the case  $\mathbf{R}(x, t) = \mathbf{R}_e(x, t)$ . See Figure 3.8 for a surface plot of  $|\psi_8(x, t)|$  (corresponding to the linear system with  $n = 4$  and  $\mathbf{R}(x, t) = \mathbf{R}_e(x, t)$ ).

<sup>7</sup> In general, the system has dimension  $4n \times 4n$ . It is also possible to obtain, via a different rational ansatz, a system of size  $k \times k$ , where  $k = 2n$  for  $k$  even and  $k = 2n - 1$  for  $k$  odd. See the appendix of [7] for details.

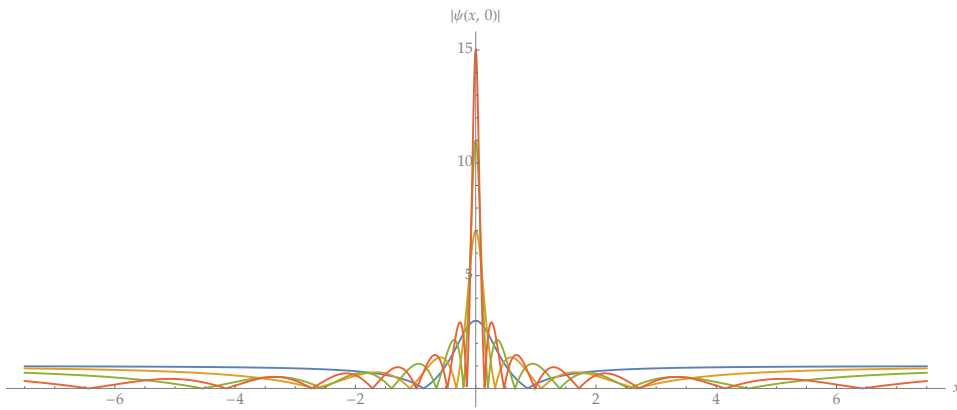


FIGURE 3.6. Moduli of the first four “odd-order” rogue waves:  $|\psi_{P_1}(x, 0)|$  (blue),  $|\psi_{P_3}(x, 0)|$  (orange),  $|\psi_{P_5}(x, 0)|$  (green), and  $|\psi_{P_7}(x, 0)|$  (red).

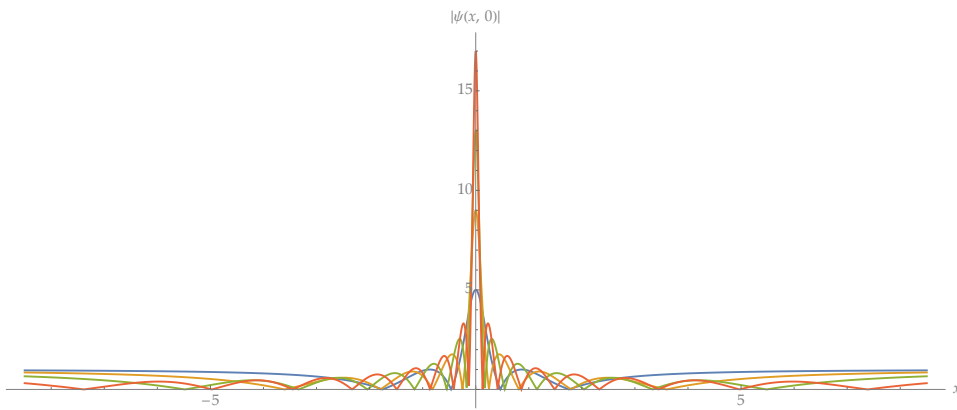


FIGURE 3.7. Moduli of the first four “even-order” rogue waves:  $|\psi_{P_2}(x, 0)|$  (blue),  $|\psi_{P_4}(x, 0)|$  (orange),  $|\psi_{P_6}(x, 0)|$  (green), and  $|\psi_{P_8}(x, 0)|$  (red).

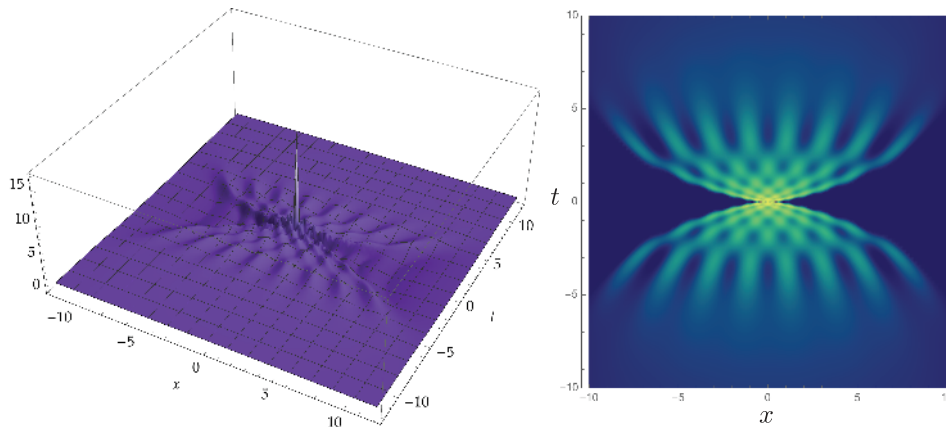


FIGURE 3.8. Modulus of the 8<sup>th</sup>-order rogue wave:  $|\psi_{P_8}(x, t)|$ , obtained from the solution of Riemann-Hilbert Problem 2 with  $n = 4$ . The pattern evident in the density plot (right-hand panel) is reminiscent of the wave patterns typical in solutions of the focusing NLS equation in the semiclassical limit [31].

### 4 Linearization of the Direct and Inverse Scattering Transforms and Application to the Peregrine Solution

#### 4.1 Complex NLS and Linearization Preliminaries

Consider the complex nonlinear Schrödinger (cNLS) for the scalar unknown functions  $(\psi, \phi)$ :

$$(4.1) \quad \begin{aligned} i\psi_t + \frac{1}{2}\psi_{xx} + (-\psi\phi - 1)\psi &= 0, \\ i\phi_t - \frac{1}{2}\phi_{xx} - (-\psi\phi - 1)\phi &= 0. \end{aligned}$$

In case the focusing symmetry  $\phi = -\psi^*$  holds, (4.1) reduces to the focusing nonlinear Schrödinger (NLS) equation (1.1). Consider the formal linearization of (4.1) around a solution  $(\psi_0, \phi_0)$  by considering solutions  $(\psi, \phi)$  that are of the form:

$$(4.2) \quad \begin{aligned} \psi &= \psi_0 + \varepsilon\psi_1 + o(\varepsilon), \\ \phi &= \phi_0 + \varepsilon\phi_1 + o(\varepsilon), \quad 0 < \varepsilon \ll 1. \end{aligned}$$



Enforcing (4.2) to solve (4.1) and retaining  $O(\varepsilon)$  terms yields the following linear partial differential equations for the perturbation functions  $(\psi_1, \phi_1)$ :

$$(4.3) \quad \begin{aligned} i\psi_{1t} + \frac{1}{2}\psi_{1xx} + (-2\psi_0\phi_0 - 1)\psi_1 - \psi_0^2\phi_1 &= 0, \\ i\phi_{1t} - \frac{1}{2}\phi_{1xx} - (-2\psi_0\phi_0 - 1)\phi_1 + \phi_0^2\psi_1 &= 0, \end{aligned}$$

which we will refer to as the linearized complex NLS equation (lcNLS), and in case the focusing symmetry holds for the unperturbed fields  $(\phi_0 = -\psi_0^*)$  and first-order perturbations  $(\phi_1 = -\psi_1^*)$ , (4.3) reduces to the linearized NLS (lNLS) equation:

$$(4.4) \quad i\psi_{1t} + \frac{1}{2}\psi_{1xx} + (2|\psi_0|^2 - 1)\psi_1 + \psi_0^2\psi_1^* = 0.$$

### 4.2 Squared Eigenfunctions

Let  $\zeta$  be an arbitrary complex number, and denote by  $\mathbf{u} = \mathbf{u}^{[a]}(\zeta; x, t)$  and  $\mathbf{u} = \mathbf{u}^{[b]}(\zeta; x, t)$  any two (not necessarily independent or distinct)  $2 \times 1$  column vector simultaneous solutions of the Lax pair associated with the solutions  $(\psi_0, \phi_0)$  of the cNLS equation for the same value of the spectral parameter  $\lambda = \zeta$ :

$$(4.5) \quad \begin{aligned} \mathbf{u}_x &= \begin{bmatrix} -i\lambda & \psi_0 \\ \phi_0 & i\lambda \end{bmatrix} \mathbf{u}, \\ \mathbf{u}_t &= \begin{bmatrix} -i\lambda^2 + \frac{i}{2}(-\psi_0\phi_0 - 1) & \lambda\psi_0 + \frac{i}{2}\psi_{0x} \\ \lambda\phi_0 - \frac{i}{2}\phi_{0x} & i\lambda^2 - \frac{i}{2}(-\psi_0\phi_0 - 1) \end{bmatrix} \mathbf{u}. \end{aligned}$$

Define the corresponding *squared eigenfunctions*

$$(4.6) \quad \begin{aligned} \mu_1^{[ab]}(\zeta; x, t) &:= u_1^{[a]}(\zeta; x, t)u_1^{[b]}(\zeta; x, t), \\ \mu_2^{[ab]}(\zeta; x, t) &:= u_1^{[a]}(\zeta; x, t)u_2^{[b]}(\zeta; x, t) + u_2^{[a]}(\zeta; x, t)u_1^{[b]}(\zeta; x, t), \\ \mu_3^{[ab]}(\zeta; x, t) &:= u_2^{[a]}(\zeta; x, t)u_2^{[b]}(\zeta; x, t). \end{aligned}$$

These squared eigenfunctions  $\mu_j = \mu_j^{[ab]}$ ,  $j = 1, 2, 3$ , satisfy the  $3 \times 3$  Lax pair:

$$(4.7) \quad \begin{aligned} \frac{\partial}{\partial x} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} &= \begin{bmatrix} -2i\lambda & \psi_0 & 0 \\ 2\phi_0 & 0 & 2\psi_0 \\ 0 & \phi_0 & 2i\lambda \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \\ \frac{\partial}{\partial t} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} &= \begin{bmatrix} -2i\lambda^2 + i(-\psi_0\phi_0 - 1) & \lambda\psi_0 + \frac{i}{2}\psi_{0x} & 0 \\ 2\lambda\phi_0 - i\phi_{0x} & 0 & 2\lambda\psi_0 + i\psi_{0x} \\ 0 & \lambda\phi_0 - \frac{i}{2}\phi_{0x} & 2i\lambda^2 - i(-\psi_0\phi_0 - 1) \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \end{aligned}$$

for the spectral parameter evaluated at  $\lambda = \zeta$ . Using the differential equations (4.7) it follows that for any  $\zeta \in \mathbb{C}$  the squared eigenfunctions

$$(\mu_1^{[ab]}(\zeta; x, t), \mu_3^{[ab]}(\zeta; x, t))$$

satisfy the following partial differential equations in which the spectral parameter does not appear explicitly:

$$(4.8) \quad \begin{aligned} i\mu_{1t}^{[ab]} + \frac{1}{2}\mu_{1xx}^{[ab]} + (-2\psi_0\phi_0 - 1)\mu_1^{[ab]} - \psi_0^2\mu_3^{[ab]} &= 0, \\ i\mu_{3t}^{[ab]} - \frac{1}{2}\mu_{3xx}^{[ab]} - (-2\psi_0\phi_0 - 1)\mu_3^{[ab]} + \phi_0^2\mu_1^{[ab]} &= 0, \end{aligned}$$

which is precisely the lcNLS given in (4.3) with  $\psi_1 := \mu_1^{[ab]}(\zeta; x, t)$  and  $\phi_1 := \mu_3^{[ab]}(\zeta; x, t)$ . Explicitly, we have the lcNLS written as the linear system

$$\begin{bmatrix} i\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial x^2} - 2\psi_0\phi_0 - 1 & -\psi_0^2 \\ \phi_0^2 & i\frac{\partial}{\partial t} - \frac{1}{2}\frac{\partial^2}{\partial x^2} + 2\psi_0\phi_0 + 1 \end{bmatrix} \begin{bmatrix} \mu_1^{[ab]}(\zeta; x, t) \\ \mu_3^{[ab]}(\zeta; x, t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Now suppose that  $\phi_0(x, t) = -\psi_0(x, t)^*$ . Unfortunately, the ordered pair of squared eigenfunctions  $(\psi_1(x, t), \phi_1(x, t)) = (\mu_1^{[ab]}(\zeta; x, t), \mu_3^{[ab]}(\zeta; x, t))$  does *not* necessarily satisfy the focusing symmetry:  $\mu_3^{[ab]}(\zeta; x, t) = -\mu_1^{[ab]}(\zeta; x, t)^*$  needed to reduce (4.3) to (4.4), but this can be remedied by applying the superposition principle making use of different values of the spectral parameter. Indeed, the focusing symmetry  $\phi_0(x, t) = -\psi_0(x, t)^*$  of the unperturbed problem implies that the vectors

$$(4.9) \quad \mathbf{u}^{[j]}(\zeta^*; x, t) := \begin{bmatrix} -u_2^{[j]}(\zeta; x, t)^* \\ u_1^{[j]}(\zeta; x, t)^* \end{bmatrix}, \quad j = a, b,$$

are simultaneous solutions of (4.5) for the spectral parameter  $\lambda = \zeta^*$ . Denoting the corresponding squared eigenfunctions defined via (4.6) from these two vectors as  $\mu_k^{[ab]}(\zeta^*; x, t)$ ,  $k = 1, 2, 3$ , it follows easily that

$$\mu_1^{[ab]}(\zeta^*; x, t) = \mu_3^{[ab]}(\zeta; x, t)^*, \quad \mu_3^{[ab]}(\zeta^*; x, t) = \mu_1^{[ab]}(\zeta; x, t)^*.$$

Now since the spectral parameter does not appear explicitly in the linear system (4.3), it is solved by both  $(\psi_1, \phi_1) = (\mu_1^{[ab]}(\zeta; x, t), \mu_3^{[ab]}(\zeta; x, t))$  and  $(\psi_1, \phi_1) = (\mu_1^{[ab]}(\zeta^*; x, t), \mu_3^{[ab]}(\zeta^*; x, t))$ , and hence by superposition

$$(4.10) \quad \begin{aligned} &(\psi_1(x, t), \phi_1(x, t)) \\ &:= (\mu_1^{[ab]}(\zeta; x, t) - \mu_1^{[ab]}(\zeta^*; x, t), \mu_3^{[ab]}(\zeta; x, t) - \mu_3^{[ab]}(\zeta^*; x, t)) \\ &\quad (\mu_1^{[ab]}(\zeta; x, t) - \mu_3^{[ab]}(\zeta; x, t)^*, \mu_3^{[ab]}(\zeta; x, t) - \mu_1^{[ab]}(\zeta; x, t)^*) \end{aligned}$$

is a solution of the system (4.3) for each  $\zeta \in \mathbb{C}$  that satisfies the focusing symmetry  $\phi_1(x, t) = -\psi_1(x, t)^*$ . Therefore

$$(4.11) \quad \begin{aligned} \psi_1 &:= \mu^{[ab]}(\zeta; x, t) \\ &:= \mu_1^{[ab]}(\zeta; x, t) - \mu_3^{[ab]}(\zeta; x, t)^* \\ &\quad u_1^{[a]}(\zeta; x, t)u_1^{[b]}(\zeta; x, t) - u_2^{[a]}(\zeta; x, t)^*u_2^{[b]}(\zeta; x, t)^* \end{aligned}$$

solves the linearized focusing NLS equation (4.4) whenever  $\mathbf{u}^{[j]}(\zeta; x, t)$ ,  $j = a, b$ , are any two simultaneous solutions of (4.5) with  $\phi_0(x, t) = -\psi_0(x, t)^*$  for the same arbitrary value  $\zeta$  of the spectral parameter. Note that if the vectors  $\mathbf{u}^{[j]}(\zeta; x, t)$ ,  $j = a, b$ , are multiplied by scalar factors  $c^{[j]}(\zeta)$ ,  $j = a, b$ , whose product is  $C(\zeta) = c^{[a]}(\zeta)c^{[b]}(\zeta)$ , then the formula (4.11) is transformed into

$$(4.12) \quad \begin{aligned} \psi_1(x, t) &= C(\zeta)u_1^{[a]}(\zeta; x, t)u_1^{[b]}(\zeta; x, t) \\ &\quad - C(\zeta)^*u_2^{[a]}(\zeta; x, t)^*u_2^{[b]}(\zeta; x, t)^*, \end{aligned}$$

which again is a solution of the linearized focusing NLS equation (4.4) regardless of the value of  $C(\zeta) \in \mathbb{C}$ . Writing  $C(\zeta) = \alpha(\zeta) + i\beta(\zeta)$  for real numbers  $\alpha(\zeta)$  and  $\beta(\zeta)$ , the solution (4.12) can be written as the real linear combination  $\alpha(\zeta)\mu^{[ab]}(\zeta; x, t) + \beta(\zeta)\nu^{[ab]}(\zeta; x, t)$  where  $\mu(\zeta; x, t)$  is defined by (4.11) and where

$$(4.13) \quad \nu^{[ab]}(\zeta; x, t) := iu_1^{[a]}(\zeta; x, t)u_1^{[b]}(\zeta; x, t) + iu_2^{[a]}(\zeta; x, t)^*u_2^{[b]}(\zeta; x, t)^*.$$

Further solutions of the linearized NLS equation (4.4) can be obtained by taking real linear combinations of  $\mu^{[ab]}(\zeta; x, t)$  and  $\nu^{[ab]}(\zeta; x, t)$  for different complex values of  $\zeta$ .

