SEMICLASSICAL WKB PROBLEM FOR THE NON-SELF-ADJOINT DIRAC OPERATOR WITH ANALYTIC POTENTIAL

SETSURO FUJIIÉ AND SPYRIDON KAMVISSIS

ABSTRACT. In this paper we examine the semiclassical behaviour of the scattering data of a non-self-adjoint Dirac operator with analytic potential decaying at infinity. In particular, employing the exact WKB method, we provide the complete rigorous uniform semiclassical analysis of the reflection coefficient and the Bohr-Sommerfeld condition for the location of the eigenvalues. Our analysis has some interesting consequences concerning the focusing cubic NLS equation, in view of the well-known fact discovered by Zakharov and Shabat that the spectral analysis of the Dirac operator is the basis of the solution of the NLS equation via inverse scattering theory.

1. Introduction: Motivation

In the last twenty years or so the analysis of the semiclassical behaviour of the focusing NLS equation has been rigorously achieved (and also numerically supported and clarified) for a certain class of real analytic decaying initial data ([17], [18], [14], [15], [22], [3]).

The problem is as follows: consider the semiclassical limit ($\epsilon \to 0$) of the solution to the initial value problem of the one-dimensional nonlinear Schrödinger equation:

$$\begin{cases} i\epsilon\partial_t\psi + \frac{\epsilon^2}{2}\partial_x^2\psi + |\psi|^2\psi = 0, \\ \psi(x,0) = A(x), \end{cases} \tag{1.1}$$

We assume here that $A(x)$ is a *real analytic* integrable function, and moreover that it is a positive "bell-shaped" function; in other words

$$A(x) > 0, \quad A(-x) = A(x), \tag{1.2}$$

and it has one single non-degenerate maximum at 0, say A_0,

$$A(0) = A_0 > 0; \quad xA'(x) < 0, \ if \ x \neq 0; \quad A''(0) < 0, \tag{1.3}$$

Suppose now that we replace the initial data by the so-called "soliton ensembles" data which are defined by replacing the scattering data for $\psi(x,0) = A(x)$ with their *formal* WKB-approximation: we set the reflection coefficient of the associated Dirac operator (see section 2) to be identically zero and replace the actual eigenvalues by their Bohr-Sommerfeld approximation (see section 5). In other words we replace the initial data by a *new* set of data which is now depending on ϵ. Suppose that we solve the focusing NLS equation under this new set of initial data. Then we have the following.

Let x_0, t_0 be any given point ($x_0 \in \mathbb{R}, t_0 > 0$). The solution $\psi(x,t)$ is asymptotically ($\epsilon \to 0$) described (locally) as a slowly modulated $G+1$ phase wavetrain. Setting $x = x_0 + \epsilon\hat{x}$ and $t = t_0 + \epsilon\hat{t}$, so that x_0, t_0 are "slow" variables while $\hat{x}, \hat{t}$ are "fast" variables, there exist parameters $a, U = (U_0, U_1, ..., U_G)^T$, $k = (k_0, k_1, ..., k_G)^T$, $w = (w_0, w_1, ..., w_G)^T$, $Y = (Y_0, Y_1, ..., Y_G)^T$, $Z = (Z_0, Z_1, ..., Z_G)^T$ depending on the slow variables x_0 and t_0 (but not $\hat{x}, \hat{t}$) such that generically $\psi(x,t) = \psi(x = x_0 + \epsilon\hat{x}, t = t_0 + \epsilon\hat{t})$ has the following leading order asymptotics as $\epsilon \to 0$:

$$\psi(x,t) \sim a(x_0,t_0) e^{\frac{iU_0(x_0,t_0)}{\epsilon}} e^{i(k_0(x_0,t_0)\hat{x} - w_0(x_0,t_0)\hat{t})} \frac{\Theta(Y(x_0,t_0) + i\frac{U(x_0,t_0)}{\epsilon} + i(k(x_0,t_0)\hat{x} - w(x_0,t_0)\hat{t}))}{\Theta(Z(x_0,t_0) + i\frac{U(x_0,t_0)}{\epsilon} + i(k(x_0,t_0)\hat{x} - w(x_0,t_0)\hat{t}))} \tag{1.4}$$

All parameters can be defined in terms of an underlying Riemann surface X which depends solely on x_0, t_0. The moduli of X vary slowly with x, t, i.e. they depend on x_0, t_0 but not on $\epsilon, \hat{x}, \hat{t}$. Θ is the G-dimensional Jacobi theta function associated with X. The genus of X can vary with x_0, t_0. In fact, the x, t-plane is divided into open regions in each of which G is constant. On the boundaries of such regions (sometimes called "caustics"; they are unions of analytic arcs), some degeneracies appear in the mathematical analysis (we may have "pinching" of the surfaces X for example) and interesting physical phenomena can appear (like the famous Peregrine rogue wave [3]). The above formulae give asymptotics which are uniform in compact (x,t)-sets not containing points on the caustics. For the exact formulae for the parameters as well as the definition of the theta functions we refer to [17] or [18]. Near the caustics the correct interpretation of (1.4) requires some more work. For an analysis of the somewhat more delicate behaviour (especially for higher order terms in ϵ) near the first caustic see [3].

The above result is interesting but somewhat unsatisfactory. The reason, of course, is that the initial data is substituted by the soliton ensembles data. A rigorous justification of this substitution requires rigorous semiclassical asympotics for the spectral data of the Dirac operator that is associated to the focusing NLS equation (see the next section). Our main aim in this paper is to show how the powerful "exact WKB method" can be used to provide the necessary rigorous asymptotic results.

The question of the semiclassical approximation of the scattering data has a deeper significance in view of the instability of the problem which appers in many levels. In fact even in the non-semiclassical regime, the *focusing* NLS is the main model for the so-called "modulational instability" ([1], [2]), although for positive fixed ϵ the initial value problem is well-posed.

Semiclassically the instabilities become more pronounced. One way to see this is related to the underlying ellipticity of the formal semiclassical limit. To be more specific, consider the well-known Madelung transformation [23].

$$\rho = |\psi|^2, \ \mu = \epsilon \mathrm{Im}\,(\bar{\psi}\psi_x). \tag{1.5}$$

Then the initial value problem becomes

$$\rho_t + \mu_x = 0, \ \mu_t + (\frac{\mu^2}{\rho} + \frac{\rho^2}{2})_x = \frac{\epsilon^2}{4}\partial_x(\rho(\log\rho)_{xx}), \tag{1.6}$$

with initial data $\rho(x,0) = (\psi(x,0))^2$ and $\mu(x,0) = 0$.

