

# Universality in the profile of the semiclassical limit solutions to the focusing Nonlinear Schrödinger equation at the first breaking curve

M. Bertola<sup>†‡12</sup>, A. Tovbis<sup>‡</sup>

<sup>†</sup> *Centre de recherches mathématiques, Université de Montréal  
C. P. 6128, succ. centre ville, Montréal, Québec, Canada H3C 3J7*

<sup>‡</sup> *Department of Mathematics and Statistics, Concordia University  
1455 de Maisonneuve W., Montréal, Québec, Canada H3G 1M8*

<sup>‡</sup> *University of Central Florida Department of Mathematics  
4000 Central Florida Blvd. P.O. Box 161364 Orlando, FL 32816-1364*

## Abstract

We consider the semiclassical (zero-dispersion) limit of the one-dimensional focusing Nonlinear Schrödinger equation (NLS) with decaying potentials. If a potential is a simple rapidly oscillating wave (the period has the order of the semiclassical parameter  $\varepsilon$ ) with modulated amplitude and phase, the space-time plane subdivides into regions of qualitatively different behavior, with the boundary between them consisting typically of collection of piecewise smooth arcs (*breaking curve(s)*). In the first region the evolution of the potential is ruled by modulation equations (Whitham equations), but for every value of the space variable  $x$  there is a moment of transition (breaking), where the solution develops fast, quasi-periodic behavior, i.e., the amplitude becomes also fastly oscillating at scales of order  $\varepsilon$ . The very first point of such transition is called the point of *gradient catastrophe*. We study the detailed asymptotic behavior of the left and right edges of the interface between these two regions at any time after the gradient catastrophe. The main finding is that the first oscillations in the amplitude are of nonzero asymptotic size even as  $\varepsilon$  tends to zero, and they display two separate natural scales; of order  $\mathcal{O}(\varepsilon)$  in the parallel direction to the breaking curve in the  $(x, t)$ -plane, and of order  $\mathcal{O}(\varepsilon \ln \varepsilon)$  in a transversal direction. The study is based upon the inverse-scattering method and the nonlinear steepest descent method.

---

<sup>1</sup>Work supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC)

<sup>2</sup>[bertola@crm.umontreal.ca](mailto:bertola@crm.umontreal.ca)

# Contents

<b>1</b>	<b>Introduction and main results</b>	<b>1</b>
1.1	Description of results . . . . .	3
<b>2</b>	<b>A short review of the zero dispersion limit of the inverse scattering transform</b>	<b>8</b>
2.1	The $g$ -function . . . . .	10
2.2	Reduction to the model RHP . . . . .	11
<b>3</b>	<b>The analysis near the right breaking curve for NLS</b>	<b>13</b>
3.1	Long time asymptotics of $S(\eta)$ along the breaking curve . . . . .	17
<b>4</b>	<b>First approximation</b>	<b>18</b>
4.1	Outer parametrix . . . . .	18
4.1.1	Reality . . . . .	20
4.2	Local parametrices near $\eta, \bar{\eta}$ and first approximation to the solution . . . . .	21
4.3	Approximate solution and error analysis . . . . .	21
4.4	Solution of the local RHP for $\mathbf{H}_K$ (4-28). . . . .	22
<b>5</b>	<b>Improved approximation</b>	<b>23</b>
5.1	Next order improved local parametrix . . . . .	23
5.2	Improved outer parametrix . . . . .	24
<b>6</b>	<b>The left breaking curve</b>	<b>27</b>
6.1	Outer parametrix . . . . .	29
6.2	Local parametrix . . . . .	29
<b>7</b>	<b>Corrections to the NLS solution</b>	<b>30</b>
7.1	Discussion . . . . .	31
<b>A</b>	<b>Explicit formulæ</b>	<b>32</b>
<b>B</b>	<b>Proofs</b>	<b>33</b>
B.1	Proof of Prop. 5.1. . . . .	33
B.2	Proof of Thm. 5.1. . . . .	34
B.3	Proof of Thm. 7.1. . . . .	36
B.4	Proof of Thm. 7.2. . . . .	37

# 1 Introduction and main results

The focusing Nonlinear Schrödinger (NLS) equation,

$$i\varepsilon\partial_t q + \frac{1}{2}\varepsilon^2\partial_x^2 q + |q|^2 q = 0, \quad (1-1)$$

where  $x \in \mathbb{R}$  and  $t \geq 0$  are space-time variables, is a basic model for self-focusing and self-modulation e.g. it governs nonlinear transmission in optical fibers; it can also be derived as a modulation equation for general nonlinear systems. It was first integrated by Zakharov and Shabat [1] who produced a Lax pair for it and used the inverse scattering procedure to describe general decaying solutions ( $\lim_{|x| \rightarrow 0} q(x, 0) = 0$ ) in terms of radiation and solitons. Throughout this work, we will use the abbreviation NLS to mean “focusing Nonlinear Schrödinger equation”.

Our interest in the semiclassical limit ( $\varepsilon \rightarrow 0$ ) of NLS stems largely from its *modulationally unstable* behavior, as shown by Forest and Lee [2], the modulation system for NLS can be expressed as a set of nonlinear PDE with complex characteristics; thus, the system is ill posed as an initial value problem with Cauchy data. As a result, initial data

$$q(x, 0, \varepsilon) = A(x)e^{i\Phi(x)/\varepsilon}, \quad (1-2)$$

i.e. a plane wave with amplitude modulated by  $A(x)$  and phase modulated by  $\Phi(x)$  are expected to break immediately into some other, presumably more disordered, wave form when the functions  $A(x)$  and  $\Phi(x)$  possess no special properties.

However, in the case of an *analytic* initial data, the NLS evolution displays some ordered structure instead of the disorder suggested by the modulational instability, as shown by Fig. 1, see [3], [4] and [5].

Fig. 1 from [5] clearly identifies regions where different types of behavior of the solution  $q(x, t, \varepsilon)$  appear. These regions are separated by some independent of  $\varepsilon$  curves in the  $x, t$  plane that are called **breaking curves** or **nonlinear caustics**. Within each region, the strong asymptotics of  $q(x, t, \varepsilon)$  can be expressed in terms of Riemann Theta-functions to within an error term of order  $\mathcal{O}(\varepsilon)$  that is uniform on compact subsets of the region ([6] for the pure soliton case and [7] for the general case). In this context, regions of different asymptotic behavior of  $q(x, t, \varepsilon)$  corresponds to the different genera of the hyperelliptic Riemann surface whose Theta-functions enter in the asymptotic description. No error estimates for the asymptotics on the breaking curve were studied so far, although it was expected that the accuracy should be of the order  $\mathcal{O}(\varepsilon^{\frac{1}{2}})$  at any regular point on the breaking curve. A visual inspection of Fig. 1 suggests that the behavior of  $q$  at the tip of the breaking curve is different from the behavior elsewhere on the breaking curve.

The tip-point is called a point of **gradient catastrophe**, or **elliptic umbilical singularity** ([8].) The main goal of this paper is to analyze the leading order asymptotic behavior of the solution  $q(x, t, \varepsilon)$  on and around the breaking curve except a vicinity of the gradient catastrophe point, and to obtain the corresponding error estimate. More precisely, we will examine a strip around the breaking curve between the genus zero and genus two regions, which has the width of order  $\mathcal{O}(\varepsilon \ln \varepsilon)$ .

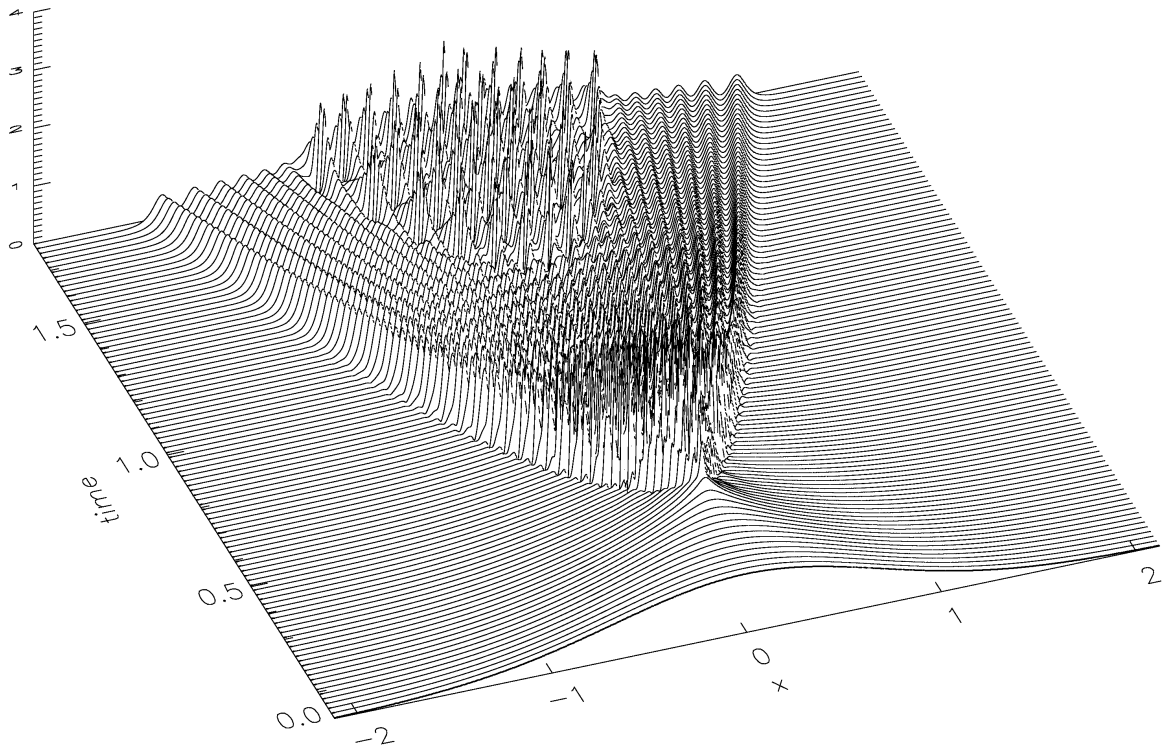


Figure 1: Absolute value  $|q(x, t, \varepsilon)|$  of a solution  $q(x, t, \varepsilon)$  to the focusing NLS (1-1) versus  $x, t$  coordinates (Reprinted from [5] with permission from Elsevier/North-Holland). Here  $A(x) = e^{-x^2}$ ,  $\Phi'(x) = -\tanh x$  and  $\varepsilon = 0.02$ .

## 1.1 Description of results

The main goal of the paper is to provide a detailed asymptotic analysis in the zero-dispersion (or semi-classical) limit ( $\varepsilon \rightarrow 0$ ) in the neighborhood of the breaking curve, that is, the frontier in  $(x, t)$ -space that separates the region of smooth behavior (*genus zero* region) from the region of fast oscillations (*genus two* region), which is clearly seen on Fig. 1.

The main highlight is an “universal” expression for the behavior of the first oscillations as we egress from the genus zero region into the genus two one. Here “universality” means that the expression does not depend upon the details of the initial data, or rather, it depends on it only through a few parameters that are explicitly computable. In order to give a taste of the expression in consideration we need to present a minimum of background notation, with more details contained in the body of the paper.

The zero-dispersion limit of (1-1) can be addressed in terms of modulation equations involving one pair of complex conjugate Riemann invariants, say  $\alpha, \bar{\alpha}$ . The invariant  $\alpha = \alpha(x, t)$  (which we choose in the upper half plane) depends on time/space according to the modulation equations (also called Whitham equations). The asymptotic behavior for the solution of the NLS equation in the genus zero region is then provided by [7]

$$\begin{aligned}
 q(x, t) &\sim -b(x, t)e^{\frac{i}{\varepsilon}\Phi(x, t)}(1 + o(1)), & \text{where} & \\
 \alpha(x, t) &= a(x, t) + ib(x, t) \\
 \partial_x \Phi(x, t) &= -2a(x, t)
 \end{aligned}
 \tag{1-3}$$

and  $o(1)$  denotes some infinitesimal quantity in  $\varepsilon$ . However, as we leave the genus-zero region, such approximation breaks down: the term that was infinitesimal  $o(1)$  becomes suddenly of finite order and a different approximation scheme must be employed. The modulation equations alone cannot detect such sudden change of behavior, since they are equations that are obtained from a formal manipulation. The modern approach to the problem requires the introduction of a special function, the ***g*-function**; this is a locally analytic function in the complex spectral plane of the associated linear problem. It is defined in terms of the scattering transform of the initial data: without entering the details we may say that in our problem the initial data is encoded in a Schwarz-symmetrical analytic function  $f_0(z)$  ( $\Im f_0(z)$  generically has a jump discontinuity along the real axis). Taking into account the time ( $t$ ) evolution and normalization with respect to  $x$ , one obtains

$$f(z) := f(z; x, t) = f_0(z) - xz - 2tz^2 . \tag{1-4}$$

The *g*-function then is obtained as solution of a scalar Riemann–Hilbert problem (RHP) on a free boundary with the jump  $f(z)$  <sup>3</sup>.

---

<sup>3</sup>This means that the contour where the jump is has to be determined implicitly along the way of the construction of the solution itself.

For the reader unfamiliar with the method we may compare the determination of the free-boundary to the determination of the steepest descent contour for Laplace-type integral. The main requirements (but this is a severely shortened list, see Section 2 for the details) is that

$$g_+(z) + g_-(z) = f(z; x, t) , \quad z \in \gamma_m \quad (1-5)$$

where  $\gamma_m$  is an arc of the said free boundary and the sign in the subscript denotes the left/right boundary value; moreover  $g$  is bounded at infinity and with some growth requirements at the endpoints (if any) of  $\gamma_m$ . The function  $g(z) = g(z; x, t)$ , defined through the above mentioned RHP, must be supplemented by several inequalities that involve

$$h(z; x, t) := 2g(z; x, t) - f(z; x, t) , \quad (1-6)$$

and the **sign distribution** of its **imaginary part**  $\Im(h)$ . All the functions  $f, g, h$  are Schwartz-symmetrical, in particular  $h(z; x, t) = \overline{h(\bar{z}; x, t)}$ ; this means that the sign-distribution requirements in the lower half-plane are the opposite of those in the upper half,  $\mathbb{C}_+$ .

One of the main requirements is that  $\Im(h)$  in  $\mathbb{C}_+$  must be **positive** along the complement  $\gamma_c$  of  $\gamma_m$  within the above-mentioned free-boundary, and must be negative on both sides of  $\gamma_m$ , the latter condition implying that  $\Im h = 0$  on  $\gamma_m$ . What happens is that one of these particular requirements fails in certain regions of the  $(x, t)$  plane; this is where the zero-dispersion asymptotic behavior of the solution changes its behavior. The typical *phase portrait* is depicted in Fig. 4. The boundary between the two regions consists of points  $(x, t)$ , for which  $\Im(h(z; x, t)) = 0$  and simultaneously  $h'(z; x, t) = 0$  at some point  $z$  of the spectral plane; we denote by  $\eta(x, t)$  the value of  $z \in \mathbb{C}_+$  satisfying  $h'(z; x, t) = 0$  for all  $x, t$  in a vicinity of the breaking curve. Because of the analyticity of  $h$  this point  $\eta$  must be a saddle-point for  $\Im h$  and  $h$  can be written as

$$h(z; x, t) = S(x, t) - \frac{C(x, t)}{2i} (z - \eta)^2 + \mathcal{O}((z - \eta)^3) . \quad (1-7)$$

(the factors are just for convenience). Note that the breaking curve is now given by  $\Im S(x, t) = 0$ . Generically, the second derivative coefficient  $C(x, t)$  does not vanish: this corresponds to the smooth part of the boundary between the two regions in the sketch of Fig. 4. The angular point of the boundary (where  $C(x_0, t_0) = 0$ ) is the so-called **point of umbilical gradient catastrophe** and will be dealt with in a forthcoming publication.

The critical value  $S(x, t) \in \mathbb{C}$  plays the first major role in our story: indeed we have

- The function  $S(x, t)$  is a local diffeomorphism for  $x \in \mathbb{R}, t \in \mathbb{R}_+$  with values in  $\mathbb{C} \sim \mathbb{R}^2$  except at the point of gradient catastrophe (Lemma 3.2).
- The function  $S(x, t)$  maps a tubular neighborhood of the breaking curve not containing the point of gradient catastrophe diffeomorphically on a tubular neighborhood of the real  $S$ -axis (Thm. 3.1).

This means that we can use  $\Im S$  as a parametrization in a transversal direction to the breaking curve, while using  $\Re S$  as a parametrization in the parallel direction. The paper focuses on the analysis of such a tubular neighborhood with a width of order  $\mathcal{O}(\varepsilon \ln \varepsilon)$ ; more precisely we introduce the **exploration parameter**  $\varkappa(x, t)$  (Def. 3.2) by

$$S(x, t) = \frac{\vartheta(x, t)}{2} + \frac{i}{2} \varkappa(x, t) \varepsilon \ln \varepsilon . \quad (1-8)$$

The parameter  $\varkappa$  is positive on the left breaking curve (negative on the right one, respectively) as we enter the oscillatory zone (in wavy lines in Fig. 4); a finite value of  $\varkappa$  corresponds to an infinitesimal *exploration*.