A natural question that arises is that of completeness of the set of particular solutions of (4.4) of the type indicated above in which  $\zeta$  is allowed to range over some subset of the complex numbers. What class of solutions of (4.4) can be represented as suitable finite or infinite superpositions of them?

### 4.3 Formal Linearization of the Inverse-Scattering Transform

A systematic way to address the completeness issue is to analyze the solution of the *nonlinear* problem with initial data  $\psi(x, 0) = \psi_0(x, 0) + \varepsilon\psi_1(x, 0)$  obtained via the inverse-scattering transform. Kaup [27] used this approach to study localized perturbations of solutions decaying to the zero background. Here we use the robust inverse-scattering transform introduced in Section 2 and consider localized perturbations of solutions decaying to the nonzero background value  $\psi \equiv 1$ .

#### Linearizing the Direct Transform

Let  $\psi_0(x, t)$  be a solution of (1.1) for which  $\psi_0 - 1 \in L^1(\mathbb{R})$  as a function of  $x$  for each  $t \in \mathbb{R}$ . Suppose that  $\psi_1(x)$  is a sufficiently localized perturbation. For  $\varepsilon > 0$  small, we consider the initial data  $\psi(x, 0; \varepsilon) = \psi_0(x, 0) + \varepsilon\psi_1(x)$ . For sufficiently large radius  $r > 0$  (independent of  $\varepsilon$ ) we calculate the core jump matrices

$$(4.14) \quad \begin{aligned} \mathbf{V}^+(\lambda; \varepsilon) &:= [a(\lambda; \varepsilon)^{-1}\mathbf{j}^{-,1}(\lambda; 0, 0; \varepsilon); \mathbf{j}^{+,2}(\lambda; 0, 0; \varepsilon)], \\ &\quad \text{Im}\{\lambda\} > 0, |\lambda| = r, \end{aligned}$$

$$(4.15) \quad \begin{aligned} \mathbf{V}^-(\lambda; \varepsilon) &:= [\mathbf{j}^{+,1}(\lambda; 0, 0; \varepsilon); a(\lambda^*; \varepsilon)^{-*}\mathbf{j}^{-,2}(\lambda; 0, 0; \varepsilon)], \\ &\quad \text{Im}\{\lambda\} < 0, |\lambda| = r, \end{aligned}$$

and

$$(4.16) \quad \mathbf{V}^{\text{out}}(\lambda; \varepsilon) := \begin{bmatrix} 1 + |R(\lambda; \varepsilon)|^2 & R(\lambda; \varepsilon)^* \\ R(\lambda; \varepsilon) & 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}, |\lambda| > r,$$

$$R(\lambda; \varepsilon) := \frac{b(\lambda; \varepsilon)}{a(\lambda; \varepsilon)},$$

where  $\mathbf{j}^{\pm, k}(\lambda; 0, 0; \varepsilon)$  denote the columns of the Jost matrices for the potential  $\psi(x, 0; \varepsilon) = \psi_0(x, 0) + \varepsilon\psi_1(x)$  evaluated at  $x = t = 0$  and where  $a(\lambda; \varepsilon)$  and  $b(\lambda; \varepsilon)$  refer to the associated scattering matrix. More to the point, the Jost column vectors are the unique solutions of the following Volterra integral equations: setting  $\mathbf{j}^{\pm, 1}(\lambda; x, 0; \varepsilon) = e^{-i\rho(\lambda)x}\mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon)$ ,  $\mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon)$  satisfies

$$(4.17) \quad \begin{aligned} & \mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon) - \mathcal{V}^{\pm, 1}[\psi_0(\cdot, 0) - 1]\mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon) \\ & = \mathbf{e}^1(\lambda) + \varepsilon\mathcal{V}^{\pm, 1}[\psi_1]\mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon), \quad \mp \text{Im}\{\rho(\lambda)\} \geq 0, \end{aligned}$$

where  $\mathcal{V}^{\pm, 1}[\psi]$  denote Volterra integral operators

$$(4.18) \quad \mathcal{V}^{\pm, 1}[\psi]\mathbf{f}(x) := \int_{\pm\infty}^x \mathbf{E}(\lambda) \begin{bmatrix} 1 & 0 \\ 0 & e^{2i\rho(\lambda)(x-y)} \end{bmatrix} \mathbf{E}(\lambda)^{-1} \begin{bmatrix} 0 & \psi(y) \\ -\psi(y)^* & 0 \end{bmatrix} \mathbf{f}(y) dy,$$

$$\mp \text{Im}\{\rho(\lambda)\} \geq 0,$$

and setting  $\mathbf{j}^{\pm, 2}(\lambda; x, 0; \varepsilon) = e^{i\rho(\lambda)x}\mathbf{k}^{\pm, 2}(\lambda; x, 0; \varepsilon)$ ,  $\mathbf{k}^{\pm, 2}(\lambda; x, 0; \varepsilon)$  satisfies

$$(4.19) \quad \begin{aligned} & \mathbf{k}^{\pm, 2}(\lambda; x, 0; \varepsilon) - \mathcal{V}^{\pm, 2}[\psi_0(\cdot, 0) - 1]\mathbf{k}^{\pm, 2}(\lambda; x, 0; \varepsilon) \\ & = \mathbf{e}^2(\lambda) + \varepsilon\mathcal{V}^{\pm, 2}[\psi_1]\mathbf{k}^{\pm, 2}(\lambda; x, 0; \varepsilon), \quad \pm \text{Im}\{\rho(\lambda)\} \geq 0, \end{aligned}$$

where  $\mathcal{V}^{\pm, 2}[\psi]$  denote Volterra operators

$$(4.20) \quad \mathcal{V}^{\pm, 2}[\psi]\mathbf{f}(x) := \int_{\pm\infty}^x \mathbf{E}(\lambda) \begin{bmatrix} e^{-2i\rho(\lambda)(x-y)} & 0 \\ 0 & 1 \end{bmatrix} \mathbf{E}(\lambda)^{-1} \begin{bmatrix} 0 & \psi(y) \\ -\psi(y)^* & 0 \end{bmatrix} \mathbf{f}(y) dy,$$

$$\pm \text{Im}\{\rho(\lambda)\} \geq 0.$$

Under the assumption that  $\psi \in L^1(\mathbb{R})$  and the indicated conditions on  $\text{Im}\{\rho(\lambda)\}$ , the operators  $\mathcal{V}^{\pm, j}[\psi]$ ,  $j = 1, 2$ , are bounded on  $L^\infty(\mathbb{R})$ , and also  $\mathcal{I} - \mathcal{V}^{\pm, j}[\psi]$ ,  $j = 1, 2$ , have bounded inverses on  $L^\infty(\mathbb{R})$ . The corresponding operator norms  $\|\mathcal{V}^{\pm, j}[\psi]\|$  and  $\|(\mathcal{I} - \mathcal{V}^{\pm, j}[\psi])^{-1}\|$ ,  $j = 1, 2$ , are uniformly bounded with respect to  $\lambda$  under the indicated conditions on  $\text{Im}\{\rho(\lambda)\}$ . Thus, (4.17) can be rewritten in the form

$$(4.21) \quad \begin{aligned} & \mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon) \\ & - \varepsilon(\mathcal{I} - \mathcal{V}^{\pm, 1}[\psi_0(\cdot, 0) - 1])^{-1} \circ \mathcal{V}^{\pm, 1}[\psi_1]\mathbf{k}^{\pm, 1}(\lambda; x, 0; \varepsilon) \\ & = \mathbf{k}_0^{\pm, 1}(\lambda; x, 0), \quad \mp \text{Im}\{\rho(\lambda)\} \geq 0, \end{aligned}$$

where

$$(4.22) \quad \mathbf{k}_0^{\pm, 1}(\lambda; x, 0) := (\mathcal{I} - \mathcal{V}^{\pm, 1}[\psi_0(\cdot, 0) - 1])^{-1}\mathbf{e}^1(\lambda),$$

$$\mp \text{Im}\{\rho(\lambda)\} \geq 0, \quad \rho(\lambda) \neq 0.$$

Similarly, (4.19) can be rewritten as

$$\begin{aligned}
 & \mathbf{k}^{\pm,2}(\lambda; x, 0; \varepsilon) \\
 (4.23) \quad & - \varepsilon(\mathcal{J} - \mathcal{V}^{\pm,2}[\psi_0(\cdot, 0) - 1])^{-1} \circ \mathcal{V}^{\pm,2}[\psi_1] \mathbf{k}^{\pm,2}(\lambda; x, 0; \varepsilon) \\
 & = \mathbf{k}_0^{\pm,2}(\lambda; x, 0), \quad \pm \operatorname{Im}\{\rho(\lambda)\} \geq 0,
 \end{aligned}$$

where

$$\begin{aligned}
 (4.24) \quad & \mathbf{k}_0^{\pm,2}(\lambda; x, 0) := (\mathcal{J} - \mathcal{V}^{\pm,2}[\psi_0(\cdot, 0) - 1])^{-1} \mathbf{e}^2(\lambda), \\
 & \pm \operatorname{Im}\{\rho(\lambda)\} \geq 0, \quad \rho(\lambda) \neq 0.
 \end{aligned}$$

In both cases, the inhomogeneous terms  $\mathbf{k}_0^{\pm,j}(\lambda; x, 0)$ ,  $j = 1, 2$ , lie in  $L^\infty(\mathbb{R})$  (the additional condition  $\rho(\lambda) \neq 0$  is required because otherwise  $\mathbf{e}^j(\lambda)$  does not exist), and if  $\rho(\lambda)$  is bounded away from 0, their norms are bounded uniformly with respect to  $\lambda$ . It follows that if  $\varepsilon\|(\mathcal{J} - \mathcal{V}^{\pm,1}[\psi_0(\cdot, 0) - 1])^{-1}\| \cdot \|\mathcal{V}^{\pm,1}[\psi_1]\| < 1$ , then (4.21) can be solved by a Neumann series convergent in  $L^\infty(\mathbb{R})$ , which takes the form of a *power series* in  $\varepsilon$ . A similar result holds for (4.23) under the condition  $\varepsilon\|(\mathcal{J} - \mathcal{V}^{\pm,2}[\psi_0(\cdot, 0) - 1])^{-1}\| \cdot \|\mathcal{V}^{\pm,2}[\psi_1]\| < 1$ . This implies that the unique solutions furnished by the sums of the convergent Neumann series are analytic functions of  $\varepsilon$  to  $L^\infty(\mathbb{R})$  at  $\varepsilon = 0$ .

Note that the inhomogeneous terms defined in (4.22) and (4.24) if also  $\rho(\lambda) \neq 0$  are simply the solutions of the integral equations (4.17) and (4.19), respectively, for  $\varepsilon = 0$ , and therefore are proportional via the exponential factors  $e^{\pm i\rho(\lambda)x}$  to the Jost solutions of the unperturbed problem. We therefore have series representations:

$$\begin{aligned}
 (4.25) \quad & \mathbf{j}^{\pm,1}(\lambda; x, 0; \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \mathbf{j}_n^{\pm,1}(\lambda; x, 0), \quad \mp \operatorname{Im}\{\rho(\lambda)\} \geq 0, \quad \rho(\lambda) \neq 0, \\
 & \mathbf{j}_n^{\pm,1}(\lambda; x, 0) = e^{-i\rho(\lambda)x} \mathbf{k}_n^{\pm,1}(\lambda; x, 0),
 \end{aligned}$$

and

$$\begin{aligned}
 (4.26) \quad & \mathbf{j}^{\pm,2}(\lambda; x, 0; \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \mathbf{j}_n^{\pm,2}(\lambda; x, 0), \quad \pm \operatorname{Im}\{\rho(\lambda)\} \geq 0, \quad \rho(\lambda) \neq 0, \\
 & \mathbf{j}_n^{\pm,2}(\lambda; x, 0) = e^{i\rho(\lambda)x} \mathbf{k}_n^{\pm,2}(\lambda; x, 0),
 \end{aligned}$$

both convergent for sufficiently small  $|\varepsilon|$ , with the convergence interpreted in a weighted  $L^\infty(\mathbb{R})$  space corresponding to unweighted  $L^\infty(\mathbb{R})$  convergence for the vector function  $\mathbf{k}^{\pm,j}(\lambda; x, 0; \varepsilon)$ ,  $j = 1, 2$ .