The formal limit as $\epsilon \to 0$ is

$$\rho_t + \mu_x = 0, \ \mu_t + (\frac{\mu^2}{\rho} + \frac{\rho^2}{2})_x = 0, \tag{1.7}$$

with initial data $\rho(x,0) = (\psi(x,0))^2$ and $\mu(x,0) = 0$.

This is an initial value problem for an elliptic system of equations and so one expects that small perturbations of the initial data (independent of ϵ) can lead to large changes in the solution, at any given time.

Another appearance of instabilities appears at the spectral analysis of the related non-self-adjoint Dirac operator (see the next section). Instability appears also at the related equilibrium measure problem (see section 6 and the appendix), the related Whitham equations (they are also elliptic) and even in the numerical studies of the problem.

As already stated, the semiclassical approximation of the scattering data results in small changes of the initial data *that depend on* ϵ. It is a priori unclear whether they can have a significant effect in the semiclassical asymptotics of the solution at a given time. Our aim is to prove that, at least for these particular initial data, they do not.

In simpler problems like the real KdV equation, or the defocusing nonlinear Schrödinger equation, one can make use of the underlying hyperbolicity of the formal limit to prove, a posteriori, that the formal semiclasssical WKB analysis of the scattering data is justified. In the focusing nonlinear Schrödinger equation, we need more delicate tools, provided by the exact WKB method. The exact WKB method was first developed for the Schrödinger operator, but here we apply it to the Dirac operator that is associated to the focusing NLS equation. The method goes back to works of Ecalle [6] and Voros [24] but here we argue along the lines of the papers of Gérard-Grigis [10] and Fujiié-Lasser-Nédélec [7]. Rather than relying on the usual formal WKB method which relies on asymptotic series that are in general divergent, we use a "resummation" of the series and in fact construct *exact* solutions in terms of *convergent* series, thus resolving a problem of "asymptotics beyond all orders".

We begin by addressing the issue of the reflection coefficient. The exact WKB method is employed to prove that it is exponentially small away from the point 0, in the spectral plane. Similarly, we give a rigorous justification of the Bohr-Sommerfeld asymptotic conditions for the locations of the eigenvalues. Our main assumptions on the initial data, i.e. the potential of the Dirac operator are two: analyticity near the real line and a mild decay estimate. Some extra technical assumptions are needed for the analysis of the scattering data near 0. These assumptions are often cumbersome to check. Still a wide enough open class of "potentials" $A(x)$ satisfies these assumptions, including positive bell-shaped rational functions in L^1 and also exponential functions like $(\cosh x)^{-1}$, e^{-x^2} (see Example 2.6, 2.7).

The plan of this paper is the following. In section 2 we state the exact assumptions on the potential and present the results on the rigorous WKB approximation. In section 3 the exact WKB method is presented. In section 4, it is applied to the reflection coefficient of the Dirac operator. In section 5, the eigenvalues are considered. In section 6, we present the application of the WKB results to the focusing NLS problem and we explain how the analysis of [17] needs to be modified in view of these results.

In the appendix, we first present the Riemann-Hilbert problem for the focusing NLS equation. We then give a rudimentary description of the change of variables needed to asymptotically deform the given Riemann-Hilbert problem into a "model" problem that can be explicitly solved. Finally, we present a discussion of the results of [14], which concern the possible obstacle of the non-analyticity of the spectral density of eigenvalues.

2. Assumptions and results

We study the semiclassical asymptotics of the reflection coefficient and the eigenvalue distribution of the Dirac operator

$$L := \begin{pmatrix} -\frac{\epsilon}{i}\frac{d}{dx} & -iA(x) \\ -iA(x) & \frac{\epsilon}{i}\frac{d}{dx} \end{pmatrix}.$$

Here, $\varepsilon > 0$ is the semiclassical small parameter and $A(x)$ is a function satisfying

(A1): $A(x)$ is a real positive smooth function on $\mathbb{R}$, and extends analytically to the complex domain

$$D_0 = D(\rho_0, \theta_0) := \{x \in \mathbb{C}; |\mathrm{Im}\, x| < \max(\rho_0, (\tan\theta_0)|\mathrm{Re}\, x|)\}$$

for some positive ρ_0 and θ_0. Moreover, there exists a positive τ such that, as $x \to \infty$ in D_0,

$$A(x) = \mathcal{O}(|x|^{-1-\tau}).$$

Under this condition, it is known that the spectrum of the non-self-adjoint operator L consists of the continuous spectrum $\mathbb{R}$ and a finite number of eigenvalues coming in complex conjugate pairs close to $i[-A_0, A_0]$, where $A_0 := \max_{x\in\mathbb{R}} A(x) > 0$ when ϵ is small.

For $\lambda \in \mathbb{R} \setminus \{0\}$, there exist a pair of solutions $\mathbf{f}_+^r(x,\varepsilon)$, $\mathbf{f}_-^r(x,\varepsilon)$ which behave, as $\mathrm{Re}\, x \to \infty$ in D, like

$$\mathbf{f}_+^r \sim \begin{pmatrix} 0 \\ e^{i\lambda x/\varepsilon} \end{pmatrix}, \quad \mathbf{f}_-^r \sim \begin{pmatrix} e^{-i\lambda x/\varepsilon} \\ 0 \end{pmatrix},$$

as well as a pair of solutions $\mathbf{f}_+^l(x,\varepsilon)$, $\mathbf{f}_-^l(x,\varepsilon)$ which behave, as $\mathrm{Re}x \to -\infty$ in D, like

$$\mathbf{f}_+^l \sim \begin{pmatrix} 0 \\ e^{i\lambda x/\varepsilon} \end{pmatrix}, \quad \mathbf{f}_-^l \sim \begin{pmatrix} e^{-i\lambda x/\varepsilon} \\ 0 \end{pmatrix}.$$