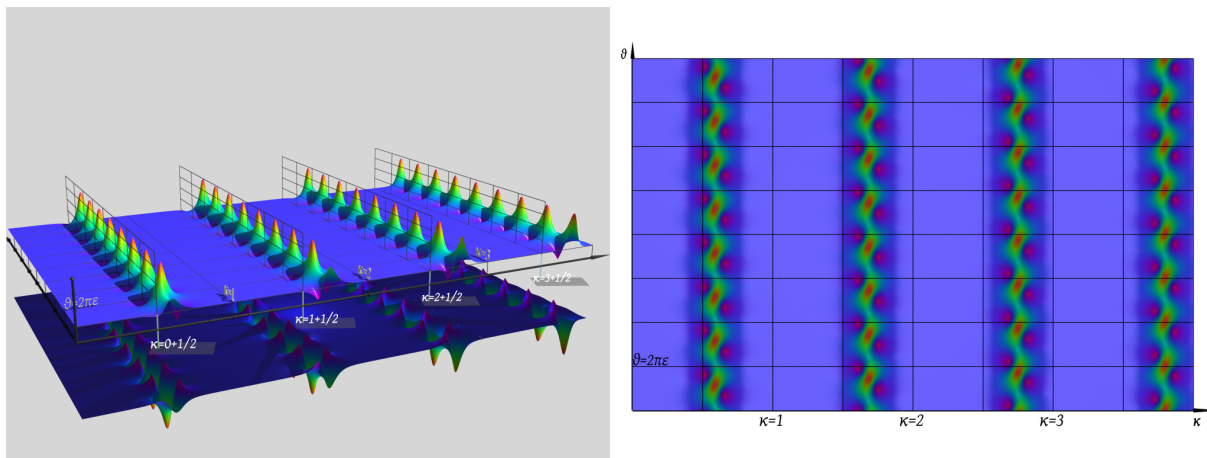


Figure 2: The relative amplitude of the corrected solution to the genus-0 solution. The plot is in the variables  $\theta, \varkappa$  and the scale is not uniform. The parameters  $C, \eta$ , etc. were obtained from an explicit genus-zero solution and are:  $\alpha = 0.8120825800 + 1.251939783 i$ ,  $\eta = -.5422408496 + 1.731859687 i$ ,  $C = .2387253889 - .4667042683 i$ ,  $\lambda_0 = -2.358599870 + 2.555370658 i$ ,  $\varepsilon = 10^{-10}$ ,  $x_0 = 0.1$ ,  $t_0 = 0.1570482549$ . It is actually necessary to set such a small epsilon because for  $\varkappa \in \mathbb{N}$  we have a discontinuity of order  $\sqrt{\varepsilon}$ , to be expected since this is the order of our approximation and at the integers the correction we have is of the exact same order of the error. Note that since  $\varkappa$  measures distances in the scale  $\varepsilon |\ln \varepsilon|$  while  $\theta$  in the scale  $\varepsilon$ , the distance between consecutive “ranges”, should be much longer than the transversal separation of the peaks. Also the successive ranges are relatively very thin (of order  $1/|\ln \varepsilon|$  relative to the longitudinal separation). Therefore the typical *size* of the hills is  $\varepsilon$  in all directions while the *separation* in the longitudinal direction is  $\varepsilon |\ln \varepsilon|$ . This picture is consistent with a full-blown genus-two regime, where the solution is quasi-periodic at the scale  $\varepsilon$ : as we progress into the genus-2 region the separation reduces to the natural scale  $\varepsilon$ . If we were to plot this in the  $(x, t)$  plane, the only difference would be a linear change of coordinates and the “ranges” would be parallel to the breaking curve. Note that the picture is strictly periodic in the  $\vartheta$  direction but only *quasi-periodic* in the  $\varkappa$  direction although it is almost not discernible in the picture: in fact the subsequent ranges are actually slightly shifted.

The main finding of the paper is the fine structure of the “shock waves” along the breaking curve in the  $x, t$ -plane, clearly visible on Fig. 1. They turn out to be not “uniform”, as one might erroneously

conclude looking at Fig. 1. As shown on Fig. 2, each wave consists of the sequence of “peaks” or “teeth” interlaced by “holes” or “depressions”. It is remarkable that the mathematical description of this phenomenon is quite similar to the description of asymptotic of “colonization” of a new interval by zeroes of certain orthogonal polynomials in [9, 10]. More precisely:

- the genus zero approximate solution (1-3), continued into the genus-two region, gets a finite order correction (w.r.t.  $\varepsilon$ ) only for **(positive) half-integer values of  $\varkappa$** , which correspond to the rows of “teeth”;
- in between half-integers, the correction is of order  $\mathcal{O}\left(\varepsilon^{\frac{1}{2}(\text{dist}(\varkappa, \mathbb{Z})-1)}\right)$ ;
- the correction term exhibits *fast oscillations* only in the direction parallel to the breaking curve, i.e., it depends on  $e^{i\frac{\vartheta(x,t)}{\varepsilon}}$  and it is periodic in  $\vartheta$ ;
- all the remaining dependence of the correction is *slow*, namely, it depends on  $(x, t)$  at a  $\mathcal{O}(1)$  scale;
- the formula for the correction term is valid with an error term of order  $\mathcal{O}(\sqrt{\varepsilon})$ ;
- the only dependence on  $\varepsilon$  in the correction terms is through the expressions  $e^{\pm i\frac{\vartheta(x,t)}{\varepsilon}}$  and  $\varepsilon^{\frac{1}{2}(1-K)}$ , where  $K = \max(0, \lfloor \varkappa \rfloor)$  with  $\lfloor \varkappa \rfloor$  denoting the integer part of  $\varkappa$ .

The picture getting painted is that the first oscillation is of *finite amplitude* and the oscillations form “mountainous ranges” with flatlands in between of order  $\varepsilon \ln \varepsilon$  and peaks separated longitudinally by distances of order  $\varepsilon$ . Possibly the most descriptive explanation is contained in Fig. 2, plotted over the  $(\varkappa, \vartheta)$  plane which, we remind, is diffeomorphic to a piece of the  $(x, t)$  plane around the breaking curve, with  $\varkappa$  growing in the transversal direction ( $\varkappa$  and  $\vartheta$  are not necessarily orthogonal; in fact one could compute their relative angle easily from (3-3))

While the correction is strictly periodic in  $\vartheta$ , it is **not periodic** in  $\varkappa$ , despite a superficial inspection of the picture: neither the formula for the correction nor the picture (at a careful inspection) are periodic in  $\varkappa$ .

The actual formula for the correction term is significantly complicated and it is inconvenient to write it here: it is contained in Thm. 7.1 and -in quite explicit terms- in Thm. 7.2. The main point that we observe at this stage is its **universality**, namely, the fact that it does not depend on the details of the initial data. To be more specific, it is an expression that contains only  $C, \vartheta, \varkappa, \alpha, \eta$ .

## Organization of the paper

We start with Section 2 where the essentials about the inverse scattering transform are recalled following [7]. In this section we summarily go through the transformation of the associated Riemann–Hilbert



problem along the lines of the Deift–Zhou steepest descent method [11] by introducing the  $g$ -function (Sect. 2.1) and the transformation of the RHP (Sect. 2.2).

In Section 3 we analyze the differential geometry of the breaking curve and establish the result that  $S(x, t)$  is a local diffeomorphism onto the spectral plane (Lemma 3.2). We also analyze the large-time behavior of  $S(x, t)$  along such breaking curve (Sect. 3.1).

In Section 4 we start the detailed construction of the approximation to the RHP: the description is made in two steps, mainly for pedagogical reasons. In the first step we provide a leading order approximation to the problem and show that the error term fails necessarily to tend to zero as  $\varepsilon \rightarrow 0$  for positive half-integer values of  $\varkappa$  (eq. 4-39). This prompts the necessity of a higher order approximation scheme, contained in Sect. 5.

The analysis is done first on the right branch of the breaking curve and hence should be repeated for the left branch; this is done in Sect. 6. The finding is that the situation is different only in some setup, but the essence of the formulas and of the phenomenon is identical.

The paper culminates in Sect. 7 where we provide the final formulæ for the leading and subleading terms asymptotic behavior of the NLS solution (Thm. 7.1, Thm. 7.2). We also provide a short discussion about the large-time behavior (along the breaking curve) of the solution.

## 2 A short review of the zero dispersion limit of the inverse scattering transform

Given an initial data (potential)  $q(x, 0)$  for the (1-1) that is decaying as  $x \rightarrow \pm\infty$ , the direct scattering transform for the NLS ([1]) produces the scattering data, namely: the reflection coefficient  $r_0(z, \varepsilon)$  and the points of discrete spectrum together with their norming constants (solitons). The time evolution of the scattering data is simple and well-known. Thus, to find the evolution of a given potential at some time  $t$ , one needs to solve the inverse scattering problem at the time  $t$ . Equivalently, we can stipulate that the initial data is assigned directly through the scattering data; thus, one can produce a solution to the NLS (1-1) by choosing some scattering data and solving the inverse scattering problem for  $t = 0$  (initial data) and for  $t > 0$  (evolution of the initial data). The latter approach allows one to avoid solving the direct scattering problem and addressing many delicate issues associated with it (see [12] for more details). Since we are interested in studying the generic structure of solutions of the NLS in the transition regions, it will be convenient for us to define a solution to the NLS (1-1) by its scattering, not by its initial, data.

In considering the semiclassical limit of (1-1), one has to consider the semiclassical limit of the corresponding scattering transform. For the case of decaying potentials of type (1-2), this limit was discussed in [13] (inverse scattering) and [12] (direct scattering), where it was shown to be a correspondence between

$$\alpha(x, t) = -\frac{1}{2}\Phi_x(x) + iA(x, t) \tag{2-1}$$

on the potential side and

$$f(z) = f_0(z) - xz - 2tz^2 \quad (2-2)$$

on the scattering side. Here  $t \geq 0$  is fixed and  $f_0(z)$  has the meaning of the “scaled” logarithm of the reflection coefficient  $r_0(z, \varepsilon)$  that corresponds to the initial data (1-2), that is,

$$f_0(z) = \frac{i}{2} \lim_{\varepsilon \rightarrow 0} \varepsilon r_0(z, \varepsilon) . \quad (2-3)$$

For example, the limiting  $f_0(z)$  that corresponds to the potential  $\cosh(x)e^{\frac{2i}{\varepsilon} \tanh x + i\pi}$ , is given by (see [7])

$$f_0(z) = (1-z) \left[ i\frac{\pi}{2} + \ln(1-z) \right] + z \ln z + \ln 2 + \frac{\pi}{2}\varepsilon, \quad \text{when } \Im z \geq 0. \quad (2-4)$$

Moreover, if  $x(\alpha)$  denotes the inverse function to  $\alpha(x, t)$ , at  $t$  fixed, then  $x(\alpha)$  and  $f_0(z)$  are connected through some Abel type linear integral transform, see details in [12].

Since we have taken the perspective that we are starting with the scattering data of the form  $r_0(z, \varepsilon) := e^{-\frac{2i}{\varepsilon} f_0(z)}$  we shall assume that ([13]):

- $f_0(z)$  is analytic in the upper half-plane and has a continuous limit on  $\mathbb{R}$ ;
- $f_0(z) = \mathcal{O}(z)$  as  $z \rightarrow \infty$ ;
- there exists an interval  $(\mu_-, \mu_+) \subset \mathbb{R}$  such that  $\Im f_0(z) > 0$  for  $z \in (\mu_-, \mu_+)$  and  $\Im f_0(z) < 0$  for  $z \in (-\infty, \mu_-) \cup (\mu_+, \infty)$ ;
- $\Im f_0(z)$ ,  $z \in \mathbb{R}$ , has simple zeroes at  $\mu_{\pm}$  and its values are separated from zero outside neighborhoods of  $\mu_{\pm}$ .

Due to the Schwarz-symmetry of the scattering data for the NLS,  $f_0(z)$  is also a Schwarz-symmetrical function, i.e.,  $f_0(\bar{z}) = \overline{f_0(z)}$ ; generically,  $\Im f_0(z)$  has a jump across the real axis. The above assumptions are not overly restrictive: it was shown that in the case of a solitonless potential the formula (2-3) yields such an  $f_0(z)$  after some process of analytic extension, or “folding” of certain logarithmic branch-cuts onto the real axis. In fact, for our goals it is sufficient to replace the analyticity requirement by a weaker condition: the contours of the RHP for the  $g$ -function (which will be introduced in the next section) always lie within the domain of analyticity of  $f_0(z)$  in  $\mathbb{C}_+$ . Generically,  $r_0(z)$  has a jump discontinuity along the real axis due to the discontinuity in  $\Im f_0$ .

At any time  $t$ , the inverse scattering problem for (1-1) (in the solitonless case) with a fixed (not infinitesimal)  $\varepsilon$  is reducible to the following matrix RHP.

**Problem 2.1** Find a matrix  $\Gamma(z)$  analytic in  $\mathbb{C} \setminus \mathbb{R}$  such that

$$\Gamma_+(z) = \Gamma_-(z) \begin{bmatrix} |r_0|^2 + 1 & \bar{r}_0 e^{-\frac{2i}{\varepsilon}(2tz^2+xz)} \\ r_0 e^{\frac{2i}{\varepsilon}(2tz^2+xz)} & 1 \end{bmatrix}, \quad z \in \mathbb{R}, \quad (2-5)$$

$$\Gamma(z) = \mathbf{1} + \frac{1}{z} \Gamma_1 + \mathcal{O}(z^{-2}), \quad z \rightarrow \infty. \quad (2-6)$$

(In the case with solitons, there are additional jumps across small circles surrounding the points of discrete spectrum.) Then

$$q(x, t, \varepsilon) := -2(\Gamma_1)_{12} \quad (2-7)$$

is the solution of the initial value problem (1-1) for the NLS equation.

The jump matrix for the RHP admits the factorization

$$\begin{bmatrix} |r|^2 + 1 & \bar{r} \\ r & 1 \end{bmatrix} = \begin{bmatrix} 1 & \bar{r} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ r & 1 \end{bmatrix}, \quad (2-8)$$

where  $r = r(z; x, t) = r_0(z)e^{\frac{2i}{\varepsilon}(2tz^2+xz)}$ .

Inspection of the RHP shows that the matrix  $\Gamma(\bar{z})^*$  (where  $*$  stands for the complex-conjugated, transposed matrix) solves the same RHP with the jump matrix  $M(z)$  replaced by  $M^{-1}(z)$  and hence

**Proposition 2.1** *The solution  $\Gamma(z)$  of the RHP for NLS has the symmetry*

$$\Gamma(z)(\Gamma(\bar{z}))^* \equiv \mathbf{1}. \quad (2-9)$$

In order to study the dispersionless limit  $\varepsilon \rightarrow 0$ , the RHP (2-5)-(2-6) undergoes a sequence of transformations (that are briefly recalled in Section 2.2) along the lines of the nonlinear steepest descent method [14, 7], which reduce it to an RHP that allows for an approximation by the so-called model RHP. The latter RHP has piece-wise constant jump matrices (parametrically dependent on  $x, t$ ) and, in general, can be solved explicitly in terms of the Riemann Theta functions, or, in simple cases, in terms of algebraic functions. The  $g$ -function, defined below, is the key element of such a reduction.

## 2.1 The $g$ -function

Given  $f_0(z)$ , we introduce the  $g$ -function  $g(z) = g(z; x, t)$  as the solution to the following scalar RHP:

1.  $g(z)$  is analytic (in  $z$ ) in  $\bar{\mathbb{C}} \setminus \gamma_m$  (including analyticity at  $\infty$ );
2.  $g(z)$  satisfies the jump condition

$$g_+ + g_- = f_0 - xz - 2tz^2 \quad \text{on } \gamma_m, \quad (2-10)$$

for  $x \in \mathbb{R}$  and  $t \geq 0$ , and;

3.  $g(z)$  has the endpoint behavior

$$g(z) = O(z - \alpha)^{\frac{3}{2}} + \text{an analytic function in a vicinity of } \alpha. \quad (2-11)$$

Here:

- $\gamma_m$  is a bounded Schwarz-symmetrical contour (called the main arc) with the endpoints  $\bar{\alpha}, \alpha$ , oriented from  $\bar{\alpha}$  to  $\alpha$  and intersecting  $\mathbb{R}$  only at  $\mu_+$ ;
- $g_{\pm}$  denote the values of  $g$  on the positive (left) and negative (right) sides of  $\gamma_m$ ;
- the function  $f_0 = f_0(z)$ , representing the initial scattering data, is Schwarz-symmetrical and Hölder-continuous on  $\gamma_m$ .

Taking into the account Schwarz symmetry, it is clear that behavior of  $g(z)$  at both endpoints  $\alpha$  and  $\bar{\alpha}$  should be the same.

Assuming  $f_0$  and  $\gamma_m$  are known, the solution  $g$  to the scalar RHP (2-10) without the endpoint condition (2-11) can be obtained by the Plemelj formula

$$g(z) = \frac{R(z)}{2\pi i} \int_{\gamma_m} \frac{f(\zeta)}{(\zeta - z)R(\zeta)_+} d\zeta, \quad (2-12)$$

where  $R(z) = \sqrt{(z - \alpha)(z - \bar{\alpha})}$ . We fix the branch of  $R$  by requiring that  $\lim_{z \rightarrow \infty} \frac{R(z)}{z} = 1$ . If  $f_0(z)$  is analytic in some region  $\mathcal{S}$  that contains  $\gamma_m \setminus \{\mu_+\}$ , the formula for  $g(z)$  can be rewritten as

$$g(z) = \frac{R(z)}{4\pi i} \int_{\hat{\gamma}_m} \frac{f(\zeta)}{(\zeta - z)R(\zeta)_+} d\zeta, \quad (2-13)$$

where  $\hat{\gamma}_m \subset \mathcal{S}$  is a negatively oriented loop around  $\gamma_m$  (which is “pinched” to  $\gamma_m$  in  $\mu_+$ , where  $f$  is not analytic) that does not contain  $z$ . Introducing function  $h = 2g - f$  we obtain

$$h(z) = \frac{R(z)}{2\pi i} \int_{\hat{\gamma}_m} \frac{f(\zeta)}{(\zeta - z)R(\zeta)_+} d\zeta, \quad (2-14)$$

where  $z$  is inside the loop  $\hat{\gamma}_m$ . The endpoint condition (2-11) can now be written as

$$h(z) = O(z - \alpha)^{\frac{3}{2}} \text{ as } z \rightarrow \alpha, \quad (2-15)$$

or, equivalently,

$$\int_{\hat{\gamma}_m} \frac{f(\zeta)}{(\zeta - \alpha)R(\zeta)_+} d\zeta = 0. \quad (2-16)$$

The latter equation is known as a *modulation equation*. The function  $h$  plays a prominent role in this paper. Using the fact that the Cauchy operator for the RHP (2-10) commutes with differentiation, we have

$$h'(z) = \frac{R(z)}{2\pi i} \int_{\hat{\gamma}_m} \frac{f'(\zeta)}{(\zeta - z)R(\zeta)_+} d\zeta, \quad (2-17)$$

where  $z$  is inside the loop  $\hat{\gamma}_m$ .