We may now solve explicitly for the first-order correction terms  $\mathbf{j}_1^{\pm,j}(\lambda; x, 0)$ . These are given at first by the expressions

$$\begin{aligned}
 (4.27) \quad & \mathbf{k}_1^{\pm,1}(\lambda; x, 0) = (\mathcal{J} - \mathcal{V}^{\pm,1}[\psi_0(\cdot, 0) - 1])^{-1} \circ \mathcal{V}^{\pm,1}[\psi_1] \mathbf{k}_0^{\pm,1}(\lambda; x, 0), \\
 & \mp \operatorname{Im}\{\rho(\lambda)\} \geq 0, \quad \rho(\lambda) \neq 0,
 \end{aligned}$$

and

$$(4.28) \quad \mathbf{k}_1^{\pm,2}(\lambda; x, 0) = (\mathcal{L} - \mathcal{V}^{\pm,2}[\psi_0(\cdot, 0) - 1])^{-1} \circ \mathcal{V}^{\pm,2}[\psi_1] \mathbf{k}_0^{\pm,2}(\lambda; x, 0),$$

$$\pm \operatorname{Im}\{\rho(\lambda)\} \geq 0, \rho(\lambda) \neq 0.$$

Multiplying these through by

$$e^{-i\rho(\lambda)x} (\mathcal{L} - \mathcal{V}^{\pm,1}[\psi_0(\cdot, 0) - 1])$$

and

$$e^{i\rho(\lambda)x} (\mathcal{L} - \mathcal{V}^{\pm,2}[\psi_0(\cdot, 0) - 1]),$$

respectively, eliminating  $\mathbf{k}_n^{\pm,j}(\lambda; x, 0)$  in favor of  $\mathbf{j}_n^{\pm,j}(\lambda; x, 0)$ ,  $j = 1, 2$ , and combining the resulting equations with their derivatives with respect to  $x$  yields in all cases the same differential equation

$$(4.29) \quad \frac{\partial \mathbf{j}_1^{\pm,j}}{\partial x}(\lambda; x, 0) = \begin{bmatrix} -i\lambda & \psi_0(x, 0) \\ -\psi_0(x, 0)^* & i\lambda \end{bmatrix} \mathbf{j}_1^{\pm,j}(\lambda; x, 0)$$

$$+ \begin{bmatrix} 0 & \psi_1(x) \\ -\psi_1(x)^* & 0 \end{bmatrix} \mathbf{j}_0^{\pm,j}(\lambda; x, 0), \quad j = 1, 2.$$

The general solution of this equation may be obtained from a fundamental matrix for the homogeneous problem and variation of parameters. To obtain  $\mathbf{j}_1^{-,1}(\lambda; x, 0)$  and  $\mathbf{j}_1^{+,2}(\lambda; x, 0)$  for  $|\lambda| = r$  with  $\operatorname{Im}\{\lambda\} > 0$ , we use the fundamental matrix  $\mathbf{U}_0^{\text{BC}}(\lambda; x, 0) = \mathbf{M}_0^{\text{BC}}(\lambda; x, 0) e^{-i\rho(\lambda)x\sigma_3}$ , which is given explicitly in terms of the Jost solutions of the unperturbed problem by

$$(4.30) \quad \mathbf{U}_0^{\text{BC}}(\lambda; x, 0) = \begin{bmatrix} a_0(\lambda)^{-1} \mathbf{j}_0^{-,1}(\lambda; x, 0); & \mathbf{j}_0^{+,2}(\lambda; x, 0) \end{bmatrix},$$

$$|\lambda| = r, \operatorname{Im}\{\lambda\} > 0,$$

where  $a_0(\lambda)$  is the Wronskian determinant that ensures that  $\det(\mathbf{U}_0^{\text{BC}}(\lambda; x, 0)) = 1$ ; note that  $a_0(\lambda) \neq 0$  by choice of sufficiently large radius  $r$ .

We also observe that  $\mathbf{V}^+(\lambda; 0) = \mathbf{U}_0^{\text{BC}}(\lambda; 0, 0)$ . Thus, the variation-of-parameters formula gives, if  $|\lambda| = r$  and  $\operatorname{Im}\{\lambda\} > 0$ ,

$$\mathbf{j}_1^{-,1}(\lambda; x, 0) =$$

$$\mathbf{U}_0^{\text{BC}}(\lambda; x, 0) \int_{-\infty}^x \mathbf{U}_0^{\text{BC}}(\lambda; y, 0)^{-1} \begin{bmatrix} 0 & \psi_1(y) \\ -\psi_1(y)^* & 0 \end{bmatrix} \mathbf{j}_0^{-,1}(\lambda; y, 0) dy,$$

where the lower limit of integration is fixed by requiring that  $\mathbf{k}_1^{-,1}(\lambda; x, 0) \rightarrow \mathbf{0}$  as  $x \rightarrow -\infty$ , and

$$\mathbf{j}_1^{+,2}(\lambda; x, 0) =$$

$$\mathbf{U}_0^{\text{BC}}(\lambda; x, 0) \int_{+\infty}^x \mathbf{U}_0^{\text{BC}}(\lambda; y, 0)^{-1} \begin{bmatrix} 0 & \psi_1(y) \\ -\psi_1(y)^* & 0 \end{bmatrix} \mathbf{j}_0^{+,2}(\lambda; y, 0) dy$$

where the lower limit of integration is fixed by requiring that  $\mathbf{k}_1^{+,2}(\lambda; x, 0) \rightarrow \mathbf{0}$  as  $x \rightarrow +\infty$ . With the squared eigenfunctions

$$(4.31) \quad \begin{aligned} \mu_1^{[jk]}(\lambda; x, t) &:= U_{0,1j}^{\text{BC}}(\lambda; x, t)U_{0,1k}^{\text{BC}}(\lambda; x, t), \\ \mu_3^{[jk]}(\lambda; x, t) &:= U_{0,2j}^{\text{BC}}(\lambda; x, t)U_{0,2k}^{\text{BC}}(\lambda; x, t), \end{aligned}$$

these formulae become

$$(4.32) \quad \begin{aligned} \mathbf{j}_1^{-,1}(\lambda; x, 0) &= a_0(\lambda)\mathbf{U}_0^{\text{BC}}(\lambda; x, 0) \\ &\cdot \int_{-\infty}^x \begin{bmatrix} \psi_1(y)\mu_3^{[12]}(\lambda; y, 0) + \psi_1(y)^*\mu_1^{[12]}(\lambda; y, 0) \\ -\psi_1(y)\mu_3^{[11]}(\lambda; y, 0) - \psi_1(y)^*\mu_1^{[11]}(\lambda; y, 0) \end{bmatrix} dy \end{aligned}$$

and

$$(4.33) \quad \begin{aligned} \mathbf{j}_1^{+,2}(\lambda; x, 0) &= \mathbf{U}_0^{\text{BC}}(\lambda; x, 0) \\ &\cdot \int_{+\infty}^x \begin{bmatrix} \psi_1(y)\mu_3^{[22]}(\lambda; y, 0) + \psi_1(y)^*\mu_1^{[22]}(\lambda; y, 0) \\ -\psi_1(y)\mu_3^{[12]}(\lambda; y, 0) - \psi_1(y)^*\mu_1^{[12]}(\lambda; y, 0) \end{bmatrix} dy. \end{aligned}$$

Since  $a(\lambda; \varepsilon) := \det([\mathbf{j}_1^{-,1}(\lambda; x, t; \varepsilon); \mathbf{j}_1^{+,2}(\lambda; x, t; \varepsilon)])$ , by using the convergent series (4.25)–(4.26) we get that

$$(4.34) \quad a(\lambda; \varepsilon) = a_0(\lambda) + \varepsilon a_1(\lambda) + O(\varepsilon^2), \quad \varepsilon \rightarrow 0,$$

holds uniformly for  $|\lambda| = r$  with  $\text{Im}\{\lambda\} \geq 0$  as well as for  $\lambda \in \mathbb{R}$ ,  $|\lambda| > r$ , where for any  $x \in \mathbb{R}$ ,

$$(4.35) \quad \begin{aligned} a_1(\lambda) &= \det([\mathbf{j}_0^{-,1}(\lambda; x, 0); \mathbf{j}_1^{+,2}(\lambda; x, 0)]) \\ &+ \det([\mathbf{j}_1^{-,1}(\lambda; x, 0); \mathbf{j}_0^{+,2}(\lambda; x, 0)]) \\ &= a_0(\lambda) \int_{-\infty}^{+\infty} (\psi_1(y)\mu_3^{[12]}(\lambda; y, 0) + \psi_1(y)^*\mu_1^{[12]}(\lambda; y, 0)) dy. \end{aligned}$$

Now, formula (4.14) admits a convergent expansion in powers of  $\varepsilon$  of the form  $\mathbf{V}^+(\lambda; \varepsilon) = \mathbf{V}_0^+(\lambda) + \varepsilon \mathbf{V}_1^+(\lambda) + O(\varepsilon^2)$ , and combining the explicit expansion of (4.14) with (4.32)–(4.33) evaluated at  $x = 0$  with  $\mathbf{U}_0^{\text{BC}}(\lambda; 0, 0) = \mathbf{V}_0^+(\lambda)$ , as well as (4.35) gives

$$(4.36) \quad \begin{aligned} \mathbf{V}_1^+(\lambda) &= [a_0(\lambda)^{-1}\mathbf{j}_1^{-,1}(\lambda; 0, 0) - a_0(\lambda)^{-2}a_1(\lambda)\mathbf{j}_0^{-,1}(\lambda; 0, 0); \mathbf{j}_1^{+,2}(\lambda; 0, 0)] \\ &= \mathbf{V}_0^+(\lambda)\mathbf{T}(\lambda)\sigma_3\sigma_1, \end{aligned}$$

where  $\mathbf{T}(\lambda)$  is a matrix of transforms given by

$$(4.37) \quad \mathbf{T}(\lambda) := \begin{bmatrix} -\int_0^{+\infty} (\psi_1\mu_3^{[22]} + \psi_1^*\mu_1^{[22]})dy & \int_0^{+\infty} (\psi_1\mu_3^{[12]} + \psi_1^*\mu_1^{[12]})dy \\ \int_0^{+\infty} (\psi_1\mu_3^{[12]} + \psi_1^*\mu_1^{[12]})dy & \int_{-\infty}^0 (\psi_1\mu_3^{[11]} + \psi_1^*\mu_1^{[11]})dy \end{bmatrix},$$

and in the integrands we use the abbreviated notation  $\psi_1 = \psi_1(y)$ ,  $\psi_1^* = \psi_1(y)^*$ , and  $\mu_{1,3}^{[jk]} = \mu_{1,3}^{[jk]}(\lambda; y, 0)$ . From the exact symmetry  $\mathbf{V}^-(\lambda; \varepsilon) = \mathbf{V}^+(\lambda^*; \varepsilon)^\dagger$  that holds for all  $\varepsilon \in \mathbb{R}$ , we deduce that the coefficients of the convergent power series in  $\varepsilon$  of both sides are also related:  $\mathbf{V}_n^-(\lambda) = \mathbf{V}_n^+(\lambda^*)^\dagger$ , and therefore if  $|\lambda| = r$  and  $\text{Im}\{\lambda\} < 0$ ,

$$(4.38) \quad \mathbf{V}_1^-(\lambda) = \mathbf{V}_1^+(\lambda^*)^\dagger = \sigma_1 \sigma_3 \mathbf{T}(\lambda^*)^\dagger \mathbf{V}_0^+(\lambda^*)^\dagger = \sigma_1 \sigma_3 \mathbf{T}(\lambda^*)^\dagger \mathbf{V}_0^-(\lambda).$$

Next, to calculate the expansion of the matrix  $\mathbf{V}^{\mathbb{R}}(\lambda; \varepsilon) = \mathbf{V}_0^{\mathbb{R}}(\lambda) + \varepsilon \mathbf{V}_1^{\mathbb{R}}(\lambda) + O(\varepsilon^2)$  for  $\lambda \in \Sigma_L \cup \Sigma_R$ , we recall that the jump matrices  $\mathbf{V}^\pm(\lambda; \varepsilon)$  make sense not only on the curves  $\Sigma_\pm$  but also (via analytic continuation) in the part of the closed half-planes  $\overline{\mathbb{C}_\pm}$  exterior to these curves, and that in terms of these the jump matrix on  $\Sigma_L \cup \Sigma_R$  can be written in factored form as  $\mathbf{V}^{\mathbb{R}}(\lambda; \varepsilon) = \mathbf{V}^-(\lambda; \varepsilon) \mathbf{V}^+(\lambda; \varepsilon)$ . Therefore, combining (4.36) and (4.38) shows that

$$(4.39) \quad \begin{aligned} \mathbf{V}_1^{\mathbb{R}}(\lambda) &= \mathbf{V}_0^-(\lambda) \mathbf{V}_0^+(\lambda) \mathbf{T}(\lambda) \sigma_3 \sigma_1 + \sigma_1 \sigma_3 \mathbf{T}(\lambda)^\dagger \mathbf{V}_0^-(\lambda) \mathbf{V}_0^+(\lambda) \\ &= \mathbf{V}_0^{\mathbb{R}}(\lambda) \mathbf{T}(\lambda) \sigma_3 \sigma_1 + \sigma_1 \sigma_3 \mathbf{T}(\lambda)^\dagger \mathbf{V}_0^{\mathbb{R}}(\lambda). \end{aligned}$$

Finally, we make the observation that, whether  $\mathbf{T}(\lambda)$  is evaluated for  $\lambda \in \Sigma_+$  or for  $\lambda \in \Sigma_L \cup \Sigma_R$  (by analytic continuation), its elements can be expressed in terms of half-line  $L^2$  inner products of  $\psi_1$  with squared eigenfunction solutions of the linearized NLS equation (4.4) evaluated at  $t = 0$ . Indeed, taking the squared eigenfunctions  $\mu^{[jk]}(\lambda; x, t)$  and  $v^{[jk]}(\lambda; x, t)$  to be defined from the two columns  $\mathbf{u}^{[1]}(\lambda; x, t)$  and  $\mathbf{u}^{[2]}(\lambda; x, t)$  of the matrix  $\mathbf{U}_{0+}(\lambda; x, t) = \mathbf{U}_0^{\text{BC}}(\lambda; x, t)$  (evaluated as a limit from  $\mathbb{C}_+$  for  $\lambda \in \Sigma_L \cup \Sigma_R$ ) as indicated in (4.11) and (4.13), we see that

$$(4.40) \quad \begin{aligned} T_{11}(\lambda) &= \text{Im} \left\{ \int_0^{+\infty} \psi_1(y) v^{[22]}(\lambda; y, 0)^* dy \right\} \\ &\quad + i \text{Im} \left\{ \int_0^{+\infty} \psi_1(y) \mu^{[22]}(\lambda; y, 0)^* dy \right\}, \\ T_{12}(\lambda) = T_{21}(\lambda) &= -\text{Im} \left\{ \int_0^{+\infty} \psi_1(y) v^{[12]}(\lambda; y, 0)^* dy \right\} \\ &\quad - i \text{Im} \left\{ \int_0^{+\infty} \psi_1(y) \mu^{[12]}(\lambda; y, 0)^* dy \right\}, \\ T_{22}(\lambda) &= -\text{Im} \left\{ \int_{-\infty}^0 \psi_1(y) v^{[11]}(\lambda; y, 0)^* dy \right\} \\ &\quad - i \text{Im} \left\{ \int_{-\infty}^0 \psi_1(y) \mu^{[11]}(\lambda; y, 0)^* dy \right\}. \end{aligned}$$

**Linearizing the Inverse Transform**

The jump matrix for Riemann-Hilbert Problem 1 has been shown to have a convergent power series expansion in  $\varepsilon$ , and here we assume that the convergence holds in the sense of  $L^\infty(\Sigma) \cap L^2(\Sigma)$ . Note that the jump matrix on the branch



cut  $\Sigma_c$  of  $\rho(\lambda)$  is independent of  $\varepsilon$ . For  $\varepsilon \ll 1$  we use the solution  $\mathbf{M}_0(\lambda; x, t)$  for  $\varepsilon = 0$  as a global parametrix; we therefore define a new unknown  $\mathbf{E}(\lambda; x, t)$  to take the place of  $\mathbf{M}(\lambda; x, t)$  by

$$(4.41) \quad \mathbf{E}(\lambda; x, t) := \mathbf{M}(\lambda; x, t)\mathbf{M}_0(\lambda; x, t)^{-1}.$$

It is easy to check that since  $\mathbf{M}(\lambda; x, t)$  and  $\mathbf{M}_0(\lambda; x, t)$  both take continuous boundary values on  $\Sigma_c$  and satisfy the same jump condition there,  $\mathbf{E}(\lambda; x, t)$  is analytic in the domain  $\mathbb{C} \setminus \Sigma'$ , where  $\Sigma' := \Sigma_+ \cup \Sigma_- \cup \Sigma_L \cup \Sigma_R = \Sigma \setminus \Sigma_c$ . Since both factors in (4.41) tend to the identity matrix  $\mathbb{I}$  as  $\lambda \rightarrow \infty$  in each half-plane, the same is true of  $\mathbf{E}(\lambda; x, t)$ . Finally, we may compute the jump conditions satisfied by  $\mathbf{E}(\lambda; x, t)$  on its jump contour:

$$(4.42) \quad \begin{aligned} &\mathbf{E}_+(\lambda; x, t) \\ &= \mathbf{E}_-(\lambda; x, t)\mathbf{M}_{0-}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{V}(\lambda; \varepsilon) \\ &\quad \cdot \mathbf{V}_0(\lambda)^{-1} e^{i\rho(\lambda)(x+\lambda t)\sigma_3} \mathbf{M}_{0-}(\lambda; x, t)^{-1} \\ &= \mathbf{E}_-(\lambda; x, t)\mathbf{U}_{0-}(\lambda; x, t)\mathbf{V}(\lambda; \varepsilon)\mathbf{V}_0(\lambda)^{-1}\mathbf{U}_{0-}(\lambda; x, t)^{-1}, \quad \lambda \in \Sigma'. \end{aligned}$$

Now observe that since  $\mathbf{V}(\lambda; \varepsilon) = \mathbf{V}_0(\lambda) + \varepsilon\mathbf{V}_1(\lambda) + O(\varepsilon^2)$ , the jump matrix for  $\mathbf{E}(\lambda; x, t)$  is a small (in  $L^\infty(\Sigma) \cap L^2(\Sigma)$ ) perturbation of the identity matrix when  $\varepsilon \ll 1$ . We may write it in the form

$$(4.43) \quad \mathbf{U}_{0-}(\lambda; x, t)\mathbf{V}(\lambda; \varepsilon)\mathbf{V}_0(\lambda)^{-1}\mathbf{U}_{0-}(\lambda; x, t)^{-1} = \mathbb{I} + \varepsilon\mathbf{W}(\lambda; x, t; \varepsilon),$$

where

$$(4.44) \quad \mathbf{W}(\lambda; x, t; \varepsilon) = \mathbf{U}_{0-}(\lambda; x, t)\mathbf{V}_1(\lambda)\mathbf{V}_0(\lambda)^{-1}\mathbf{U}_{0-}(\lambda; x, t)^{-1} + O(\varepsilon).$$

This means that  $\mathbf{E}(\lambda; x, t)$  satisfies the conditions of a Riemann-Hilbert problem of *small-norm type*.