These solutions are called *Jost solutions*. Each of these pairs is uniquely determined and forms a basis of solutions. Let $T(\lambda,\varepsilon)$ be the 2×2 constant matrix depending on λ and ε expressing the change of basis of these two pairs:

$$(\mathbf{f}_+^l, \mathbf{f}_-^l) = (\mathbf{f}_+^r, \mathbf{f}_-^r)T. \tag{2.1}$$

Then T is of the form

$$T(\lambda,\varepsilon) = \begin{pmatrix} a(\lambda,\varepsilon) & b^*(\lambda,\varepsilon) \\ b(\lambda,\varepsilon) & a^*(\lambda,\varepsilon) \end{pmatrix} \tag{2.2}$$

where a^*, b^* denote the complex conjugates of a, b. The reflection coefficient $R(\lambda,\varepsilon)$ is by definition

$$R(\lambda,\varepsilon) = \frac{b(\lambda,\varepsilon)}{a(\lambda,\varepsilon)}. \tag{2.3}$$

It is easy to see that it can be expressed by wronskians of Jost solutions:

$$R(\lambda,\varepsilon) = \frac{\mathcal{W}(\mathbf{f}_+^r, \mathbf{f}_+^l)}{\mathcal{W}(\mathbf{f}_+^l, \mathbf{f}_-^r)}. \tag{2.4}$$

For the reflection coefficient, we first have the following result for $\lambda \in \mathbb{R}$ satisfying $|\lambda| \geq \delta$ with ϵ-independent positive δ.

Theorem 2.1. *Assume (A1). Then, for any $\delta > 0$, there exists $\sigma > 0$ independent of ε such that*

$$|R(\lambda,\varepsilon)| = \mathcal{O}(e^{-\sigma/\varepsilon}),$$

as $\varepsilon \to 0$ uniformly for $\lambda \in (-\infty, \delta] \cup [\delta, \infty)$.

For the eigenvalues, we assume moreover that $A(x)$ is a "bell-shaped" function:

- **(A2):** $A(x) = A(-x)$ and $A'(x) < 0$ for $x > 0$.
- **(A3):** $A''(0) < 0$.

Let $\lambda = i\mu$ with $0 < \mu < A_0 = A(0)$. The assumption (A2) implies that there exists a unique positive $x^*(\mu)$ such that there are exactly two real numbers $x^*(\mu)$ and $-x^*(\mu)$ which satisfy $A(x) = \mu$. Let us define an "action integral"

$$S(\mu) = \int_{-x^*(\mu)}^{x^*(\mu)} \sqrt{A(x)^2 - \mu^2}\,dx. \tag{2.5}$$

Then the so-called Bohr-Sommerfeld quantization rule is written as

$$-e^{2iS(\mu)/\varepsilon} = 1.$$

The meaning of this rule is that as μ runs over the roots of this equation then the values $\lambda = i\mu$ will be good estimates of the eigenvalues on $i[-A(0), A(0)]$ as $\epsilon \to 0$. We will study the accuracy of this rule. More precisely, we will show that there exists a function $m(\mu, \epsilon)$ asymptotic to -1 in the semiclassical limit $\epsilon \to 0$ such that the eigenvalues are exactly given by $\lambda = i\mu$ with the roots of $m(\mu, \varepsilon)e^{2iS(\mu)/\varepsilon} = 1$.

Theorem 2.2. *Assume (A1), (A2) and (A3). Then there exists a function $m(\mu, \varepsilon)$ with asymptotic behavior*

$$m(\mu, \varepsilon) = -1 + \mathcal{O}(\varepsilon)$$

as $\varepsilon \to 0$ uniformly in any closed interval $I \subset (0, A(0)]$ such that $\lambda = i\mu$ where $\mu \in I$ is an eigenvalue of L if and only if

$$m(\mu, \varepsilon)e^{2iS(\mu)/\varepsilon} = 1. \tag{2.6}$$

Remark 2.3. *Klaus and Shaw proved in* [21] *that all the eigenvalues are simple and purely imaginary under the bell-shaped conditions (A1), (A2). Recently Hirota and Wittsten refined Theorem 2.2 to show that the eigenvalues are still pure imaginary even if we only impose an "energy-local" bell-shaped condition for small enough ϵ (see* [11]*).*

Now let us focus on the asymptotic behavior of the functions $R(\lambda, \varepsilon)$ and $m(\mu, \varepsilon)$ when $\lambda > 0$ or $\mu > 0$ tends to 0 together with ε. In such a case, we need a more precise assumption on the asymptotic behavior of the potential $A(x)$ as $|x| \to \infty$ in D.

We define a function

$$z(x) = i\int_0^x \sqrt{A(t)^2 + \lambda^2}\,dt, \tag{2.7}$$

where we take the branch of the square root such that it is positive at $t = 0$. This function is well-defined and holomorphic at least near the origin $x = 0$. It is extended in D_0 except at the turning points, i.e. the zeros of $A(t)^2 + \lambda^2$, around which it is multi-valued.

We first consider the case where $\lambda > 0$ is small. In this case, there is no turning point on the real axis, and the image of the real axis by the map $x \mapsto z(x)$ is the imaginary axis. Let $F(a)$ be the cone-like set

$$F(a) = \{z \in \mathbb{C}; |\mathrm{Re}z| < a|\mathrm{Im}z|\}$$

for $a > 0$. We assume

- **(A4):** For any $\lambda > 0$ small, there exist positive constants $\rho(\lambda)$, $\theta(\lambda)$ and $a(\lambda)$ such that $D(\rho(\lambda), \theta(\lambda))$ contains no turning point and its image by the map $x \mapsto z(x)$ includes $F(a(\lambda))$.

Theorem 2.4. *Assume (A1), (A2) and (A4). Then there exists a positive constant c such that*

$$R(\lambda, \varepsilon) = \mathcal{O}\left(e^{-ca(\lambda)/\varepsilon}\right),$$

as $\varepsilon \to +0$ and $\lambda \to +0$ with $\frac{\varepsilon}{a(\lambda)} \to 0$.