In order to reduce the RHP (2-5)-(2-6) to the RHP with piece-wise jump matrices, called the *model RHP*, the signs of  $\Im h(z)$  in the upper half-plane should satisfy the following conditions:

- $\Im h(z)$  is negative on both sides of the contour  $\gamma_m$ ;
- there exists a continuous contour  $\gamma_c$  (complementary arc) in  $\mathbb{C}_+$  that connects  $\alpha$  and  $\mu_-$ , so that  $\Im h(z)$  is positive along  $\gamma_c$ . Since  $\Im h(z) > 0$  on the interval  $(-\infty, \mu_-)$ , the point  $\mu_-$  in  $\gamma_c$  can be replaced by any other point of this interval, or by  $-\infty$ .

Note that the first sign requirement, together with (2-10), imply that  $\Im h(z) = 0$  along  $\gamma_m$ . Since the signs of  $\Im h(z)$  play an important role in the following discussion, we call by “sea” and “land” the regions in  $\mathbb{C}_+$ , where  $\Im h(z)$  is negative and positive respectively. In this language, the complementary arc  $\gamma_c$  goes on “land”, whereas the main arc  $\gamma_m$  is a “bridge” or a “dam”, surrounded by the sea, see Fig. 3.

## 2.2 Reduction to the model RHP

We start the transformation of the RHP (2-5)-(2-6) by deforming (preserving the orientation) the interval  $(-\infty, \mu_+)$ , which is a part of its jump contour, into some contour  $\gamma^+$  in the upper half-plane  $\mathbb{C}_+$ , such that  $\mu_+ \in \gamma^+$ . Let  $\gamma_-$  be the Schwarz symmetrical image of  $\gamma^+$ . Using the factorization (2-8) the RHP (2-5)-(2-6) can be reduced to an equivalent one where:

- the right factor of (2-8) is the jump matrix on  $\gamma_+$ ;
- the left factor of (2-8) is the jump matrix on  $\gamma_-$ ;
- the jump matrix on the remaining part of  $\mathbb{R}$  is unchanged.

It will be convenient for us to change the orientation of  $\gamma_+$ , which causes the change of sign in the off-diagonal entry of the corresponding jump matrix. On the interval  $(\mu_+, \infty)$  we have  $\Im f_0(z) < 0$  and it appears that the jump is exponentially close to the identity jump and hence it is possible to prove that it has no bearing on the leading order term of the solution (2-7) (as  $\varepsilon \rightarrow 0$ : see [7] for the case when  $f_0$  is given by (2-4) and [13] for the general case). Therefore the leading order contribution in (2-7) comes from the contour  $\gamma = \gamma_+ \cup \gamma_-$ . In the genus zero case, the contour  $\gamma$  contains points  $\alpha, \bar{\alpha}$ , which divide it into the main arc  $\gamma_m$  (contained between  $\bar{\alpha}$  and  $\alpha$ , and the complementary arc  $\gamma_c = \gamma \setminus \gamma_m$ . According to the sign requirements (2.1), the contour  $\gamma_m$  is uniquely determined as an arc of the level curve  $\Im h(z) = 0$  (bridge) that connects  $\mu_+$  and  $\alpha$ , whereas  $\gamma_c$  can be deformed arbitrarily “on the land”. Because of the Schwarz symmetry 2.1, it is sufficient to consider  $\gamma$  only in the upper half-plane, i.e., it is sufficient to consider  $\gamma_+$ .

Having found the branch-point  $\alpha$ , the  $g$ -function  $g(z)$  and the contour  $\gamma_m$ , we introduce additional contours customarily called “lenses” that join  $\alpha$  to  $\mu_+$  on both sides of  $\gamma_m$  (and symmetrically down under). These lenses are to be chosen rather freely with the only condition that  $\Im h$  must be negative along them (positive in  $\mathbb{C}_-$ ). This condition is guaranteed by (2.1).

The two spindle-shaped regions between  $\gamma_m$  and the lenses are usually called upper/lower **lips** (relative to the orientation of  $\gamma_m$ ). At this point one introduces the auxiliary matrix-valued function  $Y(z)$  as follows

$$Y(z) = e^{-\frac{2i}{\varepsilon}g(\infty)\sigma_3}\Gamma(z) \begin{cases} e^{\frac{2i}{\varepsilon}g(z)\sigma_3} & \text{outside the lips,} \\ e^{\frac{2i}{\varepsilon}g(z)\sigma_3} \begin{bmatrix} 1 & -e^{-\frac{2i}{\varepsilon}h(z)} \\ 0 & 1 \end{bmatrix} & \text{in the upper lip in } \mathbb{C}_+, \\ e^{\frac{2i}{\varepsilon}g(z)\sigma_3} \begin{bmatrix} 1 & e^{-\frac{2i}{\varepsilon}h(z)} \\ 0 & 1 \end{bmatrix} & \text{in the lower lip in } \mathbb{C}_+. \end{cases} \quad (2-18)$$

The definition of  $Y(z)$  in  $\mathbb{C}_-$  is done respecting the symmetry in Prop. 2.1, namely

$$Y(z) = (Y(\bar{z})^*)^{-1}, \quad z \in \mathbb{C}_-. \quad (2-19)$$

The jumps for the matrix  $Y(z)$  are reported in Fig. 3.

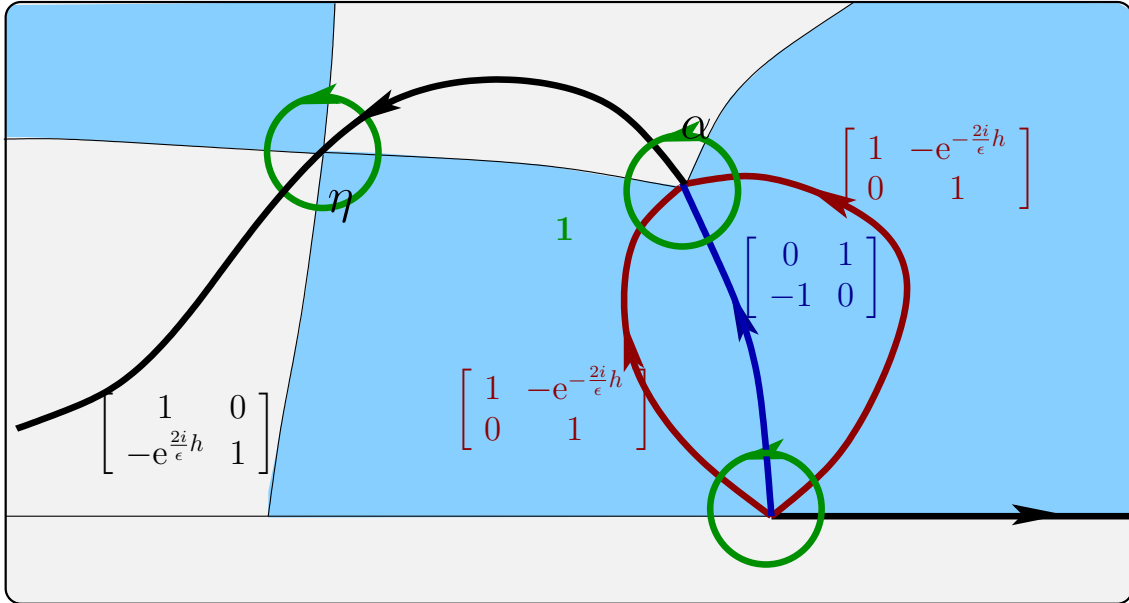


Figure 3: The jumps for the RHP for  $Y$ . The shaded region is where  $\Im h < 0$  (the “sea”). The blue contour is the main arc  $\gamma_m$ , the black contour in  $\mathbb{C}_+$  is the complementary arc  $\gamma_c$  and the red contours are the lenses.

**The model RHP.** In the limit  $\varepsilon \rightarrow 0$ , according to the signs (2.1), the jump matrices on the complementary arc  $\gamma_c$  and on the lenses are approaching the identity matrix  $\mathbf{1}$  exponentially fast. Removing these contours from the RHP for  $Y(z)$ , we will have only one remaining contour  $\gamma_m$  with the constant jump matrix  $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$  on it. This is the model RHP. Calculating the (1,2) entry of the residue at

infinity (see (2-7)) of the solution to the model RHP, one obtains the leading order term of the genus zero solution as follows ([7])

$$q_0(x, t, \varepsilon) = \Im \alpha(x, t) e^{-\frac{2i}{\varepsilon} \int_0^x \Re \alpha_0(s, t) ds} ; \quad (2-20)$$

To justify removing contours with exponentially small jump matrices, one has to calculate the error estimates coming from neighborhoods of points  $\alpha$ ,  $\bar{\alpha}$  and  $\mu_+$  (for  $\alpha$  and  $\mu_+$ , these neighborhoods are shown as green circles in Fig. 3). This is often done through the *local parametrices*. We shall consider the construction of the parametrices near these points as already done and known to the reader. We refer to the literature [7, 14]. The only information that we need is that these parametrices allow to approximate the exact solution to within an error term  $\mathcal{E}(z) = \mathbf{1} + \mathcal{O}(\varepsilon)$ .

### 3 The analysis near the right breaking curve for NLS

When  $x, t$  is on the breaking curve the asymptotic solution to NLS is about to change its genus. Although the phenomenon described in this paper is typical for any increase of genus, we will focus on the first break, i.e., on the transition from genus from 0 to genus 2. This happens by either:

- creating a **new band** (right breaking curve in Fig. 4);
- splitting an existing band (left breaking curve in Fig. 4).

Let  $(x_0, t_0)$  be on the (smooth part of the) breaking curve; this implies that there is  $\eta \in \mathbb{C}_+$

$$\Im h(\eta, x_0, t_0) = 0, \quad h'(\eta(x_0, t_0), x_0, t_0) = 0, \quad C = C(x_0, t_0) := -ih''(\eta(x_0, t_0), x_0, t_0) \neq 0, \quad (3-1)$$

where prime denotes derivative with respect to the spectral parameter  $z$ ; moreover,  $\Im h(z, x_0, t_0)$  has the required signs (2.1) on both sides of main (bands) and complementary (gaps) arcs. We keep notation  $h(z; x, t)$  for extension of  $h(z; x_0, t_0)$  in a small neighborhood of  $(x_0, t_0)$  even if **the sign-distribution requirements (2.1) are not satisfied**.

**Definition 3.1** *We denote the critical value*

$$S(x, t) := h(\eta(x, t), x, t), \quad (3-2)$$

*which satisfies  $\Im S(x_0, t_0) = 0$ .*

We will now show that  $S(x, t)$  is locally a diffeomorphism from the  $(x, t)$  plane near the smooth part of the breaking curve, into the complex  $S$ -plane. From [7], formula (4.46), we have

$$h_x(z) = R(z), \quad h_t(z) = 2(z + a)R(z), \quad R(z) = \sqrt{(z - \alpha)(z - \bar{\alpha})}, \quad (3-3)$$

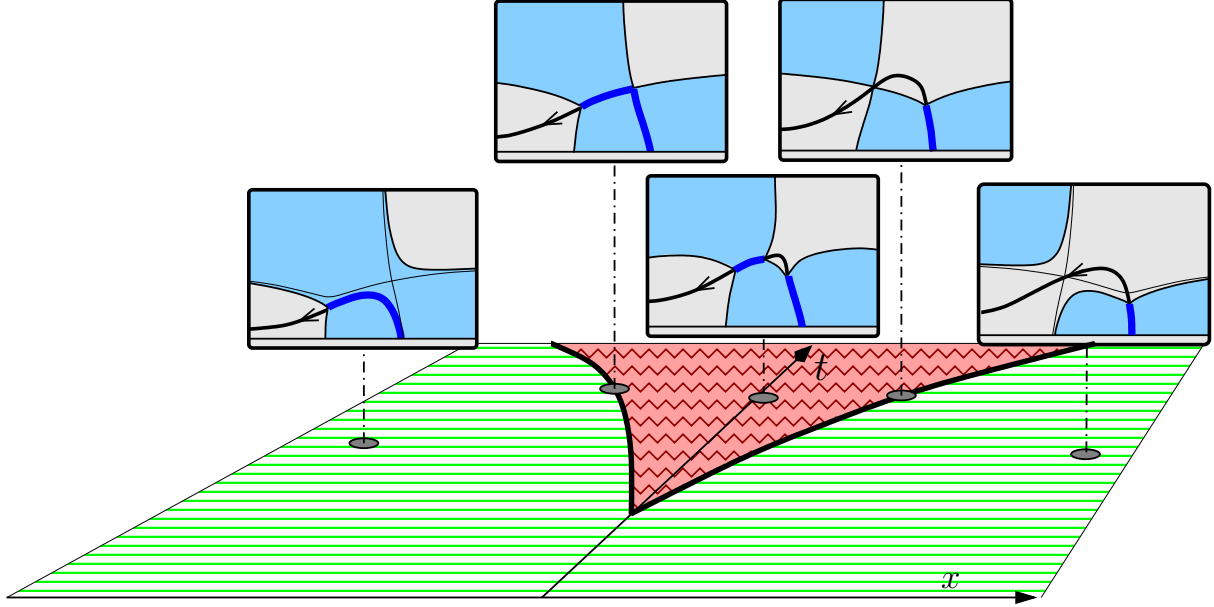


Figure 4: The typical zero-dispersion phase diagram for a one-hump initial data. Representative at different points in the  $(x, t)$ -plane are the level-curves of  $\Im(h)$ . The genus-zero region in the  $(x, t)$ -plane is the lighter area, while the wedge-shaped area is the genus 2 region. Only the upper-half spectral plane is depicted. The shape of the level curves was obtained numerically on a simple example.

where the square root is chosen to have the behavior  $R(z) \sim z$  near infinity (this is the opposite branch compared to the one used in [7]) and analytic off the main arc. Since (by definition of  $\eta$ )  $h'(\eta) = 0$  we have that the partial and total derivatives of  $S(x, t) = h(\eta(x, t), x, t)$  coincide and hence

$$S_x = R(\eta) , \quad S_t = 2(\eta + \Re(\alpha))R(\eta) . \quad (3-4)$$

We also have

**Lemma 3.1 (Lemma 4.5 in [7])** For  $t \geq 0$  and for any  $x \in \mathbb{R}$

- the function  $\Im h_x(z) = \Im R(z)$  is strictly positive in the upper half-plane except for the finite region enclosed between the main arc and the vertical segment  $[\alpha, a]$ .
- The inequality  $\Im h_t(z) = 2\Im(z + a)R(z)$  holds for  $|z|$  large enough and to the left of the zero level curve of  $\Im h_t(z)$  that emanates from  $\alpha$ . It changes sign when crossing this curve or the main arc.

In addition to this we will show that

**Lemma 3.2** The function  $S(x, t)$  is a local diffeomorphism for  $x \in \mathbb{R}$ ,  $t \in \mathbb{R}_+$  with values in  $\mathbb{C} \sim \mathbb{R}^2$  except at the point of gradient catastrophe.



**Proof.** The Jacobian of the map  $S(x, t)$  is

$$\begin{aligned} J(x, t) &= \det \begin{bmatrix} \Re S_x & \Re S_t \\ \Im S_x & \Im S_t \end{bmatrix} = \Re S_x \Im S_t - \Re S_t \Im S_x = \Im [\overline{S_x} S_t] = 2\Im(\eta + a)|R(\eta)|^2 = \\ &= 2\Im(\eta)|R(\eta)|^2. \end{aligned} \quad (3-5)$$

Thus the Jacobian does not vanish on the upper half-plane and  $S$  is a diffeomorphism as long as  $\eta \neq \alpha$ .

**Q.E.D.**

**Theorem 3.1** *Assume that on the breaking curve  $t(x)$  there is a unique gradient-catastrophe point  $(x_0, t_0)$ , i.e.,  $\eta(x_0, t_0) = \alpha(x_0, t_0)$ . Then, outside of any (small) neighborhood of the gradient-catastrophe point  $(x_0, t_0)$ , the function  $S(x, t)$  maps a strip-like neighborhood of the left breaking curve diffeomorphically to a strip around the negative real  $S$ -axis with the genus 2 phase mapped to  $\Im S > 0$ , while it maps a corresponding neighborhood of the right breaking curve to a strip around the positive real  $S$ -axis, with the genus 2 phase mapped to  $\Im S < 0$ .*

**Proof.** The full derivative of  $S$  along the breaking curve  $t(x)$  is

$$\begin{aligned} \frac{d}{dx} S(x, t) &= \frac{d}{dx} h(\eta(x, t), x, t(x)) = h'(\eta, x, t)\eta_x + h_x(\eta, x, t) + h_t(\eta, x, t)t'(x) = \\ &= h_x(\eta, x, t) + h_t(\eta, x, t)t'(x) \neq 0. \end{aligned} \quad (3-6)$$

The vectors  $h_x(z), h_t(z)$  in  $\mathbb{C}$  are not parallel to each other for any  $z \in \mathbb{C}^+ \setminus \{\alpha\}$ . But  $S(x, t)$  is a real-valued function along  $t(x)$ , so  $\frac{d}{dx} \Re S(x, t) \neq 0$ . Since  $S(x_0, t_0) = 0$  and  $S(x, t(x)) \rightarrow +\infty$  as  $x \rightarrow +\infty$  (see (3-18)), the right branch of the breaking curve is mapped onto the positive real axis of  $S$ . Similarly, the left branch of the breaking curve is mapped onto the positive real axis of  $S$ . The required sign distribution of  $\Im S(x, t)$  in the genus two region follows from diagrams at Figure 4. **Q.E.D.**

According to Thm. 3.1, the real  $S$ -axis (up to a small neighborhood of the origin) is diffeomorphic to the two branches of the first breaking curve.