To solve for  $\mathbf{E}(\lambda; x, t)$ , we subtract  $\mathbf{E}_-(\lambda; x, t)$  from both sides of (4.42) and use the Plemelj formula:

$$(4.45) \quad \mathbf{E}(\lambda; x, t) = \mathbb{I} + \frac{\varepsilon}{2\pi i} \int_{\Sigma'} \frac{\mathbf{E}_-(z; x, t)\mathbf{W}(z; x, t; \varepsilon)dz}{z - \lambda}.$$

Letting  $\lambda$  tend to the jump contour from the right yields a closed singular integral equation for the boundary value  $\mathbf{E}_-(\lambda; x, t)$ ; we write this equation in the form

$$(4.46) \quad \mathbf{F}(\lambda; x, t) - \varepsilon\mathcal{C}_-[\mathbf{F}(\cdot; x, t)\mathbf{W}(\cdot; x, t; \varepsilon)](\lambda) = \varepsilon\mathcal{C}_-[\mathbf{W}(\cdot; x, t; \varepsilon)](\lambda),$$

where  $\mathbf{F}(\lambda; x, t) := \mathbf{E}_-(\lambda; x, t) - \mathbb{I}$  and where the Cauchy projection operator  $\mathcal{C}_-$  is defined by

$$(4.47) \quad \mathcal{C}_-[\mathbf{H}(\cdot)](\lambda) := \frac{1}{2\pi i} \int_{\Sigma'} \frac{\mathbf{H}(z)dz}{z - \lambda_-}, \quad \lambda \in \Sigma',$$

where the subscript on  $\lambda$  indicates the nontangential boundary value taken from the right side of  $\Sigma'$  by orientation. Once (4.46) is solved,  $\mathbf{E}_-(\lambda; x, t) = \mathbb{I} + \mathbf{F}(\lambda; x, t)$  holds on  $\Sigma'$ , and then (4.45) gives the solution of the Riemann-Hilbert problem. It is well-known that  $\mathcal{C}_-$  is bounded on  $L^2(\Sigma')$  with norm independent of  $\varepsilon$  (the

norm only has to do with the geometry of  $\Sigma'$ ), so the fact that  $\|\mathbf{W}(\cdot; x, t; \varepsilon)\|_2$  and  $\|\mathbf{W}(\cdot; x, t; \varepsilon)\|_\infty$  are both bounded independently of  $\varepsilon \ll 1$  means that (4.46) can be solved by Neumann series with the result that  $\|\mathbf{F}(\cdot; x, t)\|_2 = O(\varepsilon)$ . Using this result in (4.45) gives

$$(4.48) \quad \mathbf{E}(\lambda; x, t) = \mathbb{I} + \frac{\varepsilon}{2\pi i} \int_{\Sigma} \frac{\mathbf{W}(z; x, t; 0) dz}{z - \lambda} + O(\varepsilon^2 \lambda^{-1}).$$

The solution  $\psi(x, t; \varepsilon)$  of the focusing NLS equation corresponding to the jump matrix  $\mathbf{V}(\lambda; x, t; \varepsilon)$  and hence to the initial data  $\psi(x, 0; \varepsilon) = \psi_0(x) + \varepsilon \psi_1(x)$  is (cf. (2.26))

$$(4.49) \quad \begin{aligned} \psi(x, t; \varepsilon) &= 2i \lim_{\lambda \rightarrow \infty} \lambda M_{12}(\lambda; x, t) \\ &= 2i \lim_{\lambda \rightarrow \infty} \lambda (\mathbf{E}(\lambda; x, t) \mathbf{M}_0(\lambda; x, t))_{12} \\ &= 2i \lim_{\lambda \rightarrow \infty} \lambda (E_{11}(\lambda; x, t) M_{0,12}(\lambda; x, t) \\ &\quad + E_{12}(\lambda; x, t) M_{0,22}(\lambda; x, t)) \\ &= \psi_0(x, t) + 2i \lim_{\lambda \rightarrow \infty} \lambda E_{12}(\lambda; x, t) \\ &= \psi_0(x, t) + \varepsilon \psi_1(x, t) + O(\varepsilon^2), \end{aligned}$$

where

$$(4.50) \quad \psi_1(x, t) := -\frac{1}{\pi} \int_{\Sigma'} W_{12}(\lambda; x, t; 0) d\lambda.$$

We now display this formula as a superposition of squared eigenfunction solutions of the INLS equation (4.4) linearized about the solution  $\psi_0(x, t)$  of the NLS equation. First we consider  $\lambda \in \Sigma_L \cup \Sigma_R \subset \mathbb{R}$ . Then, using (4.39) and  $\mathbf{U}_{0+}(\lambda; x, t) = \mathbf{U}_{0-}(\lambda; x, t) \mathbf{V}_0^{\mathbb{R}}(\lambda)$ ,

$$\begin{aligned} W_{12}(\lambda; x, t; 0) &= (\mathbf{U}_{0-}(\lambda; x, t) \mathbf{V}_1^{\mathbb{R}}(\lambda) \mathbf{V}_0^{\mathbb{R}}(\lambda)^{-1} \mathbf{U}_{0-}(\lambda; x, t)^{-1})_{12} \\ &= (\mathbf{U}_{0+}(\lambda; x, t) \mathbf{T}(\lambda) \sigma_3 \sigma_1 \mathbf{U}_{0+}(\lambda; x, t)^{-1})_{12} \\ &\quad + (\mathbf{U}_{0-}(\lambda; x, t) \sigma_1 \sigma_3 \mathbf{T}(\lambda)^\dagger \mathbf{U}_{0-}(\lambda; x, t)^{-1})_{12}. \end{aligned}$$

Recall also the symmetry  $\mathbf{M}(\lambda^*; x, t)^\dagger = \mathbf{M}(\lambda; x, t)$ , which implies the corresponding symmetry  $\mathbf{U}_0(\lambda^*; x, t)^\dagger = \mathbf{U}_0(\lambda; x, t)$ , so letting  $\lambda$  tend to  $\Sigma_L \cup \Sigma_R \subset \mathbb{R}$  from either side, we see that  $\mathbf{U}_{0-}(\lambda; x, t) = \mathbf{U}_{0+}(\lambda; x, t)^{-\dagger}$  holds for  $\lambda \in \Sigma_L \cup \Sigma_R$ . Therefore

$$(4.51) \quad \begin{aligned} W_{12}(\lambda; x, t; 0) &= (\mathbf{U}_{0+}(\lambda; x, t) \mathbf{T}(\lambda) \sigma_3 \sigma_1 \mathbf{U}_{0+}(\lambda; x, t)^{-1})_{12} \\ &\quad + (\mathbf{U}_{0+}(\lambda; x, t)^{-\dagger} \sigma_1 \sigma_3 \mathbf{T}(\lambda)^\dagger \mathbf{U}_{0+}(\lambda; x, t)^\dagger)_{12}. \end{aligned}$$

Letting  $\mathbf{u}^{[k]}(\lambda; x, t)$ ,  $k = 1, 2$ , denote the columns of the matrix  $\mathbf{U}_{0+}(\lambda; x, t)$ , and using the fact that  $\mathbf{U}_{0+}(\lambda; x, t)$  has unit determinant,

$$\begin{aligned} W_{12}(\lambda; x, t; 0) &= \begin{bmatrix} u_1^{[1]}(\lambda; x, t) & u_1^{[2]}(\lambda; x, t) \end{bmatrix} \mathbf{T}(\lambda) \begin{bmatrix} u_1^{[1]}(\lambda; x, t) \\ u_1^{[2]}(\lambda; x, t) \end{bmatrix} \\ &\quad - \begin{bmatrix} u_2^{[1]}(\lambda; x, t)^* & u_2^{[2]}(\lambda; x, t)^* \end{bmatrix} \mathbf{T}(\lambda)^\dagger \begin{bmatrix} u_2^{[1]}(\lambda; x, t)^* \\ u_2^{[2]}(\lambda; x, t)^* \end{bmatrix} \\ &= \begin{bmatrix} u_1^{[1]}(\lambda; x, t) & u_1^{[2]}(\lambda; x, t) \end{bmatrix} \mathbf{T}(\lambda) \begin{bmatrix} u_1^{[1]}(\lambda; x, t) \\ u_1^{[2]}(\lambda; x, t) \end{bmatrix} \\ &\quad - \begin{bmatrix} u_2^{[1]}(\lambda; x, t)^* & u_2^{[2]}(\lambda; x, t)^* \end{bmatrix} \mathbf{T}(\lambda)^* \begin{bmatrix} u_2^{[1]}(\lambda; x, t)^* \\ u_2^{[2]}(\lambda; x, t)^* \end{bmatrix}. \end{aligned}$$

Writing  $\mathbf{T}(\lambda) = \mathbf{A}(\lambda) + i\mathbf{B}(\lambda)$  with the elements of the matrix functions  $\mathbf{A}$  and  $\mathbf{B}$  being real, we arrive at

$$W_{12}(\lambda; x, t; 0) = \sum_{j,k=1}^2 A_{jk}(\lambda) \mu^{[jk]}(\lambda; x, t) + \sum_{j,k=1}^2 B_{jk}(\lambda) \nu^{[jk]}(\lambda; x, t),$$

where  $\mu^{[jk]}(\lambda; x, t)$  and  $\nu^{[jk]}(\lambda; x, t)$  are precisely the squared eigenfunction solutions of the linearized NLS equation (4.4) defined, respectively, by (4.11) and (4.13) in terms of the two columns  $\mathbf{u}^{[1]}(\lambda; x, t)$  and  $\mathbf{u}^{[2]}(\lambda; x, t)$  of the boundary value  $\mathbf{U}_{0+}(\lambda; x, t)$  taken by the matrix  $\mathbf{U}_0(\lambda; x, t)$  from the left on  $\Sigma_L \cup \Sigma_R$ . Since  $\mu^{[jk]}(\lambda; x, t)$  and  $\nu^{[jk]}(\lambda; x, t)$  satisfy (4.4) for each  $\lambda \in \Sigma_L \cup \Sigma_R$ , and since the space of solutions is closed under taking real linear combinations, it follows that the contributions from  $\lambda \in \Sigma_L \cup \Sigma_R \subset \mathbb{R}$  to  $\psi_1(x, t)$  given by the formula (4.50) formally satisfy (4.4).

Next consider the contributions to (4.50) from  $\Sigma_+ \cup \Sigma_-$ . We parametrize  $\Sigma_+$  by  $\lambda = \Lambda(u)$ ,  $\underline{u} < u < \bar{u}$ , and by Schwarz symmetry we similarly parametrize  $\Sigma_-$  by  $\lambda = \Lambda(u)^*$ ,  $\underline{u} < u < \bar{u}$ . Therefore

$$\begin{aligned} (4.52) \quad &\int_{\Sigma_+ \cup \Sigma_-} W_{12}(\lambda; x, t; 0) d\lambda \\ &= \int_{\underline{u}}^{\bar{u}} (W_{12}(\Lambda(u); x, t; 0) \Lambda'(u) + W_{12}(\Lambda(u)^*; x, t; 0) \Lambda'(u)^*) du. \end{aligned}$$

Substituting from (4.36), (4.38), and (4.44), and using

$$\mathbf{U}_{0+}(\lambda; x, t) = \mathbf{U}_{0-}(\lambda; x, t) \mathbf{V}_0(\lambda)$$

gives

$$\begin{aligned}
 & W_{12}(\Lambda(u); x, t; 0)\Lambda'(u) + W_{12}(\Lambda(u)^*; x, t; 0)\Lambda'(u)^* \\
 &= (\mathbf{U}_{0+}(\Lambda(u); x, t)\tilde{\mathbf{T}}(u)\sigma_3\sigma_1\mathbf{U}_{0+}(\Lambda(u); x, t)^{-1})_{12} \\
 (4.53) \quad &+ (\mathbf{U}_{0-}(\Lambda(u)^*; x, t)\sigma_1\sigma_3\tilde{\mathbf{T}}(u)^\dagger\mathbf{U}_{0-}(\Lambda(u)^*; x, t)^{-1})_{12} \\
 &= (\mathbf{U}_{0+}(\Lambda(u); x, t)\tilde{\mathbf{T}}(u)\sigma_3\sigma_1\mathbf{U}_{0+}(\Lambda(u); x, t)^{-1})_{12} \\
 &+ (\mathbf{U}_{0+}(\Lambda(u); x, t)^{-\dagger}\sigma_1\sigma_3\tilde{\mathbf{T}}(u)^\dagger\mathbf{U}_{0+}(\Lambda(u); x, t)^\dagger)_{12},
 \end{aligned}$$

where  $\tilde{\mathbf{T}}(u) := \mathbf{T}(\Lambda(u))\Lambda'(u)$ , and where in the second line we used the Schwarz symmetry  $\mathbf{U}(\lambda^*; x, t) = \mathbf{U}(\lambda; x, t)^{-\dagger}$ , which also has the effect of switching the boundary value from the right of  $\Sigma_-$  to the left of  $\Sigma_+$ . Comparing now with (4.51) we see that the rest of the argument employed in the case of  $\lambda \in \Sigma_L \cup \Sigma_R$  goes through essentially unchanged, with the result that

$$\begin{aligned}
 & W_{12}(\Lambda(u); x, t; 0)\Lambda'(u) + W_{12}(\Lambda(u)^*; x, t; 0)\Lambda'(u)^* \\
 (4.54) \quad &= \sum_{j,k=1}^2 A_{jk}^+(u)\mu^{[jk]}(\Lambda(u); x, t) + \sum_{j,k=1}^2 B_{jk}^+(u)v^{[jk]}(\Lambda(u); x, t)
 \end{aligned}$$

where  $\tilde{\mathbf{T}}(u) = \mathbf{A}^+(u) + i\mathbf{B}^+(u)$  with  $\mathbf{A}^+(\cdot)$  and  $\mathbf{B}^+(\cdot)$  being real matrix-valued functions on  $(0, 1)$ , and where  $\mu^{[jk]}(\lambda; x, t)$  and  $v^{[jk]}(\lambda; x, t)$  are again solutions of (4.4) constructed via (4.11) and (4.13), respectively, but now using the columns  $\mathbf{u}^{[1]}(\lambda; x, t)$  and  $\mathbf{u}^{[2]}(\lambda; x, t)$  of the boundary value  $\mathbf{U}_{0+}(\lambda; x, t)$  taken from the left by the matrix  $\mathbf{U}_0(\lambda; x, t)$  on the arc  $\Sigma_+$ . It therefore follows that again the contributions to (4.50) from  $\Sigma_+ \cup \Sigma_-$  taken together formally satisfy the focusing linearized NLS equation (4.4).