Next we consider the case where $\lambda = i\mu$ and $\mu > 0$ is small. In this case, there are exactly two turning points $x^*(\mu)$ and $-x^*(\mu)$ on the real axis. By the map $x \mapsto z(x)$, the real interval $(-x^*(\mu), x^*(\mu))$ is sent to the imaginary interval $(-z(x^*(\mu)), z(x^*(\mu)))$, and the half line $(x^*(\mu), \infty)$ (resp. $(-\infty, -x^*(\mu))$ is sent to the half line $z(x^*(\mu)) + \mathbb{R}_+$ (resp. $-z(x^*(\mu)) + \mathbb{R}_+$) when the square root in (2.7) is continued from $(-x^*(\mu), x^*(\mu))$ to $(x^*(\mu), \infty)$ (resp. $(-\infty, -x^*(\mu))$ passing through the upper half plane around the turning point $x^*(\mu)$ (resp. $-x^*(\mu)$). Notice that, as $\mu \to 0$, one has $x^*(\mu) \to +\infty$ and

$$|z(x^*(\mu))| \to \int_0^{+\infty} A(x)dx =: z_\infty,$$

which is a positive finite number. Let $G(b)$ be the complex subdomain of $\mathbb{C}_z$ defined by

$$\begin{aligned} G(b) = \ & \{-b < \mathrm{Re} z < 0, |\mathrm{Im} z| < |z(x^*(\mu))| + b\} \\ & \cup \{\mathrm{Re} z \geq 0, |z(x^*(\mu))| < |\mathrm{Im} z| < |z(x^*(\mu))| + b\} \end{aligned}$$

for $b > 0$. We assume

(A5): For any $\mu > 0$ small, there exist positive constants $\rho(\mu)$, $\theta(\mu)$ and $b(\mu)$ such that $D(\rho(\mu), \theta(\mu)) \cap \{z \in \mathbb{C}; \mathrm{Im} z > 0\}$ contains no turning point and its image by the map $x \mapsto z(x)$ includes $G(b(\mu))$.

Theorem 2.5. *Assume (A1), (A2), (A3) and (A5). Then there exists a function $m(\mu, \varepsilon)$ with asymptotic behavior*

$$m(\mu, \varepsilon) = -1 + \mathcal{O}\left(\frac{\varepsilon}{b(\mu)}\right)$$

as $\varepsilon \to +0$ and $\mu \to +0$ with $\frac{\varepsilon}{b(\mu)} \to 0$, such that $\lambda = i\mu$ is an eigenvalue of L if and only if (2.6) *holds.*

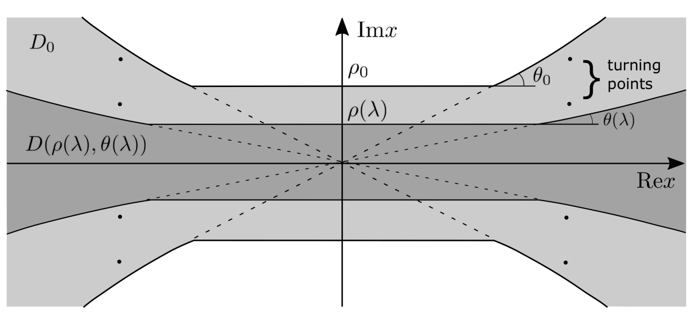

FIGURE 1. The domains D_0 and $D(\rho(\lambda), \theta(\lambda))$

Example 2.6. *Suppose $A(x)$ satifies (A1), (A2) and*

$$A(x) = Cx^{-d} + r(x) \text{ for } x > 1,$$

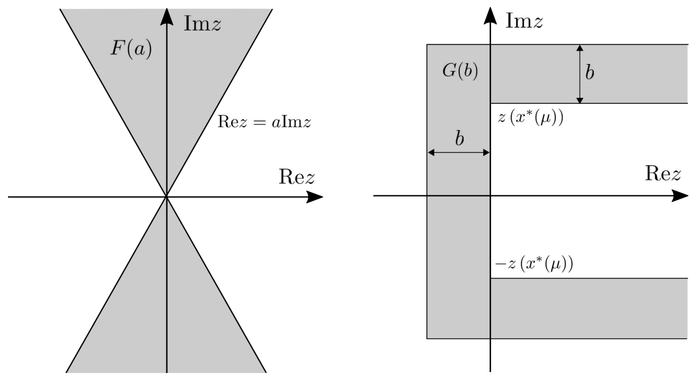

FIGURE 2. The domains $F(a)$ and $G(b)$

with $d > 1$, $C > 0$ *and* $r(x) = o(|x|^{-d-1})$, $r'(x) = o(|x|^{-d-2})$ *as* $\mathrm{Re}x \to \infty$ *in* D_0. *Then, one can take* $a(\lambda) = c\lambda$ *and* $b(\mu) = c\mu^{1+\frac{1}{2d}}$ *for some positive constant* c.

Proof. For simplicity, C is assumed to be 1 below. Let $\lambda > 0$ small. Since $r(x) = o(|x|^{-d-1})$, we see by Rouché's theorem that the turning points in this domain are $\pm e^{(k+1/2)\pi i/d}\lambda^{-1/d} + o(1)$ with integers k satisfying $|k+1/2|\pi < d\theta_0$ and the nearest turning points to the real axis are $\pm e^{\pm\pi i/(2d)}\lambda^{-1/d}+o(1)$. Hence the domain $D(\rho,\theta)$ contains no turning point for small enough λ-independent ρ and any θ smaller than $\pi/(2d)$. Its image $z(D(\rho,\theta))$ includes the domain $F(a(\lambda))$ with $a(\lambda) = c\lambda$ for some poisitive constant c. In fact, for $x \in D(\rho_1,\theta_1)$ with small enough ρ_1, θ_1,

$$\begin{aligned}|\mathrm{Re}\, z(x)| = \ & \lambda|\mathrm{Im}\, x| \int_0^1 \mathrm{Re}\sqrt{1+\lambda^{-2}A(\mathrm{Re}\, x + is\mathrm{Im}\, x)^2}ds \\ \geq \ & \tfrac{1}{2}\lambda|\mathrm{Im}\, x|,\end{aligned}$$

$$\begin{aligned}|\mathrm{Im}\, z(x)| = \ & |\mathrm{Re}\, x \int_0^1 \sqrt{A(s\mathrm{Re}\, x)^2+\lambda^2}dt \\ & -\lambda \mathrm{Im}\, x \int_0^1 \mathrm{Im}\sqrt{1+\lambda^{-2}A(\mathrm{Re}\, x + is\mathrm{Im}\, x)^2}ds| \\ & \leq 2A_0|\mathrm{Re}\, x|.\end{aligned}$$

Hence $F(c\lambda) \subset z(D(\rho_1,\theta_1))$ for $c = \frac{|\tan\theta_1|}{4A_0}$.