We will consider mostly the right-breaking curve, leaving the (not too different) analysis of the left-breaking curve to later: the genus 2 region corresponds to *negative values* of  $\Im S$ . We want to “explore” the genus-two region within a scale  $\varepsilon|\ln \varepsilon|$  and hence we introduce below a conveniently normalized **exploration parameter**  $\varkappa$ .

**Definition 3.2 (Exploration parameter)** *We define the two functions  $\vartheta(x, t), \varkappa(x, t)$  by the relation*

$$2S(x, t) = \vartheta(x, t) + i\varkappa \varepsilon \ln \varepsilon \quad (3-7)$$

This equation defines implicitly a *foliation* of the a neighborhood of  $(x_0, t_0)$ . The breaking curve is defined by the equation  $\varkappa = 0$  and  $\vartheta$  serves as a parametrization of it, thanks to Lemma 3.2. The scaling in (3-7)

means that  $\varkappa$  parametrizes an infinitesimal strip-like neighborhood of the breaking curve with a width of order  $\varepsilon \ln \varepsilon$ .

We will write

$$2ih(z) = -C(z - \eta)^2(1 + \mathcal{O}(z - \eta)) + 2iS, \quad \eta = \eta(x, t), \quad S := S(x, t), \quad C := C(x, t). \quad (3-8)$$

The complementary arc  $\gamma_c$  that goes through  $\eta$  can be deformed so that – in a finite neighborhood of  $\eta$  –

$$\Re(h(z) - S) \equiv 0 \quad \text{when} \quad z \in \gamma_c. \quad (3-9)$$

**Definition 3.3 (Scaling local coordinate)** *Define*

$$\zeta := \sqrt{-\frac{2i}{\varepsilon}(h(z) - S)} \quad (3-10)$$

where the root is chosen so that the complementary arc  $\Re(h - S) \equiv 0$  is mapped to the real  $\zeta$ -axis with the natural orientation.

The jump on the gap (complementary arc) that goes through  $\eta_0$  is

$$Y_+ = Y_- \begin{bmatrix} 1 & 0 \\ -e^{\frac{2i}{\varepsilon}h(z)} & 1 \end{bmatrix} = Y_- \begin{bmatrix} 1 & 0 \\ -e^{-\zeta^2 + \frac{2iS}{\varepsilon}} & 1 \end{bmatrix} = Y_- \begin{bmatrix} 1 & 0 \\ -\varepsilon^{-\varkappa} e^{-\zeta^2 - \frac{2i\Re(S)}{\varepsilon}} & 1 \end{bmatrix} =: Y_- \begin{bmatrix} 1 & 0 \\ -\varepsilon^{-\varkappa} e^{-\zeta^2 + i\frac{\vartheta}{\varepsilon}} & 1 \end{bmatrix} \quad (3-11)$$

Note that the oscillatory term  $e^{i\vartheta/\varepsilon}$  is not problematic because it has constant modulus, but the term  $\varepsilon^{-\varkappa}$  diverges for  $\varkappa > 0$ .

### 3.1 Long time asymptotics of $S(\eta)$ along the breaking curve

Let us recall that the scattering data  $f_0(z)$  is analytic in  $\mathbb{C}_+$  (solitonless case) and that the corresponding  $w(z) = \Im f_0(z) \text{sign}(z - \mu_+)$ , restricted to  $z \in \mathbb{R}$ , is admissible in the sense of [13]. We recall that the gist of the notion of admissibility for  $w$  is that  $|r_0(z)|$  has exponential growth on the interval  $(\mu_-, \mu_+)$  as  $\varepsilon \rightarrow 0$ , exponential decay outside  $[\mu_-, \mu_+]$ , and  $w$  is piece-wise continuously differentiable with  $w'(\mu_\pm) \neq 0$ . As it was mentioned above, the condition of the absence of solitons is not essential if the contour  $\gamma$  of the matrix RHP can be chosen so that it does not intersect singularities of  $f_0(z)$ .

Under the above assumptions and using (2-17), (2-14), the system of equations for the point  $\eta = \eta(x, t)$  and the breaking curve  $t = t_0(x)$  can be written as

$$4t + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{w'(\zeta) d\zeta}{(\zeta - \eta) |R(\zeta)|} = 0 \quad (3-12)$$

$$\Im \left\{ R(\eta) \left[ 2t(\eta + a) + x + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{w(\zeta) d\zeta}{(\zeta - \eta) |R(\zeta)|} \right] \right\} = 0. \quad (3-13)$$

According to [13], Sect. 3.4, this system implies that the right branch of the breaking curve has the asymptotics

$$t(x) = -\frac{1}{4\mu_-}x + O(1), \quad x \rightarrow +\infty \quad (3-14)$$

and correspondingly

$$\eta = (\mu_- + O(t^{-1})) + i \left( \frac{w'(\mu_-)}{4t} [1 + o(1)] \right), \quad t \rightarrow \infty. \quad (3-15)$$

Moreover it was shown in Lemma 3.6 [13] that the branch-point  $\alpha(x, t)$  approaches  $\mu_+$  as  $x, t \rightarrow \infty$  along the right branch of the breaking curve according to the following asymptotic formula

$$b := \Im \alpha = e^{\frac{2\pi t(\mu_+ - \mu_-)}{w'(\mu_+)} + O(1)}. \quad (3-16)$$

The two later equations, together with Proposition 4.3 from [15], applied to

$$h(\eta) = R(\eta) \left[ 2t(\eta + a) + x + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{w(\zeta)d\zeta}{(\zeta - \eta)|R(\zeta)|} \right] \quad (3-17)$$

yield

$$S(\eta) \sim 4t\mu_-(\mu_- - \mu_+) \quad (3-18)$$

as  $\eta \rightarrow \infty$  along the right branch of the breaking curve. Using

$$h''(\eta) = \frac{R(\eta)}{2\pi i} \int_{\tilde{\gamma}_m} \frac{f'(\zeta)}{(\zeta - \eta)^2 R(\zeta)_+} d\zeta, \quad (3-19)$$

we can, in a similar way, calculate

$$C(x, t) \sim -4it \quad (3-20)$$

along the right branch of the breaking curve. Similar asymptotic analysis can be done for the left branch of the breaking curve.

In the case of  $f_0(z)$  given by (2-4), using the explicit expression for  $h''(z)$  given in [7], Section 4.5, we obtain

$$S(\eta) = -2tR(\eta)[\eta + a] + \sinh^{-1} \frac{\eta - a}{b} - i\frac{\pi}{2}. \quad (3-21)$$

## 4 First approximation

### 4.1 Outer parametrix

The construction of the outer parametrix, namely, the solution of the model problem in Section 2.2, is accomplished along the lines of [9]. The outer parametrix (aka *model problem*) is the unique solution of the RHP

$$\Psi_+(z) = \Psi_-(z) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{on } (\alpha, \bar{\alpha}) \quad (4-1)$$

$$\Psi(z) = \mathcal{O}((z - \alpha)^{-\frac{1}{4}}) \quad \Psi(z) = \mathcal{O}((z - \bar{\alpha})^{-\frac{1}{4}}) \quad (4-2)$$

$$\Psi(z) = \mathbf{1} + \mathcal{O}(z^{-1}), \quad z \rightarrow \infty. \quad (4-3)$$

$\Psi(z)$  is explicitly given by (4-13) below. Following the ideas in [9] we introduce the so-called **Schlesinger transformation** of the model problem

**Problem 4.1 (Schlesinger-chain of model problems)** *Find a matrix  $\Psi_\kappa(z)$  analytic off  $[\alpha, \bar{\alpha}] \cup \{\eta, \bar{\eta}\}$ ,  $\eta \notin [\alpha, \bar{\alpha}]$ , such that*

$$\Psi_{\kappa+}(z) = \Psi_{\kappa-}(z) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{on } (\alpha, \bar{\alpha}) \quad (4-4)$$

$$\Psi_\kappa(z) = \mathcal{O}((z - \alpha)^{-\frac{1}{4}}) \quad \Psi_\kappa(z) = \mathcal{O}((z - \bar{\alpha})^{-\frac{1}{4}}) \quad (4-5)$$

$$\Psi_\kappa(z) = \mathbf{1} + \mathcal{O}(z^{-1}), \quad z \rightarrow \infty \quad (4-6)$$

$$\Psi_\kappa(z) = \mathcal{O}(1)(z - \eta)^{-K\sigma_3}, \quad z \rightarrow \eta \quad (4-7)$$

$$\Psi_\kappa(z) = \mathcal{O}(1)(z - \bar{\eta})^{K\sigma_3}, \quad z \rightarrow \bar{\eta}. \quad (4-8)$$

We have formulated the problem assuming that there is only one main arc (cut) but the modifications in the case of multiple arcs is straightforward (see [9]); the main point is that we have modified the model problem *only* by adding the behavior (4-7) and (4-8) near the points where a new cut is about to emerge.

**Uniformization of the plane without the cut.** Consider the complex  $z$ -plane with a slit on the segment  $[\alpha, \bar{\alpha}]$ . It is elementary to show that the map

$$z(\lambda) = \frac{\alpha - \bar{\alpha}}{4i} \left( \lambda - \frac{1}{\lambda} \right) + \frac{\alpha + \bar{\alpha}}{2} = \frac{b}{2} \left( \lambda - \frac{1}{\lambda} \right) + a, \quad (4-9)$$

where  $\alpha = a + ib$ , maps univalently and biholomorphically the outside of the unit disk in the  $\lambda$ -plane  $\{|\lambda| > 1\}$  to  $\mathbb{C} \setminus [\alpha, \bar{\alpha}]$ , with the unit  $\lambda$ -circle being mapped onto the segment two-to-one. The points  $\lambda = \pm i$  are mapped to  $\alpha, \bar{\alpha}$  respectively. If we choose another (simple) curve joining  $\alpha$  to  $\bar{\alpha}$  then its counter-image in the  $\lambda$ -plane will be some deformed curve passing through  $\lambda = \pm i$  and encircling the  $\lambda$ -origin. Note the symmetry

$$z(-\lambda^{-1}) = z(\lambda). \quad (4-10)$$

Therefore the outside and inside of  $|\lambda| = 1$  are conformally mapped to two copies of  $z \in \mathbb{C} \setminus [\alpha, \bar{\alpha}]$ . We will refer to the region  $|\lambda| > 1$  as the **physical sheet**, as opposed to the **unphysical sheet**  $|\lambda| < 1$ .

The points  $z = \eta, \bar{\eta}$  have unique pre-images  $\lambda_0, \bar{\lambda}_0$  on the physical sheet

$$z(\lambda_0) = \eta, \quad z(\bar{\lambda}_0) = \bar{\eta},$$

We then have

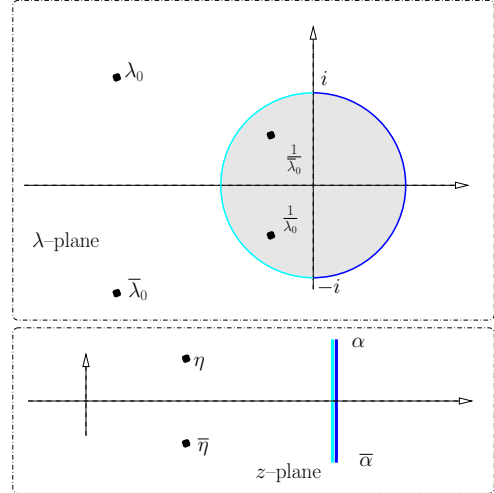


Figure 5: The uniformizing plane: the shaded region is the unphysical sheet.  $|\lambda_0| > 1$ . (4-11)

**Proposition 4.1** *The solution of the RHP (4.1) exists (unique) for any  $K \in \mathbb{Z}$  and it is given by*

$$\Psi_K = \begin{pmatrix} \bar{\lambda}_0 \\ \lambda_0 \end{pmatrix}^{K\sigma_3} \Psi_0(z) \begin{pmatrix} \lambda - \lambda_0 \\ \lambda\lambda_0 + 1 \end{pmatrix}^{-K\sigma_3} \begin{pmatrix} \lambda - \bar{\lambda}_0 \\ \lambda\bar{\lambda}_0 + 1 \end{pmatrix}^{K\sigma_3} \Big|_{\lambda=\lambda(z)}, \quad (4-12)$$

where  $\lambda(z)$  is the determination on the physical sheet of  $z \in \mathbb{C} \setminus [\alpha, \bar{\alpha}]$  and  $\lambda_0$  (defined in (4-11)) is the counter-image of  $\eta$  on the physical sheet. The matrix  $\Psi_0$  is given by

$$\Psi_0(z) := \frac{1}{2} \begin{bmatrix} -i & -1 \\ 1 & i \end{bmatrix} \begin{pmatrix} z - \alpha \\ z - \bar{\alpha} \end{pmatrix}^{\sigma_3/4} \begin{bmatrix} i & 1 \\ -1 & -i \end{bmatrix} = \begin{pmatrix} z - \alpha \\ z - \bar{\alpha} \end{pmatrix}^{\sigma_2/4}. \quad (4-13)$$

**Proof.** It is well known that the matrix  $\Psi_0$  solves the RHP

for  $K = 0$ . Since the interchange of sheets corresponds to the map  $\lambda \mapsto -\frac{1}{\lambda}$  one immediately verifies that the Szegő function

$$D(z) := \frac{\lambda\lambda_0 + 1}{\lambda - \lambda_0} \frac{\lambda - \bar{\lambda}_0}{\lambda\bar{\lambda}_0 + 1} \quad (4-14)$$

satisfies  $D_+(z)D_-(z) = 1$  on the cut, and  $\lim_{z \rightarrow \infty} D(z) = \frac{\lambda_0}{\lambda_0}$ . This is the reason for the left normalization in eq. 4-12. **Q.E.D.**

In terms of  $\lambda = \lambda(z)$ , the matrix  $\Psi_0$  is given by

$$\Psi_0(z) = \frac{1}{\sqrt{\lambda^2 + 1}} (\lambda \mathbf{1} - i\sigma_2) = \frac{1}{\sqrt{\lambda^2 + 1}} \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \quad (4-15)$$

**Remark 4.1** *Since  $\frac{z-a}{b} = (\lambda - \frac{1}{\lambda})$ , the uniformizing variable  $\lambda$  is related with the hyperbolic variable  $v$ , defined by  $\sinh v = \frac{z-a}{b}$  from Sect. 4.5, [7], through  $\lambda = e^v$ . Thus,*

$$\lambda(z) = e^v = \frac{z - a + R(z)}{b}, \quad (4-16)$$

where the square root  $R$  is chosen so that  $\lambda(z)$  is on the physical sheet.

**Lemma 4.1** *The matrices  $\Psi_K$  are related by*

$$\Psi_{K+1}(z) = R_K(z) \Psi_K(z), \quad R_K(z) = \mathbf{1} - \frac{C_K}{z - \eta} - \frac{\widehat{C}_K}{z - \bar{\eta}}, \quad R_K^{-1} = \mathbf{1} - \frac{Q_K}{z - \eta} - \frac{\widehat{Q}_K}{z - \bar{\eta}} \quad (4-17)$$

**Proof.** The matrices  $\Psi_K$  all have the same jumps and the same asymptotic behavior at the branch-points  $\alpha, \bar{\alpha}$  and at infinity. Thus  $R_K := \Psi_{K+1} \Psi_K^{-1}$  is a holomorphic matrix (without jumps) with possibly isolated singularities only at  $\eta, \bar{\eta}$ . The asymptotic behaviours of  $\Psi_{K+1}$  and  $\Psi_K$  at these two points imply that  $R_K$  may have at most simple poles there. The same reasoning can be applied to  $\Psi_K \Psi_{K+1}^{-1}$ , whence the formula for  $R_K^{-1}$ . **Q.E.D.**

It is possible to give explicit formulæ for the matrices  $C_K, \widehat{C}_K, Q_K, \widehat{Q}_K$  in terms of the columns of  $\Psi_K$  but they are of no immediate interest.

#### 4.1.1 Reality

The matrices  $\Psi_K(z)$  can be written as

$$\Psi_K(z) = \Phi_K(z)(z-\eta)^{-K\sigma_3} = \left[ \mathbf{A}_K(z), \mathbf{B}_K(z) \right] (z-\eta)^{-K\sigma_3} = \widehat{\Phi}_K(z)(z-\bar{\eta})^{K\sigma_3} = \left[ \widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z) \right] (z-\bar{\eta})^{K\sigma_3} \quad (4-18)$$

where the columns  $\mathbf{A}_K, \mathbf{B}_K$  are analytic at  $z = \eta$  and, respectively, the columns  $\widehat{\mathbf{A}}_K, \widehat{\mathbf{B}}_K$  are analytic at  $z = \bar{\eta}$ .