**Summary**

The Cauchy problem for the linearized NLS equation (4.4) (based on the solution  $\psi_0(x, t)$  of the NLS equation (1.1) for which  $\psi_0 - 1 \in L^1(\mathbb{R})$ ) with initial data  $\psi_1(x)$  having sufficiently rapid decay can be obtained formally as follows:

- For  $\lambda \in \Sigma_L \cup \Sigma_+ \cup \Sigma_R$ , let  $\mathbf{M}_{0+}(\lambda; x, t)$  denote the boundary value taken from the left of the simultaneous solution of the direct/inverse problem for the solution  $\psi_0(x, t)$  of (1.1). Taking

$$[\mathbf{u}^{[1]}(\lambda; x, t); \mathbf{u}^{[2]}(\lambda; x, t)] = \mathbf{U}_{0+}(\lambda; x, t) = \mathbf{M}_{0+}(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3},$$

construct from the columns  $\mathbf{u}^{[j]}(\lambda; x, t)$ ,  $j = 1, 2$ , the squared eigenfunction solutions

$$\mu^{[11]}(\lambda; x, t), \quad \mu^{[12]}(\lambda; x, t) = \mu^{[21]}(\lambda; x, t), \quad \mu^{[22]}(\lambda; x, t),$$

from (4.11) and

$$v^{[11]}(\lambda; x, t), \quad v^{[12]}(\lambda; x, t) = v^{[21]}(\lambda; x, t), \quad v^{[22]}(\lambda; x, t),$$

from (4.13) for  $\lambda \in \Sigma_L \cup \Sigma_+ \cup \Sigma_L$ .

- Evaluating the squared eigenfunction solutions at  $t = 0$ , compute the inner products with the Cauchy data  $\psi_1$  for the linearized problem necessary to obtain the transform matrix  $\mathbf{T}(\lambda)$  from (4.40) for  $\lambda \in \Sigma_L \cup \Sigma_+ \cup \Sigma_R$ .
- The solution  $\psi_1(x, t)$  of the linearized Cauchy problem is then  $\psi_1(x, t) = \psi_1^{\mathbb{R}}(x, t) + \psi_1^+(x, t)$ , where we have the explicit integral representations

$$\begin{aligned}
 \psi_1^{\mathbb{R}}(x, t) &:= -\frac{1}{\pi} \sum_{j,k=1}^2 \int_{\Sigma_L \cup \Sigma_R} (\operatorname{Re}\{T_{jk}(\lambda)\}\mu^{[jk]}(\lambda; x, t) \\
 &\quad + \operatorname{Im}\{T_{jk}(\lambda)\}v^{[jk]}(\lambda; x, t))d\lambda \\
 \psi_1^+(x, t) &:= -\frac{1}{\pi} \sum_{j,k=1}^2 \int_{\underline{u}}^{\bar{u}} (\operatorname{Re}\{T_{jk}(\Lambda(u))\Lambda'(u)\}\mu^{[jk]}(\Lambda(u); x, t) \\
 &\quad + \operatorname{Im}\{T_{jk}(\Lambda(u))\Lambda'(u)\}v^{[jk]}(\Lambda(u); x, t))du
 \end{aligned}
 \tag{4.55}$$

with  $\Lambda : (\underline{u}, \bar{u}) \rightarrow \Sigma_+$  being a piecewise-smooth parametrization of  $\Sigma_+$ .

**Simplification of  $\psi_1^{\mathbb{R}}(x, t)$  Using Schwarz Symmetry**

If  $\lambda \in \mathbb{R}$ , the squared eigenfunctions

$$\begin{aligned}
 \mu^{[11]}(\lambda; x, t), \quad \mu^{[12]}(\lambda; x, t) &= \mu^{[21]}(\lambda; x, t), \quad \mu^{[22]}(\lambda; x, t), \\
 v^{[11]}(\lambda; x, t), \quad v^{[12]}(\lambda; x, t) &= v^{[21]}(\lambda; x, t), \quad v^{[22]}(\lambda; x, t),
 \end{aligned}$$

are not linearly independent over the real numbers, leading to some simplifications in the formula for  $\psi_1^{\mathbb{R}}(x, t)$ . To obtain the relations among the solutions, we eliminate  $\mathbf{U}_{0-}(\lambda; x, t)$  between the jump condition  $\mathbf{U}_{0+}(\lambda; x, t) = \mathbf{U}_{0-}(\lambda; x, t)\mathbf{V}_0^{\mathbb{R}}(\lambda)$  and the Schwarz symmetry condition  $\mathbf{U}_{0+}(\lambda; x, t)^\dagger = \mathbf{U}_{0-}(\lambda; x, t)^{-1}$ , both of which hold for  $\lambda \in \Sigma_L \cup \Sigma_R$ . Expressing  $\mathbf{V}_0^{\mathbb{R}}(\lambda)$  in terms of the reflection coefficient  $R_0(\lambda)$  for the unperturbed initial condition  $\psi_0(x, 0)$  by

$$\mathbf{V}_0^{\mathbb{R}}(\lambda) = \begin{bmatrix} 1 + |R_0(\lambda)|^2 & R_0(\lambda)^* \\ R_0(\lambda) & 1 \end{bmatrix},
 \tag{4.56}$$

we systematically write all six of the squared eigenfunction solutions for a given  $\lambda \in \Sigma_L \cup \Sigma_R$  in terms of the others, obtaining a  $6 \times 6$  real-linear homogeneous system. A tedious but straightforward row reduction shows that the rank of this system is exactly 3, containing only the following three independent relations:

$$\begin{aligned}
 \mu^{[11]}(\lambda; x, t) &= 2 \operatorname{Re}\{R_0(\lambda)\}\mu^{[12]}(\lambda; x, t) - (1 + |R_0(\lambda)|^2)\mu^{[22]}(\lambda; x, t), \\
 v^{[11]}(\lambda; x, t) &= -2 \operatorname{Im}\{R_0(\lambda)\}\mu^{[12]}(\lambda; x, t) + (1 + |R_0(\lambda)|^2)v^{[22]}(\lambda; x, t), \\
 v^{[12]}(\lambda; x, t) &= -\operatorname{Im}\{R_0(\lambda)\}\mu^{[22]}(\lambda; x, t) + \operatorname{Re}\{R_0(\lambda)\}v^{[22]}(\lambda; x, t).
 \end{aligned}
 \tag{4.57}$$

It follows that for  $\lambda \in \Sigma_L \cup \Sigma_R$ ,

$$\begin{aligned}
 & \sum_{j,k=1}^2 (\operatorname{Re}\{T_{jk}(\lambda)\}\mu^{[jk]}(\lambda; x, t) + \operatorname{Im}\{T_{jk}(\lambda)\}v^{[jk]}(\lambda; x, t)) \\
 &= ((1 + |R_0(\lambda)|^2) \operatorname{Im}\{T_{11}(\lambda)\} \\
 &+ 2 \operatorname{Re}\{R_0(\lambda)\} \operatorname{Im}\{T_{12}(\lambda)\} + \operatorname{Im}\{T_{22}(\lambda)\})v^{[22]}(\lambda; x, t) \\
 (4.58) \quad &+ (2 \operatorname{Re}\{R_0(\lambda)\} \operatorname{Re}\{T_{11}(\lambda)\} - 2 \operatorname{Im}\{R_0(\lambda)\} \operatorname{Im}\{T_{11}(\lambda)\} \\
 &+ 2 \operatorname{Re}\{T_{12}(\lambda)\})\mu^{[12]}(\lambda; x, t) \\
 &+ (-(1 + |R_0(\lambda)|^2) \operatorname{Re}\{T_{11}(\lambda)\} - 2 \operatorname{Im}\{R_0(\lambda)\} \operatorname{Im}\{T_{12}(\lambda)\} \\
 &+ \operatorname{Re}\{T_{22}(\lambda)\})\mu^{[22]}(\lambda; x, t),
 \end{aligned}$$

where we have also used  $T_{21} = T_{12}$ ,  $\mu^{[jk]} = \mu^{[kj]}$ , and  $v^{[jk]} = v^{[kj]}$ . Now combining the relations (4.57) with (4.40), we observe the following identities:

$$\begin{aligned}
 & (1 + |R_0(\lambda)|^2) \operatorname{Im}\{T_{11}(\lambda)\} + 2 \operatorname{Re}\{R_0(\lambda)\} \operatorname{Im}\{T_{12}(\lambda)\} + \operatorname{Im}\{T_{22}(\lambda)\} \\
 &= -\operatorname{Im}\left\{ \int_{\mathbb{R}} \psi_1(y) \mu^{[11]}(\lambda; y, 0)^* dy \right\},
 \end{aligned}$$

$$2 \operatorname{Re}\{R_0(\lambda)\} \operatorname{Re}\{T_{11}(\lambda)\} - 2 \operatorname{Im}\{R_0(\lambda)\} \operatorname{Im}\{T_{11}(\lambda)\} + 2 \operatorname{Re}\{T_{12}(\lambda)\} = 0,$$

and

$$\begin{aligned}
 & -(1 + |R_0(\lambda)|^2) \operatorname{Re}\{T_{11}(\lambda)\} - 2 \operatorname{Im}\{R_0(\lambda)\} \operatorname{Im}\{T_{12}(\lambda)\} + \operatorname{Re}\{T_{22}(\lambda)\} \\
 &= -\operatorname{Im}\left\{ \int_{\mathbb{R}} \psi_1(y) v^{[11]}(\lambda; y, 0)^* dy \right\}.
 \end{aligned}$$

Therefore, from the definition (4.55) we see that  $\psi_1^{\mathbb{R}}(x, t)$  takes a simpler form:

$$(4.59) \quad \psi_1^{\mathbb{R}}(x, t) = \int_{\Sigma_L \cup \Sigma_R} (C_\mu(\lambda)\mu^{[22]}(\lambda; x, t) + C_\nu(\lambda)v^{[22]}(\lambda; x, t))d\lambda,$$

where the real-valued coefficients  $C_\mu(\lambda)$  and  $C_\nu(\lambda)$  are now given by explicit *whole-line* spectral transforms:

$$\begin{aligned}
 (4.60) \quad & C_\mu(\lambda) := \frac{1}{\pi} \operatorname{Im}\left\{ \int_{\mathbb{R}} \psi_1(y) v^{[11]}(\lambda; y, 0)^* dy \right\}, \\
 & C_\nu(\lambda) := \frac{1}{\pi} \operatorname{Im}\left\{ \int_{\mathbb{R}} \psi_1(y) \mu^{[11]}(\lambda; y, 0)^* dy \right\}.
 \end{aligned}$$

**4.4 Solution of the Linearized Equation for the Peregrine Solution**

Consider the case that the base solution of the NLS equation (1.1) about which we will linearize is the Peregrine breather

$$(4.61) \quad \psi_0(x, t) = \psi_P(x, t) := 1 - 4 \frac{1 + 2i(t - t_0)}{1 + 4(x - x_0)^2 + 4(t - t_0)^2}$$

centered at  $(x_0, t_0)$  in the  $(x, t)$ -plane. Without loss of generality, we will take  $x_0 = 0$  but let  $t_0$  be arbitrary. We write  $\tilde{t} := t - t_0$  below. As we have seen in Section 3.3, this solution can be obtained in the setting of the robust IST by applying a standard Darboux transformation with suitable parameters to the Riemann-Hilbert matrix that represents the background state  $\psi(x, t) \equiv 1$ . The Darboux transformation also supplies the matrix  $\mathbf{U}_0(\lambda; x, t) = \mathbf{M}_0(\lambda; x, t) e^{-i\rho(\lambda)(x+\lambda t)\sigma_3}$  associated with the solution (4.61) for  $x_0 = 0$ . For  $\lambda$  in the domain  $D_+$ , the matrix  $\mathbf{U}_0(\lambda; x, t)$  takes the form

$$(4.62) \quad \mathbf{U}_0(\lambda; x, t) = \frac{n(\lambda)}{\rho(\lambda)(1 + 4x^2 + 4\tilde{t}^2)} \mathbf{Q}(\lambda; x, t) e^{i\rho(\lambda)(x+\lambda t)\sigma_3},$$

where

$$\begin{aligned} Q_{11}(\lambda; x, t) &= 2i\lambda(2\tilde{t} - i) + \rho(\lambda)(2\tilde{t} - i)^2 + 4x(\rho(\lambda)x - i), \\ Q_{12}(\lambda; x, t) &= 4\rho(\lambda)(\lambda - \rho(\lambda))\tilde{t}^2 + 4i\tilde{t} \\ &\quad + (\lambda - \rho(\lambda))(\rho(\lambda) + 2x(\rho(\lambda)x + i)) + 2, \\ Q_{21}(\lambda; x, t) &= 4\rho(\lambda)(\lambda - \rho(\lambda))\tilde{t}^2 - 4i\tilde{t} \\ &\quad + (\lambda - \rho(\lambda))(\rho(\lambda) + 4x(\rho(\lambda)x - i)) + 2, \\ Q_{22}(\lambda; x, t) &= -2i\lambda(2\tilde{t} + i) + \rho(\lambda)(2\tilde{t} + i)^2 + 4x(\rho(\lambda)x + i). \end{aligned}$$

We observe that formula (4.62) admits analytic continuation to the whole complex plane with the branch cut  $\Sigma_c$  of  $\rho$  and  $n$  omitted. This implies that, provided  $\varepsilon\psi_1(x)$  is such that it introduces no further singularities in the solution for the perturbed potential  $\psi_0(x) + \varepsilon\psi_1(x)$ , we may replace the contour  $\Sigma_+ \cup \Sigma_-$  with a “dogbone” contour that lies against the two sides of the branch cut  $\Sigma_c$  except near the endpoints  $\lambda = \pm i$  where it is augmented by small circles of radius  $\delta \ll 1$ .

Such a deformation of  $\Sigma_{\pm}$  also results in  $\Sigma_L$  being extended to  $(-\infty, 0)$  and  $\Sigma_R$  being extended to  $(0, +\infty)$ . The contribution  $\psi_1^{\mathbb{R}}(x, t)$  therefore takes the form

$$(4.63) \quad \psi_1^{\mathbb{R}}(x, t) = \int_{\mathbb{R}} (C_{\mu}(\lambda)\mu^{[22]}(\lambda; x, t) + C_{\nu}(\lambda)\nu^{[22]}(\lambda; x, t))d\lambda,$$

where, applying (4.57) with  $R_0(\lambda) = 0$  for all  $\lambda \in \mathbb{R} \setminus \{0\}$ ,

$$(4.64) \quad \begin{aligned} C_{\mu}(\lambda) &= \frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)\nu^{[22]}(\lambda; y, 0)^* dy \right\}, \\ C_{\nu}(\lambda) &= -\frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)\mu^{[22]}(\lambda; y, 0)^* dy \right\}. \end{aligned}$$

In this situation, we therefore have a very compelling representation of  $\psi_1^{\mathbb{R}}(x, t)$  involving  $L^2(\mathbb{R})$  projections onto the functions  $\mu^{[22]}(\lambda; x, 0)$  and  $v^{[22]}(\lambda; x, 0)$  followed by superposition with the very same solutions now evaluated at  $t \neq 0$ .