For $\lambda = i\mu$ with $\mu > 0$ small, the turning points in this domain are $\pm e^{k\pi i/d}\mu^{-1/d} + o(1)$ with integers k satisfying $|k|\pi < d\theta_0$. In particular $x^*(\mu) = \mu^{-1/d} + o(1)$, and the nearest turning points to the real axis (apart from the real ones $\pm x^*(\mu)$) are $\pm e^{\pm\pi i/d}\mu^{-1/d} + o(1)$. Hence $D(\rho(\mu),\theta(\mu)) \cap \{z \in \mathbb{C}; \mathrm{Im}z > 0\}$ is turning point free for μ-independent ρ and any θ smaller than π/d. Its image by the map $z(x)$ includes the domain $G(b(\mu))$ with $b(\mu) = c\mu^{1+\frac{1}{2d}}$ for some positive constant c. To see this, we observe that

$$\int_{x^*(\mu)}^{x} \sqrt{A(t)^2-\mu^2}dt = \int_0^{x-x^*(\mu)} \sqrt{A(x^*(\mu)+s)^2-\mu^2}ds$$

and that, since $A'(x) = -dx^{-d-1} + o(|x|^{-d-1})$ as $x \to \infty$, the Taylor expansion of $A(x^*(\mu)+s)^2 - \mu^2$ in s gives

$$\sqrt{A(x^*(\mu)+s)^2-\mu^2} \sim \sqrt{2d}\,\mu^{1+\frac{1}{2d}}(-s)^{1/2},$$

as $\mu^{2+\frac{1}{d}}s \to 0$. This means that, when x runs from a point ic_0 to the right along a line $\mathrm{Im}\,x = c_0$ for a small but μ-independent positive c_0, its image $z(x)$ goes from $z(ic_0)$ near $z=0$ with $\mathrm{Re}\,z < 0$ first to the upper direction and then changes the direction to the right around $z(x^*(\mu))$ keeping a distance of order $\mu^{1+\frac{1}{2d}}$ from $z(x^*(\mu))$, and finally goes to infinity above the horizontal line $\mathrm{Im}z = \mathrm{Im}z(x^*(\mu))$. □

Example 2.7. *Suppose $A(x)$ satifies (A1), (A2) and*

$$A(x) = Ce^{-x^\sigma} + r(x) \text{ for } x > 1,$$

with $\sigma > 0$, $C > 0$ and $r(x) = o(e^{-x^\sigma})$, $r'(x) = o(x^{\sigma-1}e^{-x^\sigma})$ as $\mathrm{Re}x \to \infty$ in D_0. Then, one can take $a(\lambda) = c\frac{\lambda}{\log\frac{1}{\lambda}}$ and $b(\mu) = c\mu(\log\frac{1}{\mu})^{-1+\frac{1}{\sigma}}$ for some positive constant c.

Proof. Here also C is assumed to be 1.

For $\lambda > 0$ small, i.e. $L = \log\frac{1}{\lambda}$ large, the turning points in D_0 are $\pm\left(L + (k+\frac{1}{2})\pi i\right)^{\frac{1}{\sigma}} + o(L^{\frac{1}{\sigma}-1})$ with some integers k (the distance between two neighboring turning points is of order $L^{\frac{1}{\sigma}-1}$) and the nearest turning points to the real axis are $\pm L^{\frac{1}{\sigma}}(1 \pm \frac{\pi}{2\sigma L}i) + o(L^{\frac{1}{\sigma}-1})$. Hence the domain $D(\rho(\lambda), \theta(\lambda))$ has no turning point for $\rho(\lambda) = \frac{\pi}{4\sigma}L^{\frac{1}{\sigma}-1}$ and $\theta(\lambda) = \frac{\pi}{4\sigma}L^{-1}$. Then we see as in the previous example that its image by the map $z(x)$ includes the domain $F(a(\lambda))$ with $a(\lambda) = c\frac{\lambda}{L} = c\frac{\lambda}{\log\frac{1}{\lambda}}$ for some poisitive constant c.

For $\lambda = i\mu$ with $\mu > 0$ small i.e. $M = \log\frac{1}{\mu}$ large, the turning points in D_0 are $\pm(M + k\pi i)^{\frac{1}{\sigma}} + o(M^{\frac{1}{\sigma}-1})$ for some integers k and the nearest turning points to the real axis (apart from the real ones $\pm x^*(\mu)$) are $\pm M^{\frac{1}{\sigma}}(1 \pm \frac{\pi}{\sigma M}i) + o(M^{\frac{1}{\sigma}-1})$. Hence $D(\rho(\mu), \theta(\mu)) \cap \{z \in \mathbb{C}; \mathrm{Im}z > 0\}$ has no turning point for $\rho(\mu) = \frac{\pi}{2\sigma}M^{\frac{1}{\sigma}-1}$ and $\theta(\mu) = \frac{\pi}{2\sigma}M^{-1}$. As in the previous example, we see that its image by the map $z(x)$ includes the domain $G(b(\mu))$ with $b(\mu) = c\mu(\log\frac{1}{\mu})^{-1+\frac{1}{\sigma}}$. In fact we have in this case

$$\sqrt{A(x^*(\mu)+s)^2-\mu^2} \sim \sqrt{2\sigma}\mu M^{\frac{1}{2}-\frac{1}{2\sigma}}(-s)^{1/2},$$

and hence $|\int_{x^*(\mu)}^{x}\sqrt{A(t)^2-\mu^2}dt|$ is of order $\mu M^{-1+\frac{1}{\sigma}}$ when $|x - x^*(\mu)|$ is of order $M^{\frac{1}{\sigma}-1}$. □