Note that, since  $\det \Psi_K \equiv 1$ , we have

$$\det \left[ \mathbf{A}_K(z), \mathbf{B}_K(z) \right] \equiv 1 \equiv \det \left[ \widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z) \right]. \quad (4-19)$$

Due to the symmetry of Prop. 2.1 we have

$$\Psi_K(z)(\Psi_K(\bar{z}))^\dagger \equiv \mathbf{1}, \quad (4-20)$$

which implies  $\Phi_K(z)(\Phi_K(\bar{z}))^\dagger \equiv 1$  or

$$\left[ \mathbf{A}_K(z), \mathbf{B}_K(z) \right] \equiv \left( \left[ \widehat{\mathbf{A}}_K(\bar{z}), \widehat{\mathbf{B}}_K(\bar{z}) \right]^\dagger \right)^{-1}. \quad (4-21)$$

This is promptly shown to translate into the relations

$$\mathbf{A}_K(z) \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \overline{\widehat{\mathbf{B}}_K(\bar{z})}, \quad \mathbf{B}_K(z) \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \overline{\widehat{\mathbf{A}}_K(\bar{z})}, \quad (4-22)$$

or

$$\Phi_K(z) = \sigma_2 \overline{\widehat{\Phi}_K(\bar{z})} \sigma_2, \quad (4-23)$$

which is equivalent to  $\Phi_K(z)(\Phi_K(\bar{z}))^\dagger \equiv \mathbf{1}$ . Similar considerations lead to

$$\Phi'_K(z) = \sigma_2 \overline{\widehat{\Phi}'_K(\bar{z})} \sigma_2, \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (4-24)$$

We will denote by  $\mathbf{A}_K$  the *value*  $\mathbf{A}_K(\eta)$ , and similarly

$$\mathbf{A}_K := \mathbf{A}_K(\eta), \quad \mathbf{B}_K := \mathbf{B}_K(\eta), \quad \mathbf{A}'_K := \mathbf{A}'_K(z)|_{z=\eta}, \quad \mathbf{B}'_K := \mathbf{B}'_K(z)|_{z=\eta}, \quad (4-25)$$

$$\widehat{\mathbf{A}}_K := \widehat{\mathbf{A}}_K(\bar{\eta}), \quad \widehat{\mathbf{B}}_K := \widehat{\mathbf{B}}_K(\bar{\eta}), \quad \widehat{\mathbf{A}}'_K := \widehat{\mathbf{A}}'_K(z)|_{z=\bar{\eta}}, \quad \widehat{\mathbf{B}}'_K := \widehat{\mathbf{B}}'_K(z)|_{z=\bar{\eta}}. \quad (4-26)$$

This implies the following relations, which will be of use at a later stage in the paper

$$\begin{aligned} \det [\mathbf{A}'_K, \mathbf{A}_K] &= \overline{\det [\widehat{\mathbf{B}}'_K, \widehat{\mathbf{B}}_K]} & \det [\mathbf{B}'_K, \mathbf{B}_K] &= \overline{\det [\widehat{\mathbf{A}}'_K, \widehat{\mathbf{A}}_K]}, \\ \det [\widehat{\mathbf{A}}_K, \mathbf{B}_K] &= \mathbf{B}_K^\dagger \mathbf{B}_K > 0 & \det [\mathbf{A}_K, \widehat{\mathbf{B}}_K] &= \mathbf{A}_K^\dagger \mathbf{A}_K > 0. \end{aligned} \quad (4-27)$$

These quantities are computed explicitly in App. A

## 4.2 Local parametrices near $\eta, \bar{\eta}$ and first approximation to the solution

We only consider the point  $\eta$ , with similar considerations being repeated for  $\bar{\eta}$ . The local parametrix must solve the following **mixed RHP** where the adjective “mixed” means that the requirements are for both **columns and rows**:

$$\mathbf{H}_{K^+}(z) = \mathbf{H}_{K^-}(z) \begin{bmatrix} 1 & 0 \\ -\varepsilon^{-\varkappa} e^{i\theta/\varepsilon} e^{-\zeta^2} & 1 \end{bmatrix}; \quad (4-28)$$

$$\mathbf{H}_K(z) = (z - \eta)^{K\sigma_3} \mathcal{O}(1); \quad (4-29)$$

$$\mathbf{H}_K(z) = \mathbf{1} + o(1), \quad z \in \partial\mathbb{D}, \quad (4-30)$$

where  $o(1)$  above means an infinitesimal in  $\varepsilon$ , uniform w.r.t.  $z$  when  $z$  belongs to the boundary of the disk. When necessary to specify the values of the parameters  $\varkappa, \vartheta$  we will write

$$\mathbf{H}_{K^+}(z) = \mathbf{H}_{K^+}(z; \varkappa, \vartheta). \quad (4-31)$$

## 4.3 Approximate solution and error analysis

Suppose that we are able to solve the mixed RHP in (4-30) and define the following matrix

$$\tilde{Y}(z) := \begin{cases} \Psi_K(z) & \text{outside of the local disks,} \\ \Psi_K(z)\mathbf{H}_K(z) & \text{within the disks at } \eta, \bar{\eta}, \\ \Psi_K(z)\mathcal{A}(z) & \text{within the disks at } \alpha, \bar{\alpha}, \end{cases} \quad (4-32)$$

where  $\mathcal{A}(z)$  is the usual Airy-parametrix [14]. Note that the matrix  $\Psi_K(z)\mathbf{H}_K(z)$  is **bounded** in a neighborhood of  $\eta$  because the singular behavior of the columns of  $\Psi_K$  (4-7) is exactly canceled by the singular behavior (4-29) of the rows of  $\mathbf{H}_K$ . In order to evaluate how close the matrix  $\tilde{Y}(z)$  is to the matrix  $Y(z)$  we take their ratio

$$\mathcal{E}(z) := Y(z)\tilde{Y}(z)^{-1} \quad (4-33)$$

and consider the RHP it solves:

$$\mathcal{E}(z) \sim \mathbf{1} + \mathcal{O}(z^{-1}), \quad \mathcal{E}_+(z) = \mathcal{E}_-(z)W(z), \quad (4-34)$$

where  $W(z)$  is a matrix supported on the contours indicated in Fig. 6. In particular, on the boundary of the disk  $\mathbb{D}$  centered at  $\eta$  we have

$$\mathcal{E}_+(z) = \mathcal{E}_-(z)\Psi_K\mathbf{H}_K(z)\Psi_K^{-1}(z), \quad z \in \partial\mathbb{D}. \quad (4-35)$$

Since  $\Psi_K$  are uniformly bounded for  $z \in \partial\mathbb{D}$  as  $\varepsilon \rightarrow 0$  (in fact they are independent of  $\varepsilon$ ), it follows that the jump of  $\mathcal{E}$  on  $\partial\mathbb{D}$  is  $\mathbf{1} + o(1)$ , where this is the same

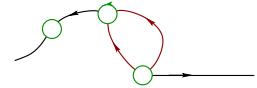


Figure 6: The contours for the RHP of the error term (in  $\mathbb{C}_+$ ).

infinitesimal that appears in eq. (4-30) (and we will promptly see below that this desired infinitesimal is not such when  $\varkappa$  is a half-integer, thus requiring a more refined approach). The usual theorem about Riemann–Hilbert problems with small-norm jumps imply that  $\mathcal{E}(z)$  is uniformly close to the identity to within the same infinitesimal.

We will see in Section 4.4 that it is possible to solve the local problem for  $\mathbf{H}_K$  formulated above with  $o(1) = \mathcal{O}(\varepsilon^{\frac{1}{2}-|\varkappa-K|})$  and hence:

- the integer  $K$  must be chosen as the nearest positive integer to  $\varkappa$ ;
- it is impossible to solve the problem as formulated for a half-integer  $\varkappa = \frac{1}{2} + \mathcal{N}$ , where  $\mathcal{N} \in \mathbb{N}$ .

We will show below how to solve the problem as formulated, why it is impossible to solve in those exceptional situation and how to circumvent the problem.

#### 4.4 Solution of the local RHP for $\mathbf{H}_K$ (4-28).

Let  $H_\ell(s)$  be the **monic** Hermite polynomials of degree  $\ell$ :

$$\int_{\mathbb{R}} H_\ell(s) H_{\ell'}(s) e^{-s^2} ds = \frac{\ell! \sqrt{\pi}}{2^\ell} \delta_{\ell\ell'} =: \mu_\ell \delta_{\ell\ell'}, \quad H_\ell(s) := (-2)^{-\ell} e^{s^2} \frac{d^\ell}{ds^\ell} e^{-s^2}. \quad (4-36)$$

We define the local parametrix as

$$\mathbf{H}_K(\zeta) := e^{-i\frac{\vartheta}{2\varepsilon}\sigma_3} \varepsilon^{\frac{\varkappa}{2}\sigma_3} \zeta^{K\sigma_3} \begin{bmatrix} \frac{-1}{\mu_{K-1}} \int_{\mathbb{R}} \frac{H_{K-1}(s) e^{-s^2}}{s-\zeta} ds & \frac{-2i\pi}{\mu_{K-1}} H_{K-1}(\zeta) \\ \frac{1}{2i\pi} \int_{\mathbb{R}} \frac{H_K(s) e^{-s^2}}{s-\zeta} ds & H_K(\zeta) \end{bmatrix} \varepsilon^{-\frac{\varkappa}{2}\sigma_3} e^{i\frac{\vartheta}{2\varepsilon}\sigma_3} \\ = e^{-i\frac{\vartheta}{2\varepsilon}\sigma_3} \begin{bmatrix} \frac{-\zeta^K}{\mu_{K-1}} \int_{\mathbb{R}} \frac{H_{K-1}(s) e^{-s^2}}{s-\zeta} ds & \frac{-2i\pi \zeta^K \varepsilon^{\varkappa}}{\mu_{K-1}} H_{K-1}(\zeta) \\ \frac{\zeta^{-K} \varepsilon^{-\varkappa}}{2i\pi} \int_{\mathbb{R}} \frac{H_K(s) e^{-s^2}}{s-\zeta} ds & \frac{H_K(\zeta)}{\zeta^K} \end{bmatrix} e^{i\frac{\vartheta}{2\varepsilon}\sigma_3} \quad (4-37)$$

$$(4-38)$$

**Boundedness.** The matrix  $\Psi_K \mathbf{H}_K$  is bounded in  $\mathbb{D}$ , because of the  $\zeta^{K\sigma_3}$  factor in (4-37) cancels out the singularity of  $\Psi_K$  (4-7).

**Error estimate.** Recall that for  $z \in \partial\mathbb{D}$  the zooming coordinate  $\zeta$  scales like  $\frac{1}{\sqrt{\varepsilon}}$ . Therefore the matrix  $\mathbf{H}_K$  on the boundary of the disk  $\mathbb{D}$  has the following behavior uniformly w.r.t.  $z$

$$\mathbf{H}_K = \begin{bmatrix} 1 + \mathcal{O}(\varepsilon) & \mathcal{O}(\varepsilon^{\frac{1}{2}+\varkappa-K}) \\ \mathcal{O}(\varepsilon^{\frac{1}{2}-\varkappa+K}) & 1 + \mathcal{O}(\varepsilon) \end{bmatrix} = \mathbf{1} + \mathcal{O}(\varepsilon^{\frac{1}{2}-|\varkappa-K|}), \quad z \in \partial\mathbb{D}. \quad (4-39)$$

Note that the diagonal estimates are  $1 + \mathcal{O}(\varepsilon)$  and not just  $1 + \mathcal{O}(\sqrt{\varepsilon})$  due to the fact that the Hermite polynomials are even/odd functions according to the parity of their degrees. For example, for the (1, 2) element in (4-37):  $H_{K-1}(\zeta) = \zeta^{K-1}(1 + \mathcal{O}(\zeta^{-1}))$  and on  $\partial\mathbb{D}$  we have  $\zeta = \mathcal{O}(1/\sqrt{\varepsilon})$  and

$$\frac{2i\pi \zeta^K e^{-i\frac{\vartheta}{\varepsilon}\sigma_3}}{\mu_{K-1}} H_{K-1} = \frac{2i\pi \zeta^{2K-1} e^{-i\frac{\vartheta}{\varepsilon}\sigma_3}}{\mu_{K-1}} (1 + \mathcal{O}(\sqrt{\varepsilon})) = \mathcal{O}(\varepsilon^{\varkappa-K+\frac{1}{2}}). \quad (4-40)$$



Similarly, for the  $(2, 1)$  element, using the orthogonality, we have that the Cauchy transform is  $\mathcal{O}(\zeta^{-K-1})$  and hence

$$-\frac{\zeta^{-K} e^{i\frac{\vartheta}{\varepsilon}} \varepsilon^{-\varkappa}}{2i\pi} \int \frac{H_K(s) e^{-s^2}}{s - \zeta} ds = \frac{\mu_K}{2i\pi} \zeta^{-2K-1} e^{i\frac{\vartheta}{\varepsilon}} \varepsilon^{-\varkappa} (1 + \mathcal{O}(\zeta^{-1})) = \mathcal{O}(\varepsilon^{K+\frac{1}{2}-\varkappa}). \quad (4-41)$$

We immediately see that –in order to meet the requirement about the decay on the boundary of the disk  $\mathbb{D}$  (4-30)– we must have

**Proposition 4.2** *The integer  $K$  must be the closest integer to  $\varkappa$ .*

The important observation is that if  $\varkappa \in \frac{1}{2} + \mathbb{Z}$  then the error term in (4-39) does not tend to zero (it is  $\mathcal{O}(1)$ ). It is understandable as these values separate regimes where the value of  $K$  jumps by one unit and the whole strong asymptotic must changes its form.

## 5 Improved approximation

### 5.1 Next order improved local parametrix

Inspecting directly and a bit closer the expression (4-37) we see that

$$\mathbf{H}_K = \begin{bmatrix} 1 + \mathcal{O}(\varepsilon) & \frac{-2i\pi\zeta^{2K-1} e^{-i\frac{\vartheta}{\varepsilon}} \varepsilon^\varkappa}{\mu_{K-1}} + \mathcal{O}(\varepsilon^{\frac{3}{2}+\varkappa-K}) \\ \frac{-\mu_K}{2i\pi\zeta^{2K+1} e^{-i\frac{\vartheta}{\varepsilon}} \varepsilon^\varkappa} + \mathcal{O}(\varepsilon^{\frac{3}{2}-\varkappa+K}) & 1 + \mathcal{O}(\varepsilon) \end{bmatrix} = \mathbf{1} + \mathcal{O}(\varepsilon^{\frac{1}{2}-|\varkappa-K|}), \quad z \in \partial\mathbb{D}. \quad (5-1)$$

and the terms that are responsible for the loss in the error term (4-39) are the off-diagonal ones. In particular:

- for  $K < \varkappa < K + \frac{1}{2}$  the leading term in the error is the **lower triangular** one;
- for  $K - \frac{1}{2} < \varkappa < K$  the leading term in the error is the **upper triangular** one.

For this reason we re-define

**Definition 5.1** *The local parametrix  $\mathbf{H}_\varkappa$ ,  $\varkappa \in \mathbb{R}_+$  is defined as*

$$\mathbf{H}_\varkappa = \begin{cases} \begin{bmatrix} 1 & 0 \\ \frac{\mu_K}{2i\pi\zeta^{2K+1} e^{-i\frac{\vartheta}{\varepsilon}} \varepsilon^\varkappa} & 1 \end{bmatrix} \mathbf{H}_K & K < \varkappa \leq K + \frac{1}{2} \\ \mathbf{H}_K & \varkappa = K \\ \begin{bmatrix} 1 & \frac{2i\pi\zeta^{2K-1} e^{-i\frac{\vartheta}{\varepsilon}} \varepsilon^\varkappa}{\mu_{K-1}} \\ 0 & 1 \end{bmatrix} \mathbf{H}_K & K - \frac{1}{2} < \varkappa < K. \end{cases} \quad (5-2)$$

**Parametrix on the disk around  $\bar{\eta}$ .** Because of the symmetry of Prop. 2.1 we have

$$\widehat{\mathbf{H}}_{\varkappa}(\zeta) := (\mathbf{H}_{\varkappa}(\bar{\zeta})^\dagger)^{-1} . \quad (5-3)$$

In the formulæ below, we will use  $\widehat{\zeta}(z) := \overline{\zeta(\bar{z})}$ .

## 5.2 Improved outer parametrix

Corresponding to the definition of  $\mathbf{H}_{\varkappa}$  (Def. 5.1) we need to define the outer parametrix  $\Psi_{\varkappa}$  with the requirements

$$\Psi_{\varkappa}(z)_+ = \Psi_{\varkappa}(z)_- \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} , \quad (5-4)$$

$$\Psi_{\varkappa}(z) = \mathcal{O}((z - \alpha)^{-\frac{1}{4}}) \quad \Psi_{\varkappa}(z) = \mathcal{O}((z - \bar{\alpha})^{-\frac{1}{4}}), \quad (5-5)$$

$$\Psi_{\varkappa}(z) = \mathbf{1} + \mathcal{O}(z^{-1}), \quad (5-6)$$

$$\Psi_{\varkappa}(z)\mathbf{H}_{\varkappa}(\zeta) = \mathcal{O}(1) , \quad z \in \mathbb{D}, \quad (5-7)$$

$$\Psi_{\varkappa}(z)\widehat{\mathbf{H}}_{\varkappa}(\widehat{\zeta}) = \mathcal{O}(1) , \quad z \in \widehat{\mathbb{D}}. \quad (5-8)$$

In correspondence with Def. 5.1 we will pose the following **Ansatz** for  $\Psi_{\varkappa}$ .

**Theorem 5.1** *The solution of the Riemann–Hilbert problem (5-4)-(5-8) always exists and is given by*

$$\Psi_{\varkappa}(z) = \begin{cases} T_{K,\varkappa}(z)\Psi_K(z) = \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} \right) \Psi_K(z) & K < \varkappa \leq K + \frac{1}{2} \\ \Psi_K(z) & \varkappa = K \in \mathbb{N} \\ U_{K,\varkappa}(z)\Psi_K(z) = \left( \mathbf{1} - \frac{F_{K,\varkappa}}{z - \eta} - \frac{\widehat{F}_{K,\varkappa}}{z - \bar{\eta}} \right) \Psi_K(z) & K - \frac{1}{2} < \varkappa < K, \end{cases} \quad (5-9)$$

where

$$G_{K,\varkappa}(\bullet) = \frac{u_K \det[\bullet, \mathbf{B}_K]}{\left| 1 + u_K \det[\mathbf{B}'_K, \mathbf{B}_K] \right|^2 + \frac{|u_K|^2 \|\mathbf{B}_K\|^4}{4\Im(\eta)^2}} \left( \left( 1 + \bar{u}_K \overline{\det[\mathbf{B}'_K, \mathbf{B}_K]} \right) \mathbf{B}_K - \frac{\bar{u}_K \|\mathbf{B}_K\|^2}{2\Im(\eta)} \sigma_2 \overline{\mathbf{B}_K} \right), \quad (5-10)$$

$$F_{K,\varkappa}(\bullet) = \frac{\ell_K \det[\mathbf{A}_K, \bullet]}{\left| 1 - \ell_K \det[\mathbf{A}'_K, \mathbf{A}_K] \right|^2 + \frac{|\ell_K|^2 \|\mathbf{A}_K\|^4}{4\Im(\eta)^2}} \left( \left( 1 - \bar{\ell}_K \overline{\det[\mathbf{A}'_K, \mathbf{A}_K]} \right) \mathbf{A}_K - \frac{\bar{\ell}_K \|\mathbf{A}_K\|^2}{2\Im(\eta)} \sigma_2 \overline{\mathbf{A}_K} \right), \quad (5-11)$$

$$\widehat{G}_{K,\varkappa} = \sigma_2 \overline{G_{K,\varkappa}} \sigma_2, \quad \widehat{F}_{K,\varkappa} = \sigma_2 \overline{F_{K,\varkappa}} \sigma_2, \quad (5-12)$$

$$u_K := \frac{K! \sqrt{\pi} C^{-K-\frac{1}{2}}}{2^{K+1} i \pi} \varepsilon^{\frac{1}{2}-\varkappa+K} e^{i\frac{\vartheta}{\varepsilon}}, \quad \ell_K := \frac{2i\pi e^{-i\frac{\vartheta}{\varepsilon}} \varepsilon^{\varkappa-K+\frac{1}{2}}}{\mu_{K-1} C^{-K+\frac{1}{2}}} . \quad (5-13)$$

The proof can be found in App. B.