Now  $\Sigma_+$  has been deformed to consist of three parts joined end to end: the vertical contour connecting  $\lambda = 0$  with  $\lambda = i(1 - \delta)$  and lying on the left side of  $\Sigma_c$ , followed by the negatively oriented circle of radius  $\delta$  centered at  $\lambda = i$ , followed by the vertical contour connecting  $\lambda = i(1 - \delta)$  with  $\lambda = 0$  and lying on the right side of  $\Sigma_c$ . Let us call the contributions to  $\psi_1^+(x, t)$  arising from integration on the circle  $\psi_1^\circ(x, t)$  and set  $\psi_1^{\parallel}(x, t) := \psi_1^+(x, t) - \psi_1^\circ(x, t)$ . We also adopt the subscript notation  $+/-$  for the boundary values of  $\mathbf{U}_0(\lambda; x, t)$  and the corresponding squared eigenfunctions taken on  $\Sigma_c$  from the left/right half-plane. Then we have the jump condition  $\mathbf{U}_{0+}(is; x, t) = \mathbf{U}_{0-}(is; x, t)i\sigma_1$  for  $0 < s < 1$ , and it follows that the squared eigenfunctions satisfy

$$(4.65) \quad \begin{aligned} \mu_{\pm}^{[11]}(is; x, t) &= -\mu_{\mp}^{[22]}(is; x, t), & \mu_{+}^{[12]}(is; x, t) &= -\mu_{-}^{[12]}(is; x, t), \\ v_{\pm}^{[11]}(is; x, t) &= -v_{\mp}^{[22]}(is; x, t), & v_{+}^{[12]}(is; x, t) &= -v_{-}^{[12]}(is; x, t), \end{aligned}$$

for  $0 < s < 1$ . Taking  $\Lambda(u) = iu$ ,  $0 < u < 1 - \delta$ , for the parametrization of the upward contribution to  $\psi_1^{\parallel}(x, t)$  and  $\Lambda(u) = i(1 - \delta - u)$ ,  $0 < u < 1 - \delta$ , for the parametrization of the downward contribution to  $\psi_1^{\parallel}(x, t)$ , and combining (4.40), (4.55), and (4.65) gives

$$\begin{aligned} \psi_1^{\parallel}(x, t) &= \int_0^{1-\delta} (D_{\mu}^{[11]}(u)\mu_{+}^{[11]}(iu; x, t) + D_{\mu}^{[22]}(u)\mu_{+}^{[22]}(iu; x, t) \\ &\quad + D_v^{[11]}(u)v_{+}^{[11]}(iu; x, t) + D_v^{[22]}(u)v_{+}^{[22]}(iu; x, t))du, \end{aligned}$$

where for  $0 < u < 1 - \delta$ ,

$$(4.66) \quad \begin{aligned} D_{\mu}^{[11]}(u) &:= \frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)\mu_{+}^{[22]}(iu; y, 0)^* dy \right\}, \\ D_{\mu}^{[22]}(u) &:= -\frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)\mu_{+}^{[11]}(iu; y, 0)^* dy \right\}, \\ D_v^{[11]}(u) &:= -\frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)v_{+}^{[22]}(iu; y, 0)^* dy \right\}, \\ D_v^{[22]}(u) &:= \frac{1}{\pi} \operatorname{Im} \left\{ \int_{\mathbb{R}} \psi_1(y)v_{+}^{[11]}(iu; y, 0)^* dy \right\}. \end{aligned}$$

Therefore, the contribution  $\psi_1^{\parallel}(x, t)$  is also represented as a real linear combination of squared eigenfunction solutions of the linearized problem with spectral transform coefficients written in terms of  $L^2(\mathbb{R})$  inner products of the initial condition  $\psi_1$  with the same solutions at  $t = 0$ . The only part of the solution formula  $\psi_1(x, t) = \psi_1^{\mathbb{R}}(x, t) + \psi_1^{\parallel}(x, t) + \psi_1^\circ(x, t)$  left in terms of transform integrals of  $\psi_1$  integrated against the squared eigenfunctions on the half-line is  $\psi_1^\circ(x, t)$ . Taking the limit  $\delta \downarrow 0$  is generally not feasible, because from the factor  $n(\lambda)/\rho(\lambda)$



in (4.62) we see that the quadratic forms  $\mu^{[jk]}(\lambda; x, t)$  and  $\nu^{[jk]}(\lambda; x, t)$  formed from the elements of the matrix  $\mathbf{U}_0(\lambda; x, t)$  blow up proportionally to  $|\lambda - i|^{-3/2}$ , and hence  $\psi_1^\circ(x, t) = O(\delta^{-2})$  as  $\delta \downarrow 0$ .

We can furthermore observe that, due to the simple exponential dependence of the columns of the matrix  $\mathbf{U}_0(\lambda; x, t)$  given in (4.62), for fixed  $t \in \mathbb{R}$  the quadratic forms  $\mu^{[jk]}(\lambda; x, t)$  and  $\nu^{[jk]}(\lambda; x, t)$  are uniformly bounded for  $x \in \mathbb{R}$  whenever  $\rho(\lambda) \in \mathbb{R}$ . The same quadratic forms are uniformly bounded for  $(x, t) \in \mathbb{R}^2$  whenever both  $\rho(\lambda) \in \mathbb{R}$  and  $\lambda \in \mathbb{R}$ . On the other hand, if  $\rho(\lambda) \in \mathbb{R}$  but  $\text{Im}\{\lambda\} \neq 0$ , then the quadratic forms are bounded with respect to  $x$  but grow exponentially in  $t$ . Moreover, the above bounds also hold uniformly as  $\lambda \rightarrow \infty$ . These results imply that under reasonable conditions the component  $\psi_1^\mathbb{R}(x, t)$  of the solution  $\psi_1(x, t)$  of the linearized NLS equation (4.4) is uniformly bounded for all time.

**THEOREM 4.1.** *Suppose that the initial condition  $\psi_1(\cdot)$  is such that  $C_\mu(\cdot)$  and  $C_\nu(\cdot)$  defined by (4.64) are in  $L^1(\mathbb{R})$ . Then  $\psi_1^\mathbb{R}(x, t)$  is uniformly bounded for  $(x, t) \in \mathbb{R}^2$ .*

**PROOF.** This is a straightforward estimate, using the fact that  $\lambda \in \mathbb{R}$  implies  $\rho(\lambda) \in \mathbb{R}$ :

$$\begin{aligned} |\psi_1^\mathbb{R}(x, t)| &\leq \int_{\mathbb{R}} (|C_\mu(\lambda)| |\mu^{[22]}(\lambda; x, t)| + |C_\nu(\lambda)| |\nu^{[22]}(\lambda; x, t)|) d\lambda \\ &\leq K \int_{\mathbb{R}} (|C_\mu(\lambda)| + |C_\nu(\lambda)|) d\lambda, \end{aligned}$$

where  $K$  is a uniform upper bound for  $|\mu^{[22]}(\lambda; x, t)|$  and  $|\nu^{[22]}(\lambda; x, t)|$ .  $\square$

On the other hand, the component  $\psi_1^\parallel(x, t)$  can grow exponentially in  $t$ . *This is for exactly the same reason that the solution of the linearized problem for the background solution  $\psi_0(x, t) \equiv 1$  typically grows exponentially.* The squared eigenfunctions  $\mu_+^{[11]}(iu; x, t)$ ,  $\mu_+^{[22]}(iu; x, t)$ ,  $\nu_+^{[11]}(iu; x, t)$ , and  $\nu_+^{[22]}(iu; x, t)$ , while bounded in  $x$  for  $0 < u < 1$  because  $\rho_+(iu) = -\sqrt{1-u^2}$ , exhibit exponential growth in  $t$  because  $\rho_+(iu)\lambda = -iu\sqrt{1-u^2}$ ; the exponential factors that appear are precisely the same ones as in the case of linearization about  $\psi_0(x, t) \equiv 1$  for which the analogue of (4.62) is simply  $\mathbf{U}_0(\lambda; x, t) = \mathbf{E}(\lambda) e^{i\rho(\lambda)(x+\lambda t)\sigma_3}$ . The linearization about the Peregrine solution  $\psi_0(x, t) = \psi_P(x, t)$  therefore predicts instability of exactly the same sort as in the case  $\psi_0(x, t) \equiv 1$ , namely the (modulational) instability of the background field on which the Peregrine solution rests. However, Theorem 4.1 predicts linearized stability for suitable perturbations of the Peregrine solution  $\psi_0(x, t) = \psi_P(x, t)$  for which the components  $\psi_1^\parallel(x, t)$  and  $\psi_1^\circ(x, t)$  vanish identically (due to the vanishing of the corresponding integral transforms in the integrand).

**Appendix A A Priori Estimates of Jost Solutions**

LEMMA A.1 (A priori estimates). *Suppose that  $\psi(x)$  is defined for all  $x \in \mathbb{R}$  with  $\Delta\psi = \psi - 1 \in L^1(\mathbb{R})$  and let  $|\lambda| \geq s > 1$ . Then we have the following a priori bounds for the indicated columns of  $\mathbf{K}^\pm(\lambda; x)$  given in (1.13):*

$$(A.1) \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{-,1}(\lambda; x)\|_{\ell^1} \leq \left(1 + \frac{1}{\sqrt{s^2 - 1}}\right) e^{c(s)\|\Delta\psi\|_1},$$

$$(A.2) \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{+,2}(\lambda; x)\|_{\ell^1} \leq \left(1 + \frac{1}{\sqrt{s^2 - 1}}\right) e^{c(s)\|\Delta\psi\|_1},$$

where  $\|\mathbf{v}\|_{\ell^1}$  denotes the  $\ell^1$ -norm of a vector  $\mathbf{v} \in \mathbb{C}^2$  and

$$(A.3) \quad c(s) = \frac{1}{2} \left(1 + \frac{s}{\sqrt{s^2 - 1}}\right) + \frac{1}{\sqrt{s^2 - 1}} + \frac{1}{2(s^2 - 1)}.$$

PROOF. Fix  $|\lambda| \geq s$  for some  $s > 1$  with  $\text{Im}\{\lambda\} \geq 0$ . Then from (1.13) it is seen that

$$(A.4) \quad \mathbf{k}^{-,1}(\lambda; x) = \mathbf{e}^1(\lambda) + \int_{-\infty}^x \mathbf{E}(\lambda) \begin{bmatrix} 1 & 0 \\ 0 & e^{2i\rho(\lambda)(x-y)} \end{bmatrix} \mathbf{E}(\lambda)^{-1} \Delta\Psi(y) \mathbf{k}^{-,1}(\lambda; y) dy,$$

$$(A.5) \quad \mathbf{k}^{+,2}(\lambda; x) = \mathbf{e}^2(\lambda) + \int_{+\infty}^x \mathbf{E}(\lambda) \begin{bmatrix} e^{-2i\rho(\lambda)(x-y)} & 0 \\ 0 & 1 \end{bmatrix} \mathbf{E}(\lambda)^{-1} \Delta\Psi(y) \mathbf{k}^{+,2}(\lambda; y) dy,$$

To solve the integral equation (A.4), we set

$$(A.6) \quad \mathbf{C}^{[1]}(\lambda; x, y) = \mathbf{E}(\lambda) \text{diag}(1, e^{2i\rho(\lambda)(x-y)}) \mathbf{E}(\lambda)^{-1} \Delta\Psi(y)$$

and introduce the Neumann series

$$(A.7) \quad \mathbf{k}^{-,1}(\lambda; x) = \sum_{n=0}^{\infty} \boldsymbol{\omega}_n^{[1]}(\lambda; x, t),$$

where

$$(A.8) \quad \boldsymbol{\omega}_0^{[1]} := \mathbf{e}^1(\lambda), \quad \boldsymbol{\omega}_{n+1}^{[1]}(\lambda; x) := \int_{-\infty}^x \mathbf{C}^{[1]}(\lambda; x, y) \boldsymbol{\omega}_n^{[1]}(\lambda; y) dy.$$

First, observe that if  $\lambda = |\lambda| e^{i\theta}$ ,  $\theta \in \mathbb{R}$ , then

$$|\rho(\lambda)| = (|\lambda|^4 + 2|\lambda|^2 \cos(2\theta) + 1)^{1/4},$$

which implies that

$$(A.9) \quad \sqrt{|\lambda|^2 - 1} \leq |\rho(\lambda)| \leq \sqrt{|\lambda|^2 + 1}.$$

Then since  $x \mapsto x/\sqrt{x^2 - 1}$  is a monotone decreasing map for  $x > 1$ , we have

$$(A.10) \quad \frac{|\lambda|}{|\rho(\lambda)|} \leq \frac{s}{\sqrt{s^2 - 1}}.$$

Now let  $\|\cdot\|$  denote the (subordinate) matrix norm induced by the  $\ell^1$ -norm on  $\mathbb{C}^2$  and observe that

$$(A.11) \quad \begin{aligned} \|\mathbf{E}(\lambda)\|\|\mathbf{E}(\lambda)^{-1}\| &= \frac{|\lambda + \rho(\lambda)|}{2|\rho(\lambda)|} (1 + |\lambda - \rho(\lambda)|)^2 \\ &= \frac{1}{|\rho(\lambda)|} + \frac{|\lambda - \rho(\lambda)| + |\lambda + \rho(\lambda)|}{2|\rho(\lambda)|} \\ &\leq \frac{1}{\rho(\lambda)} + \frac{|\lambda|}{2|\rho(\lambda)|} + \frac{1}{2} + \frac{1}{2|\rho(\lambda)|^2} \\ &\leq \frac{1}{2} \left( 1 + \frac{s}{\sqrt{s^2 - 1}} \right) + \frac{1}{\sqrt{s^2 - 1}} + \frac{1}{2(s^2 - 1)} =: c(s). \end{aligned}$$

This uniform bound implies

$$\|\mathbf{C}^{[1]}(\lambda; x, y)\| \leq c(s) \max\{1, e^{-2\text{Im}\{\rho(\lambda)\}(x-y)}\} |\Delta\psi(y)|,$$

which for  $-\infty < y < x$  becomes

$$\|\mathbf{C}^{[1]}(\lambda; x, y)\| \leq c(s) |\Delta\psi(y)|.$$

Then for the  $n^{\text{th}}$  iterate in the Neumann series we have

$$\begin{aligned} \|\omega_n^{[1]}(\lambda; x)\|_{\ell^1} &\leq \|\mathbf{e}^1(\lambda)\|_{\ell^1} \int_{\pm\infty}^x \int_{\pm\infty}^{x_1} \cdots \int_{\pm\infty}^{x_{n-1}} \|\mathbf{C}^{[1]}(\lambda; x, x_1)\| \cdots \|\mathbf{C}^{[1]}(\lambda; x_{n-1}, x_n)\| dx_n dx_{n-1} \cdots dx_1 \\ &\leq \|\mathbf{e}^1(\lambda)\|_{\ell^1} \int_{\pm\infty}^x \int_{\pm\infty}^{x_1} \cdots \int_{\pm\infty}^{x_{n-1}} |\Delta\psi(x_1)| |\Delta\psi(x_2)| \cdots |\Delta\psi(x_n)| \cdot c(s)^n dx_n \cdots dx_1. \end{aligned}$$

Setting

$$\tau(x_n) = c(s) \int_{\pm\infty}^{x_n} |\Delta\psi(x_{n-1})| dx_{n-1}, \quad x_0 = x,$$

and  $\tau_n := \tau(x_n)$  gives

$$(A.12) \quad \begin{aligned} \|\omega_n^{[1]}(\lambda; x)\|_{\ell^1} &\leq \|\mathbf{e}^1(\lambda)\|_{\ell^1} \int_{\pm\infty}^{\tau(x)} \int_{\pm\infty}^{\tau_1} \cdots \int_{\pm\infty}^{\tau_{n-1}} d\tau_n d\tau_{n-1} \cdots d\tau_1 \\ &= \|\mathbf{e}^1(\lambda)\|_{\ell^1} \frac{\tau(x)^n}{n!} \leq \|\mathbf{e}^1(\lambda)\|_{\ell^1} \frac{c(s)^n \|\Delta\psi\|_1^n}{n!}. \end{aligned}$$

Therefore the Neumann series converges absolutely and uniformly in  $x$  for  $|\lambda| > s$  restricted to  $\text{Im}\{\lambda\} \geq 0$ , and we obtain

$$(A.13) \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{-,1}(\lambda; x)\|_{\ell^1} \leq \left( 1 + \frac{1}{\sqrt{s^2 - 1}} \right) e^{c(s)\|\Delta\psi\|_1}.$$

Similarly, because  $\|\mathbf{e}^2(\lambda)\|_{\ell^1} = \|\mathbf{e}^1(\lambda)\|_{\ell^1}$  and  $\|\text{diag}(e^{-2i\rho(x-y)}, 1)\| \leq 1$  on  $x < y < +\infty$ , we have

$$(A.14) \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{+,2}(\lambda; x)\|_{\ell^1} \leq \left(1 + \frac{1}{\sqrt{s^2 - 1}}\right) e^{c(s)\|\Delta\psi\|_1}. \quad \square$$

### Appendix B Asymptotic Properties of $a(\lambda)$ for Large $\lambda$

In what follows we make no attempt to optimize the value of  $r[\Delta\psi]$  to be defined in (B.25) below.