Corollary 2.8. *Assume (A1), (A2) and (A4) with* $a(\lambda) \geq c\lambda^\beta$ *for some* $\beta > 0$ *and* $c > 0$*. Then the reflection coefficient* $R(\lambda, \epsilon)$ *is exponentially small with respect to* ϵ *uniformly for* $|\lambda| \geq \epsilon^\alpha$ *with any* $\alpha < 1/\beta$*. In particular, for potentials of Example 2.6 and 2.7,* $R(\lambda, \epsilon)$ *is exponentially small with respect to* ϵ *uniformly for* $|\lambda| \geq \epsilon^\alpha$ *with any* $\alpha < 1$*.*

Suppose that $\frac{\varepsilon}{b(\mu)}$ is small enough and let $r(\mu, \epsilon) := \log\left(-\frac{1}{m(\mu,\varepsilon)}\right)$ where the logarithm is defined near 1 with $\log 1 = 0$. Then the Bohr-Sommerfeld quantization condition (2.6) is equivalent to

$$S(\mu) = (2n+1)\pi\epsilon + i\epsilon r(\mu, \epsilon), \tag{2.8}$$

for some integer n. Let then $\mu_n = \mu_n(\epsilon)$ be the (unique) root of (2.8) and let $\mu_n^{\mathrm{WKB}} = \mu_n^{\mathrm{WKB}}(\epsilon)$ be the root of the equation

$$S(\mu) = (2n+1)\pi\epsilon. \tag{2.9}$$

By the previous theorem, we have, as $\epsilon \to 0$ with $\frac{\epsilon^2}{b(\mu_n)} \to 0$,

$$S(\mu_n) - S(\mu_n^{\mathrm{WKB}}) = i\epsilon r(\mu_n, \epsilon) = \mathcal{O}\left(\frac{\epsilon^2}{b(\mu_n)}\right).$$

For $\mu z = A(x)$, one has

$$S'(\mu) = 2\mu \int_1^{A_0/\mu} \frac{dz}{\sqrt{z^2-1}|A'(x)|}.$$

In the case of Example 2.6, $A'(x) \sim -d\mu^{1+\frac{1}{d}} z^{1+\frac{1}{d}}$, and in the case of Example 2.7 $A'(x) \sim -\sigma\mu z(\log\frac{1}{\mu z})^{1-\frac{1}{\sigma}}$, and hence we have,

$$|S'(\mu)| \geq \begin{cases} c\mu^{-\frac{1}{d}} \text{ (Example 2.6)}, \\ c\left(\log\frac{1}{\mu}\right)^{-1+\frac{1}{\sigma}} \text{ (Example 2.7)}, \end{cases}$$

for some positive constant c. Hence we have the following corollary.

Corollary 2.9. *Assume (A1), (A2), (A3) and (A5) with* $b(\mu) \geq c\mu^\beta$ *for some* $\beta > 0$ *and assume also* $|S'(\mu)| \geq c\mu^\gamma$ *for some* $c > 0$*. Then*

$$|\mu_n(\epsilon) - \mu_n^{\mathrm{WKB}}(\epsilon)| = o(\epsilon) \tag{2.10}$$

uniformly for $|\mu_n| \geq \epsilon^\alpha$ *with any* $\alpha < 1/(\beta+\gamma)$*. In particular, for potentials of Example 2.6, (2.10) holds uniformly for* $|\mu| \geq \epsilon^\alpha$ *with any* $\alpha < \frac{d}{d+1}$*. For potentials of Example 2.7, (2.10) holds uniformly for* $|\mu| \geq \epsilon^\alpha$ *with any* $\alpha < 1$*.*

In section 6, it will be Theorem 2.1 and the above corollary that will be applied to the focusing non-linear Schrödinger equation.

3. Exact WKB method for the Zakharov-Shabat system

In this section, we briefly review the exact WKB method applied to our operator L. Here we only assume (A1).

The eigenvalue problem of the operator L can be rewritten in the form

$$\frac{\varepsilon}{i}\frac{d}{dx}\mathbf{u} = M(x,\lambda)\mathbf{u}, \quad M(x,\lambda) = \begin{pmatrix} -\lambda & -iA(x) \\ iA(x) & \lambda \end{pmatrix} \tag{3.1}$$

where the unknown function $\mathbf{u}(x,\epsilon) = {}^t(u_1(x,\epsilon), u_2(x,\epsilon))$ is a column vector, ϵ is a small positive parameter, λ is a complex spectral parameter.

The zeros of $\det M(x,\lambda) = -A(x)^2 - \lambda^2$ are called *turning points*. Let Ω be a connected subdomain of D free from turning point. Then the map $x \mapsto z(x;\alpha)$ defined by

$$z(x;\alpha) = i\int_\alpha^x \sqrt{A(t)^2 + \lambda^2}\, dt\,, \tag{3.2}$$

for a fixed point α is conformal from Ω to $z(\Omega,\alpha)$.

We also define a function

$$H[z(x)] = \left(\frac{A(x) - i\lambda}{A(x) + i\lambda}\right)^{1/4},$$

which is holomorphic in Ω but multivalued around turning points, and a matrix valued function

$$Q(x) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} H(z(x))^{-1} & H(z(x))^{-1} \\ iH(z(x)) & -iH(z(x)) \end{pmatrix}.$$

Our WKB solutions are of the form

$$\mathbf{u}_\pm(x,\varepsilon) = e^{\pm z(x;\alpha)/\varepsilon} Q(x) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{1\pm 1}{2}} \mathbf{w}^\pm(z(x;\alpha),\varepsilon). \tag{3.3}$$

In the usual WKB theory, the vector valued symbols $\mathbf{w}^\pm(z,\varepsilon)$ are constructed as a power series in the parameter ε, which is in general divergent. Here we use the so-called *exact* WKB method along the lines of Gérard-Grigis [10] and Fujiié-Lasser-Nédélec [7]. This method consists in the resummation of this divergent series in the following way.