**Remark 5.1** *The idea behind this Ansatz is the following: the outer parametrices  $\Psi_K$  and  $\Psi_{K+1}$  are related by a finite Schlesinger transformation of the form (4-17) and we know that they provide the correct leading behavior for  $\varkappa \in (K - \frac{1}{2}, K + \frac{1}{2})$  and  $\varkappa \in (K + \frac{1}{2}, K + \frac{3}{2})$ , respectively. The discontinuity in this leading-order asymptotic behavior occurs at  $\varkappa = K + \frac{1}{2}$ , where the error term as per (4-39) becomes of order 1; the idea is thus to introduce some sort of interpolation between the two outer parametrices of the same form of a finite Schlesinger transformation, letting the coefficients matrices be dependent on  $\varkappa$  in such a way that their order becomes 1 as  $\varkappa$  crosses the critical value.*

Suppose  $\varkappa > K + \frac{1}{2}$  in the formula (5-10) for  $G_{K,\varkappa}(\widehat{G}_{K,\varkappa})$ ; then one sees that  $u_K$  diverges as  $\varepsilon \rightarrow 0$ . The matrix  $G_{K,\varkappa}(\widehat{G}_{K,\varkappa})$ , however, has a regular limit  $G_{K,\varkappa} \rightarrow C_K$ ,  $(\widehat{G}_{K,\varkappa} \rightarrow \widehat{C}_K)$ , while for  $\varkappa < K - \frac{1}{2}$  we have  $F_{K,\varkappa} \rightarrow Q_K$ ,  $(\widehat{F}_{K,\varkappa} \rightarrow \widehat{Q}_K)$  where  $C_K, \widehat{C}_K, Q_K, \widehat{Q}_K$  have been introduced in (4-17); indeed the defining equations for  $G, \widehat{G}_{K,\varkappa}$  turn into those of  $C_K, \widehat{C}_K$  (ditto for  $F_{K,\varkappa}, \widehat{F}_{K,\varkappa}$  versus  $Q_K, \widehat{Q}_K$ ). This means that  $G_{K,\varkappa}, F_{K,\varkappa}$  are some sort of **interpolation** between Schlesinger transformations.

**Continuity at  $\varkappa = K + \frac{1}{2}$ .** We must check that the matrix  $\Psi_\varkappa$  as defined in eq. (5-9) is continuous as  $\varkappa = K + \frac{1}{2}$ , at least within the error  $\mathcal{O}(\sqrt{\varepsilon})$ . In fact we are going to see below (Prop. 5.1) that this continuity holds exactly, namely,

$$\lim_{\varkappa \rightarrow K + \frac{1}{2} - 0} \Psi_\varkappa = \lim_{\varkappa \rightarrow K + \frac{1}{2} + 0} \Psi_\varkappa \quad (5-14)$$

or, equivalently,

$$\left( \mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} \right)_{\varkappa = K + \frac{1}{2}} \Psi_K(z) = \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z - \eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z - \bar{\eta}} \right)_{\varkappa = K + \frac{1}{2}} \Psi_{K+1}(z). \quad (5-15)$$

Even more is true as eq. (5-15) holds for **any value** of  $\varkappa$ . To show this we recall that  $\Psi_{K+1} = R_K \Psi_K$  (as per eq. 4-17). Thus, we prove the following

**Proposition 5.1** For any value of  $\varkappa \in \mathbb{R}$ , the matrices  $G_{K,\varkappa}, \widehat{G}_{K,\varkappa}, F_{K+1,\varkappa}, \widehat{F}_{K+1,\varkappa}$  are related with the  $C_K, \widehat{C}_{K+1}$  by

$$\mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} = \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z - \eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z - \bar{\eta}} \right) \left( \mathbf{1} - \frac{C_K}{z - \eta} - \frac{\widehat{C}_K}{z - \bar{\eta}} \right). \quad (5-16)$$

In particular, we have

$$G_{K,\varkappa} + \widehat{G}_{K,\varkappa} = F_{K+1,\varkappa} + \widehat{F}_{K+1,\varkappa} + C_K + \widehat{C}_K. \quad (5-17)$$

The proof can be found in Appendix B.

Thanks to Proposition 5.1 we conclude also that in Theorem 5.1

$$T_{K,\varkappa}(z) \Psi_K(z) \equiv U_{K+1,\varkappa}(z) \Psi_{K+1}(z), \quad \forall \varkappa, z, \varepsilon, \quad (5-18)$$

which allows us to dispose of the necessity for the  $F_K$  altogether and rewrite Theorem 5.1 as

**Corollary 5.1** *The solution to the RHP (5-4)-(5-8) given by Theorem 5.1 can be equivalently written as*

$$\Psi_{\varkappa}(z) = \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} \right) \Psi_K(z), \quad K = \max(0, \lfloor \varkappa \rfloor), \quad (5-19)$$

with the matrices being defined in (5-10).

We remind the reader that  $\Psi_K$  (with integer subscript) had been defined independently in Proposition 4.1.

**Remark 5.2** *We also remark that the new definition in Cor. 5.1 is **not continuous** in  $\varkappa$  at the integers (but continuous at the half-integer). However the discontinuity is to within  $\mathcal{O}(\sqrt{\varepsilon})$ . More precisely,*

$$\lim_{\varkappa \rightarrow K^-} \Psi_{\varkappa}(z) = \left( \mathbf{1} + \mathcal{O}(\sqrt{\varepsilon}) \left( \frac{1}{z - \eta} + \frac{1}{z - \bar{\eta}} \right) \right) \lim_{\varkappa \rightarrow K^+} \Psi_{\varkappa}(z) \quad (5-20)$$

with the  $\mathcal{O}(\sqrt{\varepsilon})$  factor being independent of  $z$ . This discontinuity is of no concern because the correction to the final approximation is within our overall error estimate.

## 6 The left breaking curve

The situation at the left breaking curve (Fig. 4) is at first sight quite different because it is the main arc that it is splitting into two, rather than the complementary arc. However, as we will presently see, the phenomenon is essentially identical: the informal reason is that the choice of main and complementary arcs in the construction of the function  $h$  is somewhat conventional, depending on the overall sign of  $h$  at infinity and the placement of the cuts of the  $g$ -function. The difference is nevertheless sufficiently stark to grant a separate analysis.

At the exact point of break the function  $h$  near the point  $\eta$  (where the arc is about to split) can be written

$$2ih(z) = \begin{cases} -C(z - \eta)^2(\mathbf{1} + \mathcal{O}(z - \eta)), & z \in \mathbb{D}_1, \\ C(z - \eta)^2(\mathbf{1} + \mathcal{O}(z - \eta)), & z \in \mathbb{D}_0, \end{cases} \quad (6-1)$$

where the region  $\mathbb{D}_1$  is the region (within the local disk) to the right of the two main arcs and  $\mathbb{D}_0$  its complement. If we let evolve the level curves according to the modulation equations while respecting the branch-cuts we would have the situation depicted on Fig. 8, with two pairs of close cuts running off to infinity (i.e. a new main arc appears to compensate the split of the original one).

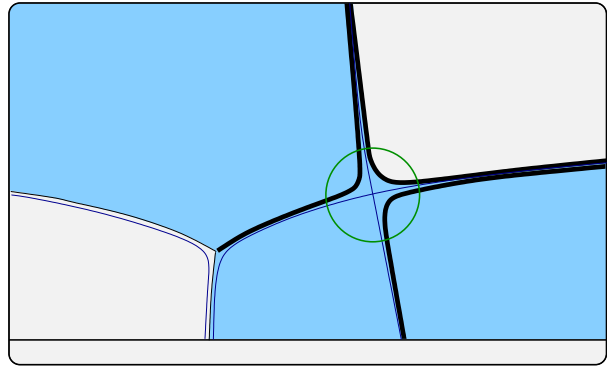


Figure 8: The main arc has split into two and we are forced to introduce a new main arc on the zero-level set of  $\Im(h)$ .

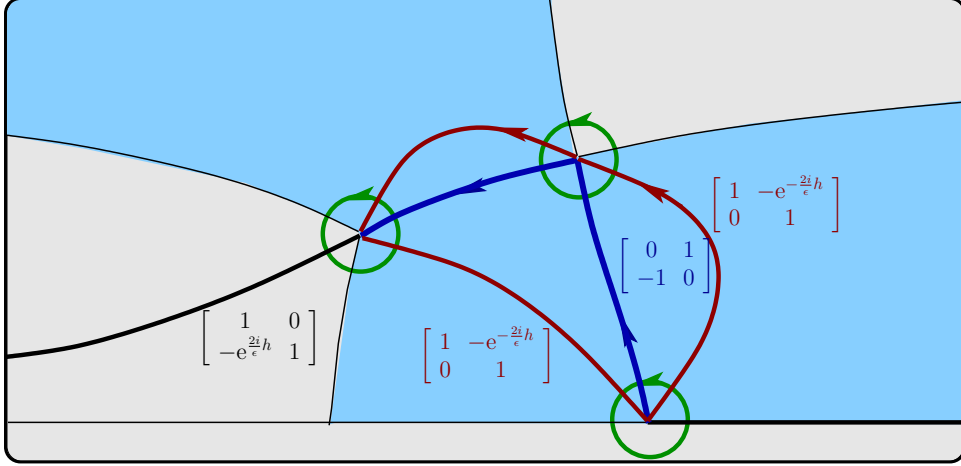


Figure 7: The typical shape of the main, complementary arcs and lens contours at the precise point on the left breaking curve.

We then must re-define the function  $h$  by introducing a small main arc within the local disk  $\mathbb{D}$  according to Fig. 9

The difference between the two pictures on Fig. 8 and Fig. 9 is that we have changed the sign of  $h$  within the region that has changed color; the result is a function that we still denote by  $h(z)$ , which is harmonic analytic on the upper complex plane minus the main arc (marked by the thick line). In the sequel we will denote by  $h(z)$  the function defined this way: it has the property that it is now **analytic** in a small neighborhood of  $z = \eta$  and behaves

$$2ih(z) = -C(z - \eta)^2(1 + \mathcal{O}(z - \eta)) + 2iS, \quad (6-2)$$

$$C = C(x, t), \quad \eta = \eta(x, t), \quad S := S(x, t). \quad (6-3)$$

The arc made by the two lenses can be deformed to pass through the level-curve  $\Re(h - S) = 0$  and the zooming coordinate  $\zeta$  can be defined as the conformal change of coordinate

$$\zeta^2 = \begin{cases} \frac{2i}{\epsilon}(h(z) - S), & z \in \mathbb{D}_1, \\ -\frac{2i}{\epsilon}(h(z) - S), & z \in \mathbb{D}_0. \end{cases} \quad (6-4)$$

We note that now we have  $\Im(S) > 0$  and hence it is convenient to parametrize it by

$$S := 2\vartheta + i2\kappa\epsilon \ln \epsilon, \quad (6-5)$$

where –contrary to the case of the right breaking curve– **negative** values of  $\kappa$  correspond to the exploration into the genus–two region. In Fig. 10 we have shown the jumps and contours for the solution  $Y(z)$  in the local disk  $\mathbb{D}$ . The model problem solves the usual jump  $\Psi_+ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Psi_-$ .

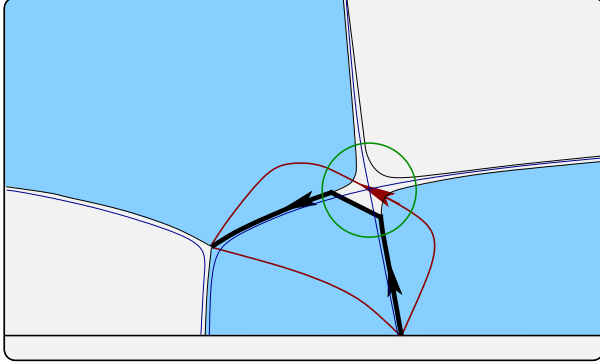


Figure 9: [On the left] The main arc has split into two and we are modifying the function  $h$  by introducing a small “artificial” main arc connecting the two pieces of the broken one.

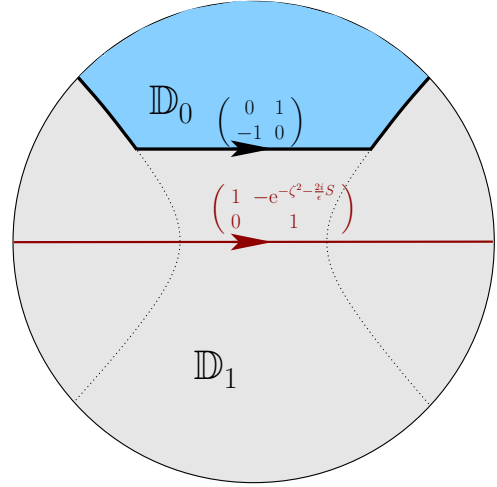


Figure 10: [On the right] The exact jumps of  $Y$  in terms of the conformal coordinate  $\zeta$  near the point  $\eta$ .

## 6.1 Outer parametrix

The situation from this point on is quite similar to the previous; the outer parametrix  $\Psi_J$  will be taken as defined by the same Riemann–Hilbert Problem 4.1, with the only difference that the jump condition now runs on the piecewise-smooth arc. Note that the point  $\eta$  (by our construction) lies to the right of the main arc; worth mentioning is also that this time we will have to use the formula for  $\Psi_J$  for the closest **negative integer** to  $\varkappa$ . The formulæ for  $\Psi_J$  are identical.

## 6.2 Local parametrix

Due to the different shape of the jumps within the local disk  $\mathbb{D}$ , the relationship between the present case and the previous ones are less immediate than for the model problem.

**Problem 6.1** *The local parametrix is a matrix-valued function  $\mathcal{P}_J(\zeta) = \mathcal{P}_J(\zeta; \varkappa, \vartheta)$  of the zooming coordinate  $\zeta$  satisfying the following **mixed RHP** ( $J \in -\mathbb{N}$ ).*

$$\mathcal{P}_{J+}(\zeta) = \begin{bmatrix} 1 & -\varepsilon^\varkappa e^{-\zeta^2 - \frac{i}{\varepsilon} \vartheta} \\ 0 & 1 \end{bmatrix} \mathcal{P}_{J-}(\zeta), \quad \zeta \in \mathbb{R}, \quad (6-6)$$

$$\mathcal{P}_{J+}(\zeta) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \mathcal{P}_{J-}(\zeta) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \zeta \in \text{main arc}, \quad (6-7)$$

$$\mathcal{P}_J(\zeta) = \zeta^{J\sigma_3} \mathcal{O}(1), \quad \zeta \rightarrow 0, \quad (6-8)$$

$$\mathcal{P}_J(\zeta) = \mathbf{1} + o(1), \quad \zeta \in \partial\mathbb{D}. \quad (6-9)$$

We anticipate that we will need this problem for negative integer  $K$ . If we disregard the condition (6-7) then this problem is essentially (up to the triangularity of the jump) the same problem as for  $\mathbf{H}_K$ ; more precisely, let  $K \in -\mathbb{N}$  and  $\varkappa \in -\mathbb{R}_+$  and consider

$$\mathbf{H}_{(-J)}(\zeta; -\varkappa, -\vartheta) \quad (6-10)$$

is the solution of (4-28), (4-29), (4-30) with  $K = -J$  and  $\varkappa, \vartheta$  replaced by  $-\varkappa, -\vartheta$ . Then the solution of the RHP 6.1 is

$$\mathcal{P}_J(\zeta; \varkappa, \vartheta) := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{H}_{(-J)}(\zeta; -\varkappa, -\vartheta) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \zeta \in \mathbb{D}_1, \quad (6-11)$$

$$\mathcal{P}_J(\zeta; \varkappa, \vartheta) := e^{\frac{i\pi}{2}\sigma_3} \mathbf{H}_{(-J)}(\zeta; -\varkappa, -\vartheta) e^{-\frac{i\pi}{2}\sigma_3}, \quad \zeta \in \mathbb{D}_0. \quad (6-12)$$

It is promptly seen that  $\mathcal{P}_J$  solves the jump on the real axis and –on the boundary between  $\mathbb{D}_0, \mathbb{D}_1$ – is satisfies the condition (6-7).

It should be apparent now that the sequel of the analysis is quite parallel to the previous case; the parametrices  $\Psi_\varkappa$  and  $\mathcal{P}_\varkappa$  will be constructed by the same methods. The explicit formulæ for the matrices  $G_J, \widehat{G}_J$  etc. have only minor cosmetic differences.