PROOF OF LEMMA 1.2. Since  $a(\lambda; t) = a(\lambda; 0) =: a(\lambda)$ , we omit the time dependence of the quantities involved in the proof. We suppose that  $\psi(x; 0)$  is not identically 1 since otherwise  $a(\lambda) \equiv 1$ , and we also a priori assume that  $|\lambda| \geq 2$ . Then, we apply Lemma A.1 with  $c = 2$ . From the following obvious inequalities,

$$(B.1) \quad 1 + \frac{1}{\sqrt{2^2 - 1}} < 2 \quad \text{and} \quad c(2) < 2,$$

where for  $c(s)$  we refer to the definition (A.11), Lemma A.1 implies the a priori bounds

$$(B.2) \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{-,1}(\lambda; x)\|_{\ell^1} \leq 2e^{2\|\Delta\psi\|_1}, \quad \sup_{x \in \mathbb{R}} \|\mathbf{k}^{+,2}(\lambda; x)\|_{\ell^1} \leq 2e^{2\|\Delta\psi\|_1}.$$

Note that  $\mathbf{E}(\lambda) = n(\lambda)(\mathbb{I} + i(\lambda - \rho(\lambda))\sigma_1)$ . We write

$$\mathbf{k}^{-,1}(\lambda; x) =: [u_1(\lambda; x) \quad v_1(\lambda; x) e^{2i\rho(\lambda)x}]^\top.$$

Then the second component of (A.4) reads

$$(B.3) \quad \begin{aligned} & e^{2i\rho(\lambda)x} v_1(\lambda; x) \\ &= in(\lambda)(\lambda - \rho(\lambda)) \\ &+ n(\lambda)^2 \left\{ - \int_{-\infty}^x e^{2i\rho(\lambda)(x-y)} \Delta\psi(y)^* u_1(\lambda; y) dy \right. \\ &\quad + i(\lambda - \rho(\lambda)) \int_{-\infty}^x (1 - e^{2i\rho(\lambda)(x-y)}) \Delta\psi(y) e^{2i\rho(\lambda)y} v_1(\lambda; y) dy \\ &\quad \left. - (\lambda - \rho(\lambda))^2 \int_{-\infty}^x \Delta\psi(y)^* u_1(\lambda; y) dy \right\}. \end{aligned}$$

Note that since  $\Delta\psi, \psi' \in L^1(\mathbb{R})$ , we have  $\psi, \Delta\psi \in L^\infty(\mathbb{R})$ . Using this together with the boundedness of  $u_1(\lambda; y)$  for  $-\infty < y \leq x$ ,  $\text{Im } \rho(\lambda) > 0$ , integrating by

parts in the first integral, and using the differential equation (1.8) gives

$$\begin{aligned}
 & e^{2i\rho(\lambda)x}v_1(\lambda; x) \\
 &= in(\lambda)(\lambda - \rho(\lambda)) \\
 \text{(B.4)} \quad & + n(\lambda)^2 \left\{ \frac{\Delta\psi(x)^*u_1(\lambda; x)}{2i\rho(\lambda)} + \frac{1}{2i\rho(\lambda)} \mathcal{T}_1[\Delta\psi](e^{2i\rho(\lambda)\cdot}v_1(\lambda; \cdot))(x) \right. \\
 & \quad - \frac{1}{2i\rho(\lambda)} \int_{-\infty}^x e^{2i\rho(\lambda)(x-y)}(\psi'(y)^* + i(\rho(\lambda) - \lambda)\Delta\psi(y)^*)u_1(\lambda; y)dy \\
 & \quad \left. - (\lambda - \rho(\lambda))^2 \int_{-\infty}^x \Delta\psi(y)^*u_1(\lambda; y)dy \right\},
 \end{aligned}$$

where  $\mathcal{T}_1[\Delta\psi]$  denotes the following Volterra integral operator on  $L^\infty(\mathbb{R})$ :

$$\begin{aligned}
 & \mathcal{T}_1[\Delta\psi]h(x) \\
 \text{(B.5)} \quad & := - \int_{-\infty}^x e^{2i\rho(\lambda)(x-y)} \Delta\psi(y)^* \psi(y)h(y)dy \\
 & \quad - 2\rho(\lambda)(\lambda - \rho(\lambda)) \int_{-\infty}^x (1 - e^{2i\rho(\lambda)(x-y)}) \Delta\psi(y)h(y)dy.
 \end{aligned}$$

The assumption  $|\lambda| \geq 2$  implies that  $|\lambda - \rho(\lambda)||\rho(\lambda)| < 1$  and hence the bound

$$\text{(B.6)} \quad \|\mathcal{T}_1[\Delta\psi]\|_\infty \leq \|\Delta\psi\|_1(\|\psi\|_\infty + 4)$$

on the operator norm of  $\mathcal{T}_1[\Delta\psi]$  on  $L^\infty(\mathbb{R})$ . Thus, with the uniform bounds (B.2) and the fact that  $|n(\lambda)|^2 \leq 1 + |\rho(\lambda)|^{-1} \leq 1 + 1/\sqrt{3} < 2$  when  $|\lambda| \geq 2$ , if  $|\rho(\lambda)| > \|\Delta\psi\|_1(\|\psi\|_\infty + 4)$ , we then have

$$\begin{aligned}
 & \|e^{2i\rho(\lambda)\cdot}v_1(\lambda; \cdot)\|_\infty \\
 & \leq \left\| \left( \mathcal{I} - \frac{n(\lambda)^2}{2i\rho(\lambda)} \mathcal{T}_1[\Delta\psi] \right)^{-1} \right\|_\infty \\
 \text{(B.7)} \quad & \cdot \frac{1}{|\rho(\lambda)|} \left( 2 + \left[ \|\Delta\psi\|_\infty + \|\psi'\|_1 + \frac{2\|\Delta\psi\|_1}{|\rho(\lambda)|} \right] 2e^{2\|\Delta\psi\|_1} \right) \\
 & \leq \left( 1 - \frac{\|\Delta\psi\|_1(\|\psi\|_\infty + 4)}{|\rho(\lambda)|} \right)^{-1} \\
 & \cdot \frac{1}{|\rho(\lambda)|} \left( 2 + \left[ \|\Delta\psi\|_\infty + \|\psi'\|_1 + \frac{2\|\Delta\psi\|_1}{|\rho(\lambda)|} \right] 2e^{2\|\Delta\psi\|_1} \right),
 \end{aligned}$$

where we have again used the bound  $|\lambda - \rho(\lambda)||\rho(\lambda)| < 1$ , together with

$$|\lambda - \rho(\lambda)|^2|\rho(\lambda)|^2 < 1/2 \quad \text{and} \quad |n(\lambda)|^2 < 2 < 4$$

for  $|\lambda| \geq 2$ .

Similarly, the discrepancy from unity for the first component of (A.4) reads

$$\begin{aligned}
 u_1(\lambda; x) - 1 = & n(\lambda) - 1 \\
 & + n(\lambda)^2 \left\{ \int_{-\infty}^x \Delta\psi(y) e^{2i\rho(\lambda)y} v_1(\lambda; y) dy \right. \\
 & - i(\lambda - \rho(\lambda)) \int_{-\infty}^x (e^{2i\rho(\lambda)(x-y)} - 1) \Delta\psi(y)^* u_1(\lambda; y) dy \\
 & \left. + (\lambda - \rho(\lambda))^2 \int_{-\infty}^x e^{2i\rho(\lambda)(x-y)} \Delta\psi(y) e^{2i\rho(\lambda)y} v_1(\lambda; y) dy \right\},
 \end{aligned}$$

from which we directly obtain the bound

$$\begin{aligned}
 \|u_1(\lambda; \cdot) - 1\|_\infty \leq & \frac{1}{2|\rho(\lambda)|^2} \\
 (B.8) \quad & + 2 \left( \|e^{2i\rho(\lambda)\cdot} v_1(\lambda; \cdot)\|_\infty \left( 1 + \frac{1}{2|\rho(\lambda)|^2} \right) \|\Delta\psi\|_1 \right. \\
 & \left. + \frac{2}{|\rho(\lambda)|} \|\Delta\psi\|_1 2e^{2\|\Delta\psi\|_1} \right).
 \end{aligned}$$

We now proceed to get the analogous estimates for the elements of the other column:

$$(B.9) \quad \mathbf{k}^{+,2}(\lambda; x) =: [e^{-2i\rho(\lambda)x} u_2(\lambda; x) \quad v_2(\lambda; x)]^\top.$$

The first component of (A.5) reads:

$$\begin{aligned}
 & e^{-2i\rho(\lambda)x} u_2(\lambda; x) \\
 = & in(\lambda)(\lambda - \rho(\lambda)) \\
 & + n(\lambda)^2 \left\{ \int_{+\infty}^x e^{-2i\rho(\lambda)(x-y)} \Delta\psi(y) v_2(\lambda; y) dy \right. \\
 & - i(\lambda - \rho(\lambda)) \int_{+\infty}^x (1 - e^{-2i\rho(\lambda)(x-y)}) \Delta\psi(y)^* e^{-2i\rho(\lambda)y} u_2(\lambda; y) dy \\
 & \left. + (\lambda - \rho(\lambda))^2 \int_{+\infty}^x \Delta\psi(y) v_2(\lambda; y) dy \right\}.
 \end{aligned}$$

As before, integrating by parts in the first integral and using the differential equation (1.8) to eliminate  $v_{2y}(\lambda; y)$  gives

$$\begin{aligned} & e^{-2i\rho(\lambda)x} u_2(\lambda; x) \\ &= in(\lambda)(\lambda - \rho(\lambda)) \\ &+ n(\lambda)^2 \left\{ \frac{\Delta\psi(x)v_2(\lambda; x)}{2i\rho(\lambda)} + \frac{1}{2i\rho(\lambda)} \mathcal{F}_2[\Delta\psi](e^{-2i\rho(\lambda)\cdot} u_2(\lambda; \cdot))(x) \right. \\ &\quad - \frac{1}{2i\rho(\lambda)} \int_{+\infty}^x e^{-2i\rho(\lambda)(x-y)} (\psi'(y) + i(\lambda - \rho(\lambda))\Delta\psi(y)) v_2(\lambda; y) dy \\ &\quad \left. + (\lambda - \rho(\lambda))^2 \int_{+\infty}^x \Delta\psi(y) v_2(\lambda; y) dy \right\}, \end{aligned}$$

where

$$\begin{aligned} & \mathcal{F}_2[\Delta\psi]h(x) \\ (B.10) \quad & := \int_{+\infty}^x e^{-2i\rho(\lambda)(x-y)} \Delta\psi(y) \psi(y)^* h(y) dy \\ & + 2\rho(\lambda)(\lambda - \rho(\lambda)) \int_{+\infty}^x (1 - e^{-2i\rho(\lambda)(x-y)}) \Delta\psi(y)^* h(y) dy. \end{aligned}$$

Again the assumption  $|\lambda| \geq 2$  implies the bound

$$(B.11) \quad \|\mathcal{F}_2[\Delta\psi]\|_\infty \leq \|\Delta\psi\|_1 (\|\psi\|_\infty + 4)$$

on the operator norm of  $\mathcal{F}_2[\Delta\psi]$  on  $L^\infty(\mathbb{R})$ . Thus, if  $|\rho(\lambda)| > \|\Delta\psi\|_1 (\|\psi\|_\infty + 4)$ , then we obtain

$$\begin{aligned} & \|e^{-2i\rho(\lambda)\cdot} u_2(\lambda; \cdot)\|_\infty \\ (B.12) \quad & \leq \left( 1 - \frac{\|\Delta\psi\|_1 (\|\psi\|_\infty + 4)}{|\rho(\lambda)|} \right)^{-1} \\ & \cdot \frac{1}{|\rho(\lambda)|} \left( 2 + \left[ \|\Delta\psi\|_\infty + \|\psi'\|_1 + \frac{2\|\Delta\psi\|_1}{|\rho(\lambda)|} \right] 2e^{2\|\Delta\psi\|_1} \right), \end{aligned}$$

for  $|\lambda| \geq 2$ . Similarly, for the discrepancy from unity for the second component of (A.5) we have

$$\begin{aligned} &v_2(\lambda; x) - 1 \\ &= n(\lambda) - 1 \\ &\quad - n(\lambda)^2 \left\{ \int_{+\infty}^x \Delta\psi(y)^* e^{-2i\rho(\lambda)y} u_2(\lambda; y) dy \right. \\ &\quad \quad + i(\lambda - \rho(\lambda)) \int_{+\infty}^x (e^{-2i\rho(\lambda)(x-y)} - 1) \Delta\psi(y) v_2(\lambda; y) dy \\ &\quad \quad \left. - (\lambda - \rho(\lambda))^2 \int_{+\infty}^x e^{-2i\rho(\lambda)(x-y)} \Delta\psi(y)^* e^{-2i\rho(\lambda)y} u_2(\lambda; y) dy \right\} \end{aligned}$$

from which we can directly obtain the estimate

$$\begin{aligned} &\|v_2(\lambda; \cdot) - 1\|_\infty \\ &\leq \frac{1}{2|\rho(\lambda)|^2} \\ \text{(B.13)} \quad &+ 2 \left( \|e^{-2i\rho(\lambda)\cdot} u_2(\lambda; \cdot)\|_\infty \left( 1 + \frac{1}{2|\rho(\lambda)|^2} \right) \|\Delta\psi\|_1 \right. \\ &\quad \left. + \frac{2}{|\rho(\lambda)|} \|\Delta\psi\|_1 2e^{2\|\Delta\psi\|_1} \right). \end{aligned}$$

For convenience we denote the (identical) bound in the estimates (B.7) and (B.12) by  $m(\lambda)$ , namely

$$\text{(B.14)} \quad \max\{\|e^{-2i\rho(\lambda)\cdot} u_2(\lambda; \cdot)\|_\infty, \|e^{2i\rho(\lambda)\cdot} v_1(\lambda; \cdot)\|_\infty\} \leq m(\lambda),$$

where

$$\begin{aligned} m(\lambda) := &\left( 1 - \frac{\|\Delta\psi\|_1(\|\psi\|_\infty + 4)}{|\rho(\lambda)|} \right)^{-1} \\ &\cdot \frac{1}{|\rho(\lambda)|} \left( 2 + \left[ \|\Delta\psi\|_\infty + \|\psi'\|_1 + \frac{2\|\Delta\psi\|_1}{|\rho(\lambda)|} \right] 2e^{2\|\Delta\psi\|_1} \right), \end{aligned}$$

which is well-defined for  $|\rho(\lambda)| > \|\Delta\psi\|_1(\|\psi\|_\infty + 4)$  and  $m(\lambda) > 0$ , with  $m(\lambda) = O(\lambda^{-1})$  as  $\lambda \rightarrow \infty$  since  $\rho(\lambda) = O(\lambda)$  as  $\lambda \rightarrow \infty$ . Then from (B.8) and (B.13) we obtain

$$\begin{aligned} &\max\{\|u_1(\lambda; \cdot) - 1\|_\infty, \|v_2(\lambda; \cdot) - 1\|_\infty\} \\ \text{(B.15)} \quad &\leq \frac{1}{2|\rho(\lambda)|^2} + 2m(\lambda)\|\Delta\psi\|_1 + \frac{m(\lambda)\|\Delta\psi\|_1}{|\rho(\lambda)|^2} \\ &\quad + \frac{8\|\Delta\psi\|_1 e^{2\|\Delta\psi\|_1}}{|\rho(\lambda)|} =: q(\lambda), \end{aligned}$$



where clearly  $q(\lambda) > 0$  and  $q(\lambda) = O(\lambda^{-1})$  as  $\lambda \rightarrow \infty$ . Then from definition (1.15) of  $a(\lambda)$  and the estimates (B.14) and (B.15), it follows that

$$(B.16) \quad |a(\lambda) - 1| \leq 2q(\lambda) + q(\lambda)^2 + m(\lambda)^2 = O(\lambda), \quad \lambda \rightarrow \infty, \quad \text{Im}\{\rho(\lambda)\} > 0,$$

which proves the first claim. To guarantee that  $|a(\lambda)| > \frac{1}{2}$  outside a disk centered at the origin in the  $\lambda$ -plane, it suffices to ensure that  $|q(\lambda)| \leq \frac{1}{8}$  and  $m(\lambda) \leq \frac{1}{4}$  by the inequality in (B.16). First, from (B.15) it is seen that if

$$(B.17) \quad |\rho(\lambda)| \geq \max\{4, 256\|\Delta\psi\|_1 e^{2\|\Delta\psi\|_1}\}$$

and

$$(B.18) \quad m(\lambda) \leq \min\left\{\frac{1}{4}, \frac{1}{33\|\Delta\psi\|_1}\right\},$$

then  $q(\lambda) \leq \frac{1}{8}$ . Note also that we have the condition

$$(B.19) \quad |\rho(\lambda)| > \|\Delta\psi\|_1(\|\psi\|_\infty + 4)$$

for  $m(\lambda)$  to be well-defined, and it remains to obtain a lower bound on  $|\rho(\lambda)|$  to guarantee (B.18).