We take a point x_0 in Ω and construct $\mathbf{w}^\pm(z,\varepsilon)$ of the form

$$\mathbf{w}^\pm(z,\varepsilon) = \sum_{n=0}^\infty \mathbf{w}_n^\pm(z,\varepsilon) = \sum_{n=0}^\infty \begin{pmatrix} w_{2n}^\pm \\ w_{2n-1}^\pm \end{pmatrix} =: \begin{pmatrix} w_{\rm even}^\pm \\ w_{\rm odd}^\pm \end{pmatrix}, \tag{3.4}$$

where the scalar functions $w_n^\pm$ are defined inductively by

$$w_{-1}^\pm \equiv 0, \quad w_0^\pm \equiv 1, \tag{3.5}$$

and for $n \geq 1$,

$$\begin{cases} \dfrac{d}{dz} w_{2n}^{\pm} &= \mathcal{H}(z)\, w_{2n-1}^{\pm}, \\ \left(\dfrac{d}{dz} \pm \dfrac{2}{\varepsilon}\right) w_{2n-1}^{\pm} &= \mathcal{H}(z)\, w_{2n-2}^{\pm}. \end{cases} \tag{3.6}$$

with initial conditions at $z_0 = z(x_0)$

$$w_n^{\pm}|_{z=z_0} = 0 \quad (n \geq 1). \tag{3.7}$$

Here we defined

$$\mathcal{H}(z) := \frac{H_z'(z)}{H(z)} = \frac{d}{dz} \log H(z) = \frac{\lambda}{2} \frac{A_x'(x)}{(A(x)^2 + \lambda^2)^{3/2}}.$$

Notice that $\mathcal{H}(z)$ is holomorphic in $z(\Omega)$, and if β is a turning point of order k, it behaves, as $z \to z(\beta)$, like

$$\mathcal{H}(z) = \frac{\mp ik}{(2k+4)(z - z(\beta))}(1 + \mathcal{O}((z - z(\beta))^{2/(k+2)})), \tag{3.8}$$

where $\mp$ corresponds to whether β is zero of $A - i\lambda$ or $A + i\lambda$ ($\lambda \neq 0$).

The recurrence equations uniquely determine (at least in a neighborhood of x_0) the sequence of scalar functions $\{w_n^{\pm}(z, \varepsilon; z_0)\}_{n=-1}^{\infty}$, and hence the sequence of vector-valued functions $\{\mathbf{w}_n^{\pm}(z, \varepsilon; z_0)\}_{n=0}^{\infty}$.

The recursive relations (3.6) ,(3.7) can be written in the integral form:

$$w_{2n} = J(w_{2n-1}), \quad w_{2n-1} = I_{\pm}(w_{2n-2}) \quad (n \geq 1), \tag{3.9}$$

with two integral operators

$$J(f)(z) := \int_{\Gamma} \mathcal{H}(\zeta) f(\zeta) d\zeta, \tag{3.10}$$

$$I_{\pm}(f)(z) := \int_{\Gamma} e^{\pm 2(\zeta - z)/\varepsilon} \mathcal{H}(\zeta) f(\zeta) d\zeta, \tag{3.11}$$

where $\Gamma = \Gamma(z; z_0)$ is the image by $z = z(x; \alpha)$ of a path $\gamma(x; x_0)$ in Ω starting from x_0 and ending at x.

Thus we have constructed formal solutions, which we write from now on $\mathbf{u}_{\pm}(x, \varepsilon; \alpha, x_0)$, or simply $\mathbf{u}_{\pm}(x; \alpha, x_0)$ depending on a base point α for the phase and a base point x_0 for the symbol. This solution has the following important properties:

Theorem 3.1. *(i) The formal series are absolutely convergent in a neighborhood of* x_0.

(ii) Let $\Omega_{\pm}$ *be the set of* $x \in \Omega$ *such that there exists a path* $\gamma(x; x_0)$ *from* x_0 *to* x *in* Ω *along which* $\pm \mathrm{Re}\, z(x)$ *increases strictly (we will call such a path progressive). Then we have for each* $N \in \mathbb{N}$

$$\mathbf{w}^{\pm} - \sum_{n=0}^{N-1} \mathbf{w}_n^{\pm} = \mathcal{O}(\varepsilon^N),$$

$$w_{\text{even}}^{\pm} - \sum_{n=0}^{N-1} w_{2n}^{\pm} = \mathcal{O}(\varepsilon^N), \quad w_{\text{odd}}^{\pm} - \sum_{n=0}^{N-1} w_{2n+1}^{\pm} = \mathcal{O}(\varepsilon^N),$$

as $\varepsilon \to 0$, uniformly in any compact subset of $\Omega_\pm$. In particular, there we have

$$w_{\text{even}}^{\pm} = 1 + \mathcal{O}(\varepsilon), \quad w_{\text{odd}}^{\pm} = \mathcal{O}(\varepsilon).$$

(iii) The Wronskian of any two exact WKB solutions with different base points of amplitude are given by

$$\mathcal{W}(\mathbf{u}^+(x,\varepsilon;\alpha,x_0), \mathbf{u}^-(x,\varepsilon;\alpha,x_1)) = 4i\, w_{\text{even}}^{+}(z_1;z_0), \tag{3.12}$$

$$\mathcal{W}(\mathbf{u}^+(x,\varepsilon;\alpha,x_0), \mathbf{u}^+(x,\varepsilon;\alpha,x_1)) = -4i\, e^{2z_1/\varepsilon} w_{\text{odd}}^{+}(z_1;z_0), \tag{3.13}$$

where $z_j = z(x_j;\alpha)$ for $j = 0,1$ and $\mathcal{W}(\mathbf{f},\mathbf{g})$ is by definition the determinant of the matrix $(\mathbf{f},\mathbf{g})$.

Proof. The proof is almost the same as in references [10], [7] and [8], so we only point out the essence.

The main point lies in the "transport" equation (3.6) or equivalently (3.9). In the usual WKB construction in powers of ϵ, each coefficient is determined as an integral of the second derivative of the previous coefficient, which makes the sum divergent in general, whereas in the above construction, w_n is an integral of w_{n-1} itself, which makes the sums $\sum w_{2n}$ and $\sum w_{2n+1}$ convergent. More precisely, let K be any compact set in $z(\Omega)$. Then one has an estimate

$$|w_n^{\pm}(z,z_0)| \le C(AL)^n/n!$$

with some positive constant C and

$$L = \text{diam } (K), \quad A = \max_{z\in K} |\mathcal{H}(z)| \cdot \max(1, e^{2L/\varepsilon}).$$