## 7 Corrections to the NLS solution

**Theorem 7.1** *The solution of the NLS (1-1) defined by the scattering data  $f_0(z)$  behaves as*

$$q(x, t, \varepsilon) = \left( -b(x, t) e^{2iK\psi(x, t)} - 2 \left( G_{K, \varkappa}(x, t, \varepsilon) + \widehat{G}_{K, \varkappa}(x, t, \varepsilon) \right)_{12} + \mathcal{O}(\sqrt{\varepsilon}) \right) e^{\frac{4i}{\varepsilon} g(\infty, x, t)} \quad (7-1)$$

as  $\varepsilon \rightarrow 0$ . Here  $(x, t)$  are such that

$$\Im(S(x, t)) = \frac{\varkappa}{2} \varepsilon \ln \varepsilon, \quad \varkappa \in \mathbb{R}, \quad (7-2)$$

where  $K = \max(0, \lfloor \varkappa \rfloor)$ ,  $e^{i\psi} := \frac{\bar{\lambda}_0}{\lambda_0}$  (for  $\lambda_0$  see (4-11)),  $G_{K, \varkappa}, \widehat{G}_{K, \varkappa}$  have been defined in (5-10). The term  $(G_{K, \varkappa} + \widehat{G}_{K, \varkappa})_{12}$  has the following behavior

$$2(G_{K, \varkappa} + \widehat{G}_{K, \varkappa})_{12} = \begin{cases} \mathcal{O}(\varepsilon^{\frac{1}{2} + K - \varkappa}), & \varkappa \in [K, K + \frac{1}{2}], \\ -b(x, t) \left[ e^{2iK\psi} - e^{-2i(K+1)\psi} \right] + \mathcal{O}(\varepsilon^{\varkappa - K - \frac{1}{2}}), & \varkappa \in (K + \frac{1}{2}, K + 1). \end{cases} \quad (7-3)$$

In particular the correction is of the order  $\mathcal{O}(1)$  one as  $\varepsilon \rightarrow 0$  only for  $\varkappa \in \frac{1}{2} + \mathbb{N}$ .

The proof can be found in Appendix B.

**Remark 7.1** *Considering  $\varkappa$  as a parameter in a direction transversal to the breaking curve, we notice that centers of adjacent ripples are at unit  $\varkappa$ -distance from each other. On the other hand, according to (7-3), the “support” of each ripple is of the order  $\mathcal{O}(\frac{1}{\ln \varepsilon})$  in terms of the exploration parameter  $\varkappa$ . Since the order of  $\varkappa$  is  $\mathcal{O}(\varepsilon \ln \varepsilon)$ , this means that the support of each ripple is of order  $\mathcal{O}(\varepsilon)$ , exactly as in the bulk of the genus 2 region. However their spacing is of the same order as  $\varkappa$ , so that they look like ranges of mountains/valleys separated by vast flatlands (see Fig. 2).*

While Theorem 7.1 contains the substantial qualitative information about the relevant features of the behavior, the reader may wish for a slightly more concrete formula. It is possible to write the correction term in full detail and –for example– use some computer algebra software to achieve the plot in Fig. 2. We restate the theorem with some more explicit expressions.

**Theorem 7.2** *The solution of the NLS (1-1) defined by the scattering data  $f_0(z)$  behaves as*

$$q(x, t) = -b(x, t)e^{\frac{4ig(\infty)}{\varepsilon} + 2iK\psi} \left( 1 - \frac{2e^{-2iK\psi}}{b(x, t)\Delta_K} \left\{ |u_K|^2 \chi_K + u_K \mathbf{B}_{K,1}^2 + \bar{u}_K \overline{\mathbf{B}_{K,2}}^2 \right\} + \mathcal{O}(\sqrt{\varepsilon}) \right), \quad (7-4)$$

where

$$\mathbf{B}_{K,1} = \frac{\tau^K e^{iK\psi}}{\sqrt{1 + \lambda_0^2}}, \quad \mathbf{B}_{K,2} = \frac{\tau^K e^{-iK\psi} \lambda_0}{\sqrt{1 + \lambda_0^2}}, \quad (7-5)$$

$$\tau := |\tau| e^{i\phi} = \frac{\lambda_0^2(1 + |\lambda_0|^2)}{ib\Im(\lambda_0)(1 + \lambda_0^2)^2}, \quad (7-6)$$

$$u_K = -i \frac{K!}{\sqrt{\pi} 2^{K+1} C^{K+\frac{1}{2}}} \exp\left(i \frac{\vartheta}{\varepsilon}\right) \varepsilon^{\frac{1}{2} - \varkappa + K}, \quad (7-7)$$

$$\Delta_K = 1 + 4\Re\left(\frac{u_K \lambda_0 \tau^{2K}}{b(1 + \lambda_0^2)^2}\right) + \frac{4|u_K|^2 |\lambda_0|^4 |\tau|^{4K}}{b^2 |1 + \lambda_0^2|^4} \left(1 + \frac{|1 + \lambda_0^2|^2}{4\Im(\lambda_0)^2}\right), \quad (7-8)$$

$$\chi_K = \frac{e^{2iK\psi}}{b|1 + \lambda_0^2|^2} \left( \frac{2\bar{\lambda}_0^2 \tau^{2K}}{1 + \bar{\lambda}_0^2} + \frac{2|\lambda_0|^4 \tau^{2K}}{1 + \lambda_0^2} - \frac{ib\bar{\lambda}_0 |1 + \lambda_0^2|}{\Im\eta} \right). \quad (7-9)$$

The proof can be found in App. B.

## 7.1 Discussion

**Universality.** The correction depends only on  $\Psi_K$ ,  $\eta$  and the constant  $C$  that appears in  $u_K$ ; in particular,  $\Psi_K$  depends only on the branch-point  $\alpha$  and the integer  $K$  which counts the number of first ripples. Thus, universality here is taken to mean that no details of  $f_0(z)$ , that defines a particular solution to the NLS (1-1), enter into the formula.

Since the gradient of  $\Im S(x, t)$  does not vanish, the natural scale of modulation of  $q(x, t)$  in the direction transversal to the breaking curve is  $\varepsilon \ln \frac{1}{\varepsilon} \|\nabla \Im S\|$  and  $\varkappa$  serves as exploration parameter in this scale.

Consider the case  $\varkappa = K + \frac{1}{2} \in \frac{1}{2} + \mathbb{N}$ , i.e. on the “crest” of the  $K$ -th ripple; for this value of the parameter the correction to  $q(x, t)$  is of order  $\mathcal{O}(1)$ . In this case  $u_K$  is a fast oscillatory function: if we fix a small  $\varepsilon$  then the correction term to the amplitude has a slow modulation in  $(x, t)$  due to the dependence of  $C, \eta, \alpha$  etc. and a fast change along the direction parallel to the breaking curve due to the fast oscillation of  $u_K$ .



**Large time limit of the profile.** According to [15], [13], together with (4-16), we have the following asymptotics along the right breaking curve as  $t \rightarrow \infty$ :

$$\alpha = a + ib \sim \mu_+ + ie^{\frac{2\pi(\mu_+ - \mu_-)t}{w'(\mu_+)}} , \quad (7-10)$$

$$\eta \sim \mu_- + i\frac{w'(\mu_-)}{4t} , \quad (7-11)$$

$$\lambda_0 = \lambda(\eta) \sim \frac{-2(\mu_+ - \mu_-)}{b} + i\frac{w'(\mu_-)}{2tb} , \quad (7-12)$$

$$\tau \sim -2i\frac{e^{i\psi}t}{w'(\mu_-)} , \quad w(x) := \Im(f_0(x)_+) \text{sign}(x - \mu_+) \quad (7-13)$$

Then

$$\bar{\mathbf{B}}_{K1}^2 \sim e^{2iK\psi} i^{2K} \left( \frac{2t}{w'(\mu_-)} \right)^{2K} , \quad (7-14)$$

whereas  $\det[\mathbf{B}'_K, \mathbf{B}_K]$  and  $\mathbf{B}_{K1}$  are decaying exponentially as  $t \rightarrow \infty$ . Thus

$$r_K = \Delta_K \sim 1 + |u_K|^2 \left( \frac{2t}{w'(\mu_-)} \right)^{4K+2} \quad (7-15)$$

and, according to (3-20) and (7-7),

$$u_K \sim e^{-\frac{i\pi}{4}} e^{-i\frac{\theta}{\varepsilon}} \frac{\varepsilon^{K+\frac{1}{2}-\varkappa} i^K K!}{\sqrt{\pi} 2^{3K+2} t^{K+\frac{1}{2}}} . \quad (7-16)$$

In order to gain some insight on the profile and relative height of the ripples for large time we take  $\varkappa$  a half-integer (similar consideration hold in a double-scaling approach as long as the expression  $\varepsilon^{K+\frac{1}{2}-\varkappa} t^{K+\frac{1}{2}}$  approaches infinity). Then we obtain ( $\varkappa = K + \frac{1}{2}$ )

$$(G_{K, K+\frac{1}{2}} - \bar{G}_{K, K+\frac{1}{2}})_{12} \sim -\sqrt{\pi} e^{-\frac{i\pi}{4}} i^K e^{2iK\psi} e^{i\frac{\theta}{\varepsilon}} \frac{2^K w'(\mu_-)^{2K+2}}{K! t^{K+\frac{3}{2}}} . \quad (7-17)$$

In the particular example, considered in the paper for numerical computations,  $w'(\mu_-) = \frac{1}{2}\pi$ . Notice that the ripples has algebraic decay as  $t \rightarrow \infty$  whereas the amplitude of the corresponding genus zero solution decays exponentially fast. Thus,  $F_{\text{ripple}}$  is the main contribution to the solution along the ripples for large  $t$ . Expression (7-17) also shows that the rate of decay for large  $t$  increases with the number  $K$ ; however, it may be offset by the  $K!$  growth of the coefficient in front of  $t^{-K-\frac{3}{2}}$ . That is consistent with  $\mathcal{O}(t^{-\frac{1}{2}})$  magnitude of the solution in the bulk of the genus two region, see [15].

**Remark 7.2** *Similar results are true for transition through any breaking curve separating regions of genus  $2N$  and  $2N + 2$ .*

**Remark 7.3** *This type of scales are to be expected from the similarities with random matrix models [16].*

## A Explicit formulæ

Using (4-13), the expression for  $\Phi_K(z)$  can be written as

$$\Phi_K(z) = \left(\frac{\bar{\lambda}_0}{\lambda_0}\right)^{K\sigma_3} \Psi_0(z) [(z - \eta)D(z)]^{K\sigma_3}. \quad (\text{A.1})$$

To express it in terms of  $\lambda$ , we calculate

$$(z - \eta)D(z) = \frac{b(\lambda - \bar{\lambda}_0)(1 + \lambda\lambda_0)^2}{2\lambda\lambda_0(1 + \lambda\bar{\lambda}_0)}, \quad (\text{A.2})$$

so that (A.1) becomes

$$\Phi_K = \frac{1}{\sqrt{\lambda^2 + 1}} \left(\frac{\bar{\lambda}_0}{\lambda_0}\right)^{K\sigma_3} \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \left(\frac{b(\lambda - \bar{\lambda}_0)(1 + \lambda\lambda_0)^2}{2\lambda\lambda_0(1 + \lambda\bar{\lambda}_0)}\right)^{K\sigma_3}. \quad (\text{A.3})$$

The quantities involved in (4-27) require the computation of  $\Phi'_K(z)$ , which is given by

$$\Phi'_K(z) = \left(\frac{\bar{\lambda}_0}{\lambda_0}\right)^{K\sigma_3} \left(\frac{z - \alpha}{z - \bar{\alpha}}\right)^{\frac{\sigma_2}{4}} \left[ \frac{ib\sigma_2}{2R^2(z)} + K\sigma_3 (\ln(z - \eta) + \ln D(z))' \right] ((z - \eta)D(z))^{K\sigma_3}. \quad (\text{A.4})$$

The expression in the square brackets is

$$\frac{2\lambda^2}{b(1 + \lambda^2)} \left[ \frac{i}{1 + \lambda^2} \sigma_2 + K\nu(\lambda)\sigma_3 \right], \quad (\text{A.5})$$

where

$$\nu(\lambda) = \left( \frac{2\lambda_0}{1 + \lambda\lambda_0} - \frac{\bar{\lambda}_0}{1 + \lambda\bar{\lambda}_0} + \frac{1}{\lambda - \bar{\lambda}_0} - \frac{1}{\lambda} \right). \quad (\text{A.6})$$

Thus,

$$\Phi'_K(z) = \frac{2\lambda^2}{b(1 + \lambda^2)^{\frac{3}{2}}} \left(\frac{\bar{\lambda}_0}{\lambda_0}\right)^{K\sigma_3} \begin{bmatrix} K\lambda\nu(\lambda) + \frac{1}{1 + \lambda^2} & K\nu(\lambda) + \frac{\lambda}{1 + \lambda^2} \\ K\nu(\lambda) - \frac{\lambda}{1 + \lambda^2} & -K\lambda\nu(\lambda) + \frac{1}{1 + \lambda^2} \end{bmatrix} \left(\frac{b(\lambda - \bar{\lambda}_0)(1 + \lambda\lambda_0)^2}{2\lambda\lambda_0(1 + \lambda\bar{\lambda}_0)}\right)^{K\sigma_3}. \quad (\text{A.7})$$

## B Proofs

### B.1 Proof of Prop. 5.1.

Equation (5-17) follows immediately from (5-16) by comparing the behavior of both sides of the equation as  $z \rightarrow \infty$ . To verify this we recall that the matrices  $G_{K,\varkappa}, \widehat{G}_{K,\varkappa}, F_{K+1,\varkappa}, \widehat{F}_{K+1,\varkappa}, C_K, \widehat{C}_K$  are uniquely characterized by the equations (5-7, 5-8) that can be written as

$$\left\{ \begin{array}{l} \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z - \eta} & 1 \end{bmatrix} = \mathcal{O}(1) \quad z \rightarrow \eta \\ \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z - \eta} - \frac{\widehat{G}_{K,\varkappa}}{z - \bar{\eta}} \right) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] \begin{bmatrix} 1 & \widehat{u}_K \\ 0 & z - \bar{\eta} \end{bmatrix} = \mathcal{O}(1), \quad z \rightarrow \bar{\eta}, \end{array} \right. \quad (\text{B.1})$$

$$\left\{ \begin{array}{l} \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z-\eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z-\bar{\eta}} \right) [\mathbf{A}_{K+1}(z), \mathbf{B}_{K+1}(z)] \begin{bmatrix} 1 & \ell_{K+1} \\ 0 & 1 \end{bmatrix} = \mathcal{O}(1) \quad z \rightarrow \eta \\ \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z-\eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z-\bar{\eta}} \right) [\widehat{\mathbf{A}}_{K+1}(z), \widehat{\mathbf{B}}_{K+1}(z)] \begin{bmatrix} 1 & 0 \\ \widehat{\ell}_{K+1} & 1 \end{bmatrix} = \mathcal{O}(1), \quad z \rightarrow \bar{\eta}, \end{array} \right. \quad (\text{B.2})$$

$$\ell_K := \frac{2i\pi e^{-i\frac{\theta}{\varepsilon}} \varepsilon^{\varkappa-K+\frac{1}{2}}}{\mu_{K-1} C^{-K+\frac{1}{2}}}, \quad \widehat{\ell}_K := \frac{2i\pi e^{i\frac{\theta}{\varepsilon}} \varepsilon^{\varkappa-K+\frac{1}{2}}}{\mu_{K-1} \widehat{C}^{-K+\frac{1}{2}}} \quad (\text{B.3})$$

$$\left\{ \begin{array}{l} \left( \mathbf{1} - \frac{C_K}{z-\eta} - \frac{\widehat{C}_K}{z-\bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] = [\mathbf{A}_{K+1}(z), \mathbf{B}_{K+1}(z)] (z-\eta)^{-\sigma_3} \\ \left( \mathbf{1} - \frac{C_K}{z-\eta} - \frac{\widehat{C}_K}{z-\bar{\eta}} \right) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] = [\widehat{\mathbf{A}}_{K+1}(z), \widehat{\mathbf{B}}_{K+1}(z)] (z-\bar{\eta})^{\sigma_3}. \end{array} \right. \quad (\text{B.4})$$

We note that from (5-13,B.3) it follows

$$u_K \ell_{K+1} = 1 = \widehat{u}_K \widehat{\ell}_{K+1}. \quad (\text{B.5})$$

So the check consists in verifying that the matrix

$$M(z) := \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z-\eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z-\bar{\eta}} \right) \left( \mathbf{1} - \frac{C_K}{z-\eta} - \frac{\widehat{C}_K}{z-\bar{\eta}} \right) \quad (\text{B.6})$$

solves

$$\left\{ \begin{array}{l} M(z) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \mathcal{O}(1) \\ M(z) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] \begin{bmatrix} 1 & \widehat{u}_K \\ 0 & 1 \end{bmatrix} = \mathcal{O}(1). \end{array} \right. \quad (\text{B.7})$$

Indeed we have

$$\begin{aligned} & M(z) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \\ & = \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z-\eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z-\bar{\eta}} \right) \left( \mathbf{1} - \frac{C_K}{z-\eta} - \frac{\widehat{C}_K}{z-\bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \\ & [\text{From (B.4)}] = \left( \mathbf{1} - \frac{F_{K+1,\varkappa}}{z-\eta} - \frac{\widehat{F}_{K+1,\varkappa}}{z-\bar{\eta}} \right) [\mathbf{A}_{K+1}(z), \mathbf{B}_{K+1}(z)] (z-\eta)^{-\sigma_3} \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \\ & [\text{From (B.2)}] = \mathcal{O}(1) \begin{bmatrix} 1 & -\frac{\ell_{K+1}}{z-\eta} \\ 0 & 1 \end{bmatrix} (z-\eta)^{-\sigma_3} \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \\ & = \mathcal{O}(1) \begin{bmatrix} \frac{1-\ell_{K+1}u_K}{z-\eta} & -\ell_{K+1} \\ u_K & (z-\eta). \end{bmatrix} \quad (\text{B.8}) \end{aligned}$$

This last expression is analytic because  $\ell_{K+1}u_K = 1$ . This proves the identity (5-16). One should repeat the verification near  $\bar{\eta}$  but it is completely parallel. This proves the continuity. **Q.E.D.**

## B.2 Proof of Thm. 5.1.

The existence of the solution will be implied by the formula and by noticing that all denominators are strictly positive. The symmetries (5-12) are obtained from the symmetries (4-23), (4-24).