We first strengthen (B.19) and demand

$$(B.20) \quad |\rho(\lambda)| \geq 4\|\Delta\psi\|_1(\|\psi\|_\infty + 4).$$

Then using (B.17) together with (B.20) we obtain

$$(B.21) \quad m(\lambda) \leq \frac{4}{3|\rho(\lambda)|} \left(2 + \frac{1}{64} + 2(\|\Delta\psi\|_\infty + \|\psi'\|_1) e^{2\|\Delta\psi\|_1}\right).$$

Thus, to ensure (B.18), one needs to choose  $\lambda$  such that

$$(B.22) \quad |\rho(\lambda)| \geq \frac{1}{3} \max\{16, 132\|\Delta\psi\|_1\} \cdot \left(2 + \frac{1}{64} + 4\|\psi'\|_1 e^{2\|\Delta\psi\|_1}\right)$$

since  $\|\Delta\psi\|_\infty \leq \|\psi'\|_1$ . Since  $\|\Delta\psi\|_1(\|\psi\|_\infty + 4) \leq \|\Delta\psi\|_1(\|\psi'\|_1 + 5)$ , to guarantee all of the conditions (B.17), (B.20), and (B.22) simultaneously,  $\rho(\lambda)$  needs to satisfy

$$(B.23) \quad |\rho(\lambda)| \geq \max\left\{4, 256\|\Delta\psi\|_1 e^{2\|\Delta\psi\|_1}, 4\|\Delta\psi\|_1(\|\psi'\|_1 + 5), h[\Delta\psi], \frac{33}{4}\|\Delta\psi\|_1 h[\Delta\psi]\right\},$$

where

$$(B.24) \quad h[\Delta\psi] := \frac{16}{3}(2 + 2^{-6} + 4\|\psi'\|_1 e^{2\|\Delta\psi\|_1}).$$

Note that the initial assumption  $|\lambda| \geq 2$  is absorbed in this condition. Using (A.9), it is enough to choose

$$(B.25) \quad |\lambda| \geq \sqrt{1 + \max\left\{4, 256\|\Delta\psi\|_1 e^{2\|\Delta\psi\|_1}, d_1[\Delta\psi], h[\Delta\psi], \frac{33}{4}\|\Delta\psi\|_1 h[\Delta\psi]\right\}^2} \\ =: r[\Delta\psi],$$

where  $d_1[\Delta\psi] := 4\|\Delta\psi\|_1(\|\psi'\|_1 + 5)$  to guarantee (B.23). With this choice  $m(\lambda) \leq \frac{1}{4}$  and  $q(\lambda) \leq \frac{1}{8}$ , and hence

$$(B.26) \quad |a(\lambda) - 1| \leq \frac{21}{64} < \frac{1}{2},$$

which implies the second claim.  $\square$

**Acknowledgment.** The authors are grateful to Robbie Buckingham, Liming Ling, and an anonymous referee for useful feedback and discussions. Both authors were partially supported by the National Science Foundation under grant DMS-1513054. In addition, D. B. was supported by an AMS-Simons Travel Grant.

### Bibliography

- [1] Ablowitz, M. J.; Kaup, D. J.; Newell, A. C.; Segur, H. The inverse scattering transform-Fourier analysis for nonlinear problems. *Studies in Appl. Math.* **53** (1974), no. 4, 249–315.
- [2] Akhmediev, N.; Ankiewicz, A.; Soto-Crespo, J. M. Rogue waves and rational solutions of the nonlinear Schrödinger equation. *Phys. Rev. E.* **80** (2009), no. 2, 026601. doi:10.1103/PhysRevE.80.026601
- [3] Akhmediev, N. N.; Korneev, V. I. Modulation instability and periodic solutions of the nonlinear Schrödinger equation. *Teoret. Mat. Fiz.* **69** (1986), no. 2, 189–194.
- [4] Ankiewicz, A.; Clarkson, P. A.; Akhmediev, N. Rogue waves, rational solutions, the patterns of their zeros and integral relations. *J. Phys. A* **43** (2010), no. 12, 122002, 9 pp. doi:10.1088/1751-8113/43/12/122002
- [5] Beals, R.; Coifman, R. R. Scattering and inverse scattering for first order systems. *Comm. Pure Appl. Math.* **37** (1984), no. 1, 39–90. doi:10.1002/cpa.3160370105
- [6] Bertola, M.; Tovbis, A. Universality for the focusing nonlinear Schrödinger equation at the gradient catastrophe point: rational breathers and poles of the tritronquée solution to Painlevé I. *Comm. Pure Appl. Math.* **66** (2013), no. 5, 678–752. doi:10.1002/cpa.21445
- [7] Bilman, D.; Ling, L.; Miller, P. D. Extreme superposition: rogue waves of infinite order and the Painlevé-III hierarchy. Preprint, 2018. arXiv:1806.00545 [nlin.SI]
- [8] Biondini, G.; Kovačič, G. Inverse scattering transform for the focusing nonlinear Schrödinger equation with nonzero boundary conditions. *J. Math. Phys.* **55** (2014), no. 3, 031506, 22 pp. doi:10.1063/1.4868483
- [9] Biondini, G.; Mantzavinos, D. Long-time asymptotics for the focusing nonlinear Schrödinger equation with nonzero boundary conditions at infinity and asymptotic stage of modulational instability. *Comm. Pure Appl. Math.* **70** (2017), no. 12, 2300–2365. doi:10.1002/cpa.21701
- [10] Chabchoub, A.; Fink, M. Time-reversal generation of rogue waves. *Phys. Rev. Lett.* **112** (2014), no. 12, 124101, 5 pp. doi:10.1103/PhysRevLett.112.124101
- [11] Chabchoub, A.; Hoffmann, N. P.; Akhmediev, N. Rogue wave observation in a water wave tank. *Phys. Rev. Lett.* **106** (2011), no. 20, 204502, 4 pp. doi:10.1103/PhysRevLett.106.204502

- [12] Coddington, E. A.; Levinson, N. *Theory of ordinary differential equations*. McGraw-Hill, New York–Toronto–London, 1955.
- [13] Cousins, W.; Sapsis, T. P. Reduced order prediction of rare events in unidirectional nonlinear water waves. *J. Fluid Mech.* **790** (2016), 368–388. doi:10.1017/jfm.2016.13
- [14] Deift, P.; Zhou, X. Direct and inverse scattering on the line with arbitrary singularities. *Comm. Pure Appl. Math.* **44** (1991), no. 5, 485–533. doi:10.1002/cpa.3160440502
- [15] Deift, P.; Zhou, X. A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the mKdV equation. *Ann. of Math. (2)* **137** (1993), no. 2, 295–368. doi:10.2307/2946540
- [16] Demontis, F.; Prinari, P.; van der Mee, C.; Vitale, F. The inverse scattering transform for the defocusing nonlinear Schrödinger equations with nonzero boundary conditions. *Stud. Appl. Math.* **131** (2013), no. 1, 1–40. doi:10.1111/j.1467-9590.2012.00572.x
- [17] Dubard, P.; Gaillard, P.; Klein, C.; Matveev, V. On multi-rogue wave solutions of the NLS equation and positon solutions of the KdV equation. *Eur. Phys. J. Spec. Top.* **185** (2010), no. 1, 247–258. doi:10.1140/epjst/e2010-01252-9
- [18] Faddeev, L. D.; Takhtajan, L. A. *Hamiltonian methods in the theory of solitons*. Springer Series in Soviet Mathematics. Springer, Berlin, 1987. doi:10.1007/978-3-540-69969-9
- [19] Fokas, A. S. A unified transform method for solving linear and certain nonlinear PDEs. *Proc. Roy. Soc. London Ser. A* **453** (1997), no. 1962, 1411–1443. doi:10.1098/rspa.1997.0077
- [20] Fokas, A. S. *A unified approach to boundary value problems*. CBMS-NSF Regional Conference Series in Applied Mathematics, 78. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2008. doi:10.1137/1.9780898717068
- [21] Gaillard, P. Degenerate determinant representation of solutions of the nonlinear Schrödinger equation, higher order Peregrine breathers and multi-rogue waves. *J. Math. Phys.* **54** (2013), no. 1, 013504, 32 pp. doi:10.1063/1.4773096
- [22] Gaillard, P. Higher order Peregrine breathers, their deformations and multi-rogue waves. *J. Phys. Conf. Ser.* **482** (2014), no. 1, 012016. doi:10.1088/1742-6596/482/1/012016
- [23] Guo, B.; Ling, L.; Liu, Q. P. Nonlinear Schrödinger equation: generalized Darboux transformation and rogue wave solutions. *Phys. Rev. E* **85** (2012), no. 2, 026607, 9 pp. doi:10.1103/PhysRevE.85.026607
- [24] Guo, B.; Tian, L.; Yan, Z.; Ling, L.; Wang, Y.-F. *Rogue waves. Mathematical theory and applications in physics*. De Gruyter, Berlin, 2017.
- [25] Hammani, K.; Kibler, B.; Finot, C.; Morin, P.; Fatome, J.; Dudley, J. M.; Millot, G. Peregrine soliton generation and breakup in standard telecommunications fiber. *Opt. Lett.* **36** (2011), no. 2, 112–114. doi:10.1364/OL.36.000112
- [26] He, J. S.; Zhang, H. R.; Wang, L. H.; Porsezian, K.; Fokas, A. S. Generating mechanism for higher-order rogue waves. *Phys. Rev. E* **87** (2013), no. 5, 052914. doi:10.1103/PhysRevE.87.052914
- [27] Kaup, D. J. Closure of the squared Zakharov–Shabat eigenstates. *J. Math. Anal. Appl.* **54** (1976), no. 3, 849–864. doi:10.1016/0022-247X(76)90201-8
- [28] Kharif, C.; Pelinovsky, E.; Slunyaev, A. *Rogue waves in the ocean*. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer, Berlin, 2009.
- [29] Kibler, B.; Fatome, J.; Finot, C.; Millot, G.; Dias, F.; Genty, G.; Akhmediev, N.; Dudley, J. M. The Peregrine soliton in nonlinear fibre optics. *Nat. Phys.* **6** (2010), no. 10, 790–795. doi:10.1038/nphys1740
- [30] Kuznetsov, E. A. Solitons in parametrically unstable plasma. *Dokl. Akad. Nauk SSR.* **22** (1977), no. 9, 507–508.
- [31] Lyng, G. D.; Miller, P. D. The  $N$ -soliton of the focusing nonlinear Schrödinger equation for  $N$  large. *Comm. Pure Appl. Math.* **60** (2007), no. 7, 951–1026. doi:10.1002/cpa.20162
- [32] Ma, Y.-C. The perturbed plane-wave solutions of the cubic Schrödinger equation. *Stud. Appl. Math.* **60** (1979), no. 1, 43–58. doi:10.1002/sapm197960143

- [33] Matveev, V. B.; Salle, M. A. *Darboux transformations and solitons*. Springer Series in Nonlinear Dynamics. Springer, Berlin, 1991. doi:10.1007/978-3-662-00922-2
- [34] <http://news.mit.edu/2016/prediction-tool-rogue-waves-0225> (2016)
- [35] Peregrine, D. H. Water waves, nonlinear Schrödinger equations and their solutions. *J. Austral. Math. Soc. Ser. B* **25** (1983), no. 1, 16–43. doi:10.1017/S0334270000003891
- [36] Suret, P.; Kousseaifi, R. E.; Tikan, A.; Evain, C.; Randoux, S.; Szwaj, C.; Bielawski, S. Single-shot observation of optical rogue waves in integrable turbulence using time microscopy. *Nat. Commun.* **7** (2016), 13136. doi:10.1038/ncomms13136
- [37] Tajiri, M.; Watanabe, Y. Breather solutions to the focusing nonlinear Schrödinger equation. *Phys. Rev. E* (3) **57** (1998), no. 3, part B, 3510–3519. doi:10.1103/PhysRevE.57.3510
- [38] Tikan, A.; Billet, C.; El, G.; Tovbis, A.; Bertola, M.; Sylvestre, T.; Gustave, F.; Randoux, S.; Genty, G.; Suret, P.; Dudley, J. M. Universality of the Peregrine soliton in the focusing dynamics of the cubic nonlinear Schrödinger equation. *Phys. Rev. Lett.* **119** (2017), no. 3, 033901, 6 pp. doi:10.1103/PhysRevLett.119.033901
- [39] Walczak, P.; Randoux, S.; Suret, P. Optical rogue waves in integrable turbulence. *Phys. Rev. Lett.* **114** (2015), no. 14, 143903, 5 pp. doi:10.1103/PhysRevLett.114.143903
- [40] Wang, L.; Yang, C.; Wang, J.; He, J. The height of an  $n$ th-order fundamental rogue wave for the nonlinear Schrödinger equation. *Phys. Lett. A* **381** (2017), no. 20, 1714–1718. doi:10.1016/j.physleta.2017.03.023
- [41] Zakharov, V. E.; Shabat, A. B. A scheme for integrating the nonlinear equations of mathematical physics by the method of the inverse scattering problem. I. *Funct. Anal. Appl.* **8** (1974), no. 3, 226–235. doi:10.1007/BF01075696
- [42] Zhou, X. Direct and inverse scattering transforms with arbitrary spectral singularities. *Comm. Pure Appl. Math.* **42** (1989), no. 7, 895–938. doi:10.1002/cpa.3160420702
- [43] Zhou, X. The Riemann-Hilbert problem and inverse scattering. *SIAM J. Math. Anal.* **20** (1989), no. 4, 966–986. doi:10.1137/0520065

DENIZ BILMAN  
 University of Michigan  
 530 Church St., East Hall 2074  
 Ann Arbor, MI 48104  
 USA  
 E-mail: bilman@umich.edu

PETER D. MILLER  
 University of Michigan  
 530 Church St., East Hall 2074  
 Ann Arbor, MI 48104  
 USA  
 E-mail: millerpd@umich.edu

Received October 2017.