The conditions that determine uniquely  $G_{K,\varkappa}, \widehat{G}_{K,\varkappa}, F_{K,\varkappa}, \widehat{F}_{K,\varkappa}$  are to be read off (5-7,5-8): we will show the details only for the pair  $G_{K,\varkappa}, \widehat{G}_{K,\varkappa}$ , the computation for  $F_{K,\varkappa}, \widehat{F}_{K,\varkappa}$  being completely similar.

**Case  $K < \varkappa < K + \frac{1}{2}$ .** The conditions (5-7, 5-8) translate into

$$\left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] (z-\eta)^{-K\sigma_3} \begin{bmatrix} 1 & 0 \\ \frac{\mu_K}{2i\pi\zeta^{2K+1}e^{-i\frac{\vartheta}{\varepsilon}\varepsilon^\varkappa}} & 1 \end{bmatrix} (z-\eta)^{K\sigma_3} = \mathcal{O}(1), \quad (\text{B.9})$$

$$\left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] (z-\bar{\eta})^{K\sigma_3} \begin{bmatrix} 1 & \frac{\mu_K}{2i\pi\widehat{\zeta}^{2K+1}e^{i\frac{\vartheta}{\varepsilon}\varepsilon^\varkappa}} \\ 0 & 1 \end{bmatrix} (z-\bar{\eta})^{-K\sigma_3} = \mathcal{O}(1) \quad (\text{B.10})$$

This can be rewritten

$$\left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \left( \frac{z-\eta}{\zeta\sqrt{\varepsilon}} \right)^{2K} \frac{\mu_K}{2i\pi\zeta e^{-i\frac{\vartheta}{\varepsilon}\varepsilon^{\varkappa-K}}} & 1 \end{bmatrix} = \mathcal{O}(1), \quad (\text{B.11})$$

$$\left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] \begin{bmatrix} 1 & \left( \frac{z-\bar{\eta}}{\widehat{\zeta}\sqrt{\varepsilon}} \right)^{2K} \frac{\mu_K}{2i\pi\widehat{\zeta} e^{i\frac{\vartheta}{\varepsilon}\varepsilon^{\varkappa-K}}} \\ 0 & 1 \end{bmatrix} = \mathcal{O}(1). \quad (\text{B.12})$$

From (3-8) it follows that

$$\sqrt{\varepsilon}\zeta = \sqrt{C}(z-\eta)(1 + \mathcal{O}(z-\eta)), \quad \sqrt{\varepsilon}\widehat{\zeta} = \sqrt{\widehat{C}}(z-\bar{\eta})(1 + \mathcal{O}(z-\bar{\eta})). \quad (\text{B.13})$$

In particular, we have

$$\lim_{z \rightarrow \eta} \frac{z-\eta}{\zeta\sqrt{\varepsilon}} = \frac{1}{\sqrt{C}}, \quad \lim_{z \rightarrow \bar{\eta}} \frac{z-\bar{\eta}}{\widehat{\zeta}\sqrt{\varepsilon}} = \frac{1}{\sqrt{\widehat{C}}} \quad (\text{B.14})$$

and that one can replace the above condition by the following

$$\begin{aligned} & \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\mathbf{A}_K(z), \mathbf{B}_K(z)] \begin{bmatrix} 1 & 0 \\ \frac{u_K}{z-\eta} & 1 \end{bmatrix} = \mathcal{O}(1), \\ & \left( \mathbf{1} - \frac{G_{K,\varkappa}}{z-\eta} - \frac{\widehat{G}_{K,\varkappa}}{z-\bar{\eta}} \right) [\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] \begin{bmatrix} 1 & \frac{\widehat{u}_K}{z-\bar{\eta}} \\ 0 & 1 \end{bmatrix} = \mathcal{O}(1), \end{aligned} \quad (\text{B.15})$$

$$u_K := \frac{\mu_K C^{-K-\frac{1}{2}}}{2i\pi e^{-i\frac{\vartheta}{\varepsilon}\varepsilon^{\varkappa-K-\frac{1}{2}}}}, \quad \widehat{u}_K := \frac{\mu_K \widehat{C}^{-K-\frac{1}{2}}}{2i\pi e^{i\frac{\vartheta}{\varepsilon}\varepsilon^{\varkappa-K-\frac{1}{2}}}}, \quad \mu_K := \frac{K!\sqrt{\pi}}{2^K}. \quad (\text{B.16})$$

We remark that

$$\widehat{u}_K = -\bar{u}_K. \quad (\text{B.17})$$

This implies the following system of equations

$$\begin{cases} G_{K,\varkappa} \mathbf{B}_K = 0 \\ \widehat{G}_{K,\varkappa} \widehat{\mathbf{A}}_K = 0 \\ -G_{K,\varkappa} \mathbf{A}_K + u_K \left( \mathbf{B}_K - G_K \mathbf{B}'_K - \frac{\widehat{G}_{K,\varkappa} \mathbf{B}_K}{\eta - \bar{\eta}} \right) = 0 \\ -\widehat{G}_{K,\varkappa} \widehat{\mathbf{B}}_K + \widehat{u}_K \left( \widehat{\mathbf{A}}_K - \frac{G_{K,\varkappa} \widehat{\mathbf{A}}_K}{\bar{\eta} - \eta} - \widehat{G}_{K,\varkappa} \widehat{\mathbf{A}}'_K \right) = 0, \end{cases} \quad (\text{B.18})$$

where –for brevity–  $\mathbf{A}_K := \mathbf{A}_K(\eta)$ ,  $\mathbf{A}'_K := \mathbf{A}'_K(\eta)$  etc. The solution is

$$G_{K,\varkappa}(\bullet) = \frac{u_K \det[\bullet, \mathbf{B}_K] (\eta - \bar{\eta})^2}{(1 + u_K \det[\mathbf{B}'_K, \mathbf{B}_K]) \left( 1 - \widehat{u}_K \det[\widehat{\mathbf{A}}'_K, \widehat{\mathbf{A}}_K] \right) (\eta - \bar{\eta})^2 + u_K \widehat{u}_K \det[\widehat{\mathbf{A}}_K, \mathbf{B}_K]^2} \times \quad (\text{B.19})$$

$$\times \left( \left( 1 - \widehat{u}_K \det[\widehat{\mathbf{A}}'_K, \widehat{\mathbf{A}}_K] \right) \mathbf{B}_K + \frac{\widehat{u}_K \det[\mathbf{B}_K, \widehat{\mathbf{A}}_K]}{(\eta - \bar{\eta})} \widehat{\mathbf{A}}_K \right), \quad (\text{B.20})$$

$$\widehat{G}_{K,\varkappa}(\bullet) = \frac{-\widehat{u}_K \det[\bullet, \widehat{\mathbf{A}}_K] (\bar{\eta} - \eta)^2}{(1 + u_K \det[\mathbf{B}'_K, \mathbf{B}_K]) \left( 1 - \widehat{u}_K \det[\widehat{\mathbf{A}}'_K, \widehat{\mathbf{A}}_K] \right) (\bar{\eta} - \eta)^2 + \widehat{u}_K u_K \det[\widehat{\mathbf{A}}_K, \mathbf{B}_K]^2} \times \quad (\text{B.21})$$

$$\times \left( (1 + u_K \det[\mathbf{B}'_K, \mathbf{B}_K]) \widehat{\mathbf{A}}_K - \frac{u_K \det[\widehat{\mathbf{A}}_K, \mathbf{B}_K]}{(\bar{\eta} - \eta)} \mathbf{B}_K \right), \quad (\text{B.22})$$

where we have used that

$$\det[\mathbf{A}_K(z), \mathbf{B}_K(z)] \equiv \det[\widehat{\mathbf{A}}_K(z), \widehat{\mathbf{B}}_K(z)] \equiv 1. \quad (\text{B.23})$$

Using the reality relations (4-27) and (B.17) we can rewrite them in the slightly simpler form which is in the statement of the theorem.

**Case**  $K - \frac{1}{2} < \varkappa < K$ . The derivation is completely analog. In particular we will see below that the formulæ for  $F_{K,\varkappa}$ ,  $\widehat{F}_{K,\varkappa}$  are not necessary for our purposes if we have those for  $G_{K,\varkappa}$ ,  $\widehat{G}_{K,\varkappa}$ . **Q.E.D.**

### B.3 Proof of Thm. 7.1.

According to (2-7), the solution to NLS is recovered from the solution  $\Gamma(z)$  of the *original* RHP as

$$q(x, t, \varepsilon) = -2 \lim_{z \rightarrow \infty} z \Gamma_{12}(z). \quad (\text{B.24})$$

Away from the cuts the approximation is such that

$$e^{-\frac{2i}{\varepsilon} g(\infty) \sigma_3} \Gamma(z) e^{\frac{2i}{\varepsilon} g(z) \sigma_3} = \mathcal{E}(z) \Psi_{\varkappa}(z), \quad (\text{B.25})$$

where  $\mathcal{E}(z)$  is the error term matrix which is (uniformly)  $\mathcal{E}(z) = \mathbf{1} + \mathcal{O}(\sqrt{\varepsilon})$  and, according to (2-12),

$$g(z; x, t) = g(\infty; x, t) + \mathcal{O}(z^{-1}), \quad \phi \in \mathbb{R}. \quad (\text{B.26})$$

Thus we have

$$q(x, t, \varepsilon) = -2e^{\frac{4i}{\varepsilon}g(\infty; x, t)} \lim_{z \rightarrow \infty} z \Psi_{\varkappa, 12}(z). \quad (\text{B.27})$$

According to Thm. 5.1, Cor. 5.1 and Proposition 4.1,

$$\Psi_{\varkappa}(z) = T_{K, \varkappa}(z) \left( \frac{\bar{\lambda}_0}{\lambda(z)} \right)^{K\sigma_3} \Psi_0(z) (D(z))^{K\sigma_3} \quad K = \max(0, \lfloor \varkappa \rfloor). \quad (\text{B.28})$$

Then

$$\begin{aligned} -2 \lim_{z \rightarrow \infty} z [\Psi_{\varkappa}]_{12} &= -2 \left( \frac{\bar{\lambda}_0}{\lambda_0} \right)^{2K} \lim_{z \rightarrow \infty} z [\Psi_0]_{12} + 2(G_{K, \varkappa} + \widehat{G}_{K, \varkappa})_{12} = \\ &= -b(x, t) e^{2iK\psi} + 2(G_{K, \varkappa} - \widehat{G}_{K, \varkappa})_{12} \\ e^{i\psi} &:= \frac{\bar{\lambda}_0}{\lambda_0}. \end{aligned} \quad (\text{B.29})$$

where  $\alpha(x, t) = a(x, t) + ib(x, t)$ . This is the leading order in the approximation: the correction to the amplitude  $b(x, t)$  of the solution  $q(x, t, \varepsilon)$  is purely a phase-shift in units of  $\psi$ , where  $e^{i\psi} = \frac{\bar{\lambda}_0}{\lambda_0}$ ; the integer  $K$  counts the number of ‘‘ripples’’ as we explore the genus two region. So, finally,

$$\begin{aligned} q(x, t, \varepsilon) &= -2e^{\frac{4i}{\varepsilon}g(\infty)} \lim_{z \rightarrow \infty} z (\mathcal{E}(z) \Psi_{\varkappa}(z))_{12} = \\ &= \left( -b(x, t) e^{2iK\psi} + 2(G_{K, \varkappa} + \widehat{G}_{K, \varkappa})_{12} + \mathcal{O}(\sqrt{\varepsilon}) \right) e^{\frac{2i}{\varepsilon}g(\infty)} \quad K \leq \varkappa < K+1 \end{aligned} \quad (\text{B.30})$$

This proves formula (7-1). To verify the assertion regarding the different behaviors (7-3) we recall here equation (5-17)

$$G_{K, \varkappa} + \widehat{G}_{K, \varkappa} = F_{K+1, \varkappa} + \widehat{F}_{K+1, \varkappa} + C_K + \widehat{C}_K. \quad (\text{B.31})$$

We know that the matrices  $F_{K+1, \varkappa}$ ,  $\widehat{F}_{K+1, \varkappa}$  both are of order  $\mathcal{O}(\varepsilon^{\varkappa-K-\frac{1}{2}})$  (by inspection of (5-11) and (5-13)). On the other hand

$$C_K + \widehat{C}_K = \lim_{z \rightarrow \infty} z (\Psi_K \Psi_{K+1}^{-1}) = \frac{-b(x, t)}{2} \left[ e^{i2K\psi} - e^{-2(K+1)\psi} \right]. \quad (\text{B.32})$$

And this concludes the proof. **Q.E.D.**

## B.4 Proof of Thm. 7.2.

Looking at the formulæ (5-10) and recalling formulæ (4-27) and  $\widehat{u}_K = -\overline{u_K}$  we find

$$(G_{K, \varkappa} + \widehat{G}_{K, \varkappa})_{12} = \frac{-1}{\Delta_K} \left\{ |u_K|^2 \chi_K + u_K \mathbf{B}_{K,1}^2 + \overline{u_K} \overline{\mathbf{B}_{K,2}}^2 \right\}, \quad (\text{B.33})$$

$$\chi_K := \left( \det[\mathbf{B}'_K, \mathbf{B}_K] \mathbf{B}_{K,1}^2 + \det[\mathbf{B}'_K, \mathbf{B}_K] \overline{\mathbf{B}_{K,2}}^2 - 2 \frac{\|\mathbf{B}_K\|^2 \overline{\mathbf{B}_{K,2}} \mathbf{B}_{K,1}}{\bar{\eta} - \eta} \right), \quad (\text{B.34})$$

$$\Delta_K := \left| 1 + u_K \det[\mathbf{B}'_K, \mathbf{B}_K] \right|^2 + \frac{|u_K|^2 \|\mathbf{B}_K\|^4}{4(\Im(\eta))^2} > 0. \quad (\text{B.35})$$

At this point one needs to use (A.3) (which defines the column vectors  $\mathbf{A}_K, \mathbf{B}_K$ ) and (A.7) (which gives  $\mathbf{A}'_K, \mathbf{B}'_K$ ), which produce (7-5). Plugging into the various object we obtain

$$\det[\mathbf{B}'_K, \mathbf{B}_K] = \frac{2\lambda_0^2}{b(1 + \lambda_0^2)^2} \tau^{2K} \quad \|\mathbf{B}_K\|^2 = \frac{1 + |\lambda_0|^2}{|1 + \lambda_0^2|} |\tau|^{2K} \quad (\text{B.36})$$

The rest is simply an algebraic manipulation which is as straightforward as uninteresting. **Q.E.D.**

## References

- [1] V. E. Zaharov and A. B. Šabat. Integration of the nonlinear equations of mathematical physics by the method of the inverse scattering problem. II. *Funktsional. Anal. i Prilozhen.*, 13(3):13–22, 1979.
- [2] M. Gregory Forest and Jong Eao Lee. Geometry and modulation theory for the periodic nonlinear Schrödinger equation. In *Oscillation theory, computation, and methods of compensated compactness (Minneapolis, Minn., 1985)*, volume 2 of *IMA Vol. Math. Appl.*, pages 35–69. Springer, New York, 1986.
- [3] Peter D. Miller and Spyridon Kamvissis. On the semiclassical limit of the focusing nonlinear Schrödinger equation. *Phys. Lett. A*, 247(1-2):75–86, 1998.
- [4] E. R. Tracy and H. H. Chen. Nonlinear self-modulation: an exactly solvable model. *Phys. Rev. A* (3), 37(3):815–839, 1988.
- [5] David Cai, David W. McLaughlin, and Kenneth T. R. McLaughlin. The nonlinear Schrödinger equation as both a PDE and a dynamical system. In *Handbook of dynamical systems, Vol. 2*, pages 599–675. North-Holland, Amsterdam, 2002.
- [6] Spyridon Kamvissis, Kenneth D. T.-R. McLaughlin, and Peter D. Miller. *Semiclassical soliton ensembles for the focusing nonlinear Schrödinger equation*, volume 154 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2003.
- [7] Alexander Tovbis, Stephanos Venakides, and Xin Zhou. On semiclassical (zero dispersion limit) solutions of the focusing nonlinear Schrödinger equation. *Comm. Pure Appl. Math.*, 57(7):877–985, 2004.
- [8] B. Dubrovin, T. Grava, and C. Klein. On universality of critical behavior in the focusing nonlinear Schrödinger equation, elliptic umbilic catastrophe and the Tritonquée solution to the Painlevé-I equation. *J. Nonlinear Sci.*, 19(1):57–94, 2009.
- [9] M. Bertola and S. Y. Lee. First colonization of a spectral outpost in random matrix theory. *Constr. Approx.*, 2008 (in press).

- [10] M. Bertola and S. Y. Lee. First colonization of a hard-edge in random matrix theory. *Constr. Approx.*, In press, 2009.
- [11] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems. *Bull. Amer. Math. Soc. (N.S.)*, 26(1):119–123, 1992.
- [12] Alexander Tovbis and Stephanos Venakides. In preparation. Technical report.
- [13] Alexander Tovbis, Stephanos Venakides, and Xin Zhou. Semiclassical focusing nonlinear Schrödinger equation I: inverse scattering map and its evolution for radiative initial data. *Int. Math. Res. Not. IMRN*, (22):Art. ID rnm094, 54, 2007.
- [14] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.*, 52(11):1335–1425, 1999.
- [15] Alexander Tovbis, Stephanos Venakides, and Xin Zhou. On the long-time limit of semiclassical (zero dispersion limit) solutions of the focusing nonlinear Schrödinger equation: pure radiation case. *Comm. Pure Appl. Math.*, 59(10):1379–1432, 2006.
- [16] B Eynard. Universal distribution of random matrix eigenvalues near the ‘birth of a cut’ transition. *Journal of Statistical Mechanics: Theory and Experiment*, Jan 2006.