As for the asymptotic property (ii), let us define a norm

$$\|f\| := \sup_{\Gamma(z;z_0)} |f| + \varepsilon \sup_{\Gamma(z;z_0)} |f'|$$

for holomorphic functions f on $z(\Omega)$. For $I_+(f)$, we have, by a change of variable $s = (\zeta - z)/\varepsilon$ and the Taylor expansion of $(\mathcal{H}f)(z+\varepsilon s)$ at z in the integral expression,

$$I_+(f)(z) = \frac{\varepsilon}{2}(1 - e^{2(z_0-z)/\varepsilon})(\mathcal{H}f)(z) + \varepsilon^2 \int_{(z_0-z)/\varepsilon}^{0} s e^{2s} \int_0^1 (\mathcal{H}f)'(z+s\varepsilon t)dtds. \tag{3.14}$$

It follows from the fact $\operatorname{Re} z < \operatorname{Re} z_0$ that

$$\sup_{\Gamma(z;z_0)} |J \circ I_+(f)| \le C\varepsilon \|f\|.$$

Moreover, using that $\frac{d}{dz}(J \circ I_+(f)) = \mathcal{H} I_+(f)$, we obtain

$$\|J \circ I_+(f)\| \le C\varepsilon \|f\|. \tag{3.15}$$

for some positive constant C. Hence we conclude that

$$\|w_{2n}^{+}\| = \|(J \circ I_+)^n(1)\| \le (C\varepsilon)^n,$$

$$\|w^+_{2n+1}\| = \|(J\circ I_+)^n(w^+_1)\| \le (C\varepsilon)^n\|w^+_1\|,$$

which prove (ii).

It remains to check (iii). We only prove (3.12). From the fact that $\det Q = -4i$, we immediately have

$$\begin{aligned}&\mathcal{W}(\mathbf{u}^+(x,\varepsilon;\alpha,x_0),\mathbf{u}^-(x,\varepsilon;\alpha,x_1))\\ &= \det Q\ \mathcal{W}\left(\begin{pmatrix}0&1\\1&0\end{pmatrix}\mathbf{w}^+(z,z_0),\mathbf{w}^-(z,z_1)\right)\\ &= -4i\ (w^+_{\mathrm{odd}}(z;z_0)w^-_{\mathrm{odd}}(z;z_1) - w^+_{\mathrm{even}}(z;z_0)w^-_{\mathrm{even}}(z;z_1)).\end{aligned}$$

This must be independent of x since the matrix M is trace free. Hence we can replace x in the right hand side by a particular point, say $x = x_1$. Then taking the previous into account, we get the proof for (3.12). The proof for the other formula is similar. □

Remark 3.2. *The constant C in (3.15) may depend on the energy λ. In fact, it becomes large when a turning point approaches the path $\gamma = \gamma(x;x_0)$. More precisely, let $\rho(\lambda)$ be the distance between the path γ and the "nearest" turning point x^* measured after the map $x\mapsto z = z(x)$ (see (3.2)):*

$$\rho(\lambda) = \mathrm{dist}(z(x^*), z(\gamma)).$$

Then we have, with a constant C' independent of λ,

$$C = \frac{C'}{\rho(\lambda)}. \tag{3.16}$$

This fact has already been proved and used in the Schrödinger case in [10] and [8] for the study of eigenvalues or resonances close to a barrier top of the potential, where the two turning points near the non-degenerate maximal point "pinch" a path along which the wronskian of two solutions on the opposite side of the barrier should be computed. Here we use this fact for the study of eigenvalues and the reflection coefficient near $\lambda = 0$. We briefly sketch the proof of (3.16) below.

To estimate $J\circ I_+(f)$, we should study integrals of type

$$I_1 = \int_{z_1}^z |\mathcal{H}(\tau)||\mathcal{H}(\tau + ths)|d\tau, \quad I_2 = \int_{z_1}^z |\mathcal{H}(\tau)||\mathcal{H}'(\tau + ths)|d\tau.$$

Because of the Cauchy-Schwarz inequality and (3.14) we only need to estimate $\int_{z_1}^z|\mathcal{H}(t)|^2dt$, $\int_{z_1}^z|\mathcal{H}'(t)|^2dt$. Since the singularity of the function $\mathcal{H}(z)$ at the image of turning points is like (3.8), these two integrals are of type $\int_{\mathbb{R}}\frac{dt}{t^2+\rho^2}$, $\int_{\mathbb{R}}\frac{dt}{(t^2+\rho^2)^2}$ respectively. This gives $I_1\le\frac{c}{\rho}$, $I_2\le\frac{c}{\rho^2}$, and consequently

$$\sup|J\circ I_+(f)| \le C_1\left(\frac{\varepsilon}{\rho} + (\frac{\varepsilon}{\rho})^2\right)\sup|f| + C_1\frac{\varepsilon^2}{\rho}\sup|f'|.$$

Using again $\frac{d}{dz}(J\circ I_+(f)) = \mathcal{H}I_+(f)$, we get (3.16) in the estimate (3.15).

The level curves of $\operatorname{Re} z(x)$ in the x-complex plane

$$\{x \in D; \mathrm{Re} z(x) = \text{const.}\}$$

are called *Stokes curves*. (Sometimes in the literature it is the level curves of $\mathrm{Im} z(x)$, especially only those passing through turning points, that are called Stokes curves, but we employ the former definition.) The geometric configuration of Stokes curves is useful for us to know the domain of validity of the asymptotic expansion of WKB solutions.

Notice that from a simple turning point exactly three Stokes curves emanate and the angles between two of them are all $2\pi/3$ at that point.

4. Reflection coefficient

Here we compute the reflection coefficient for real positive λ. The computation for negative λ is similar.

We construct the Jost solutions as exact WKB solutions. We define four exact WKB solutions:

$$\widetilde{\mathbf{v}}^r_{\pm}(x,\varepsilon) = e^{\pm z^r(x)/\varepsilon} Q(x) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{1\pm 1}{2}} \mathbf{w}^{\pm}_r(x,\varepsilon), \tag{4.1}$$

$$\widetilde{\mathbf{v}}^l_{\pm}(x,\varepsilon) = e^{\pm z^l(x)/\varepsilon} Q(x) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{1\pm 1}{2}} \mathbf{w}^{\pm}_l(x,\varepsilon), \tag{4.2}$$

where the phase function are

$$z^r(x) := i\lambda x + i\int_{+\infty}^{x} (\sqrt{A(t)^2+\lambda^2} - \lambda)dt,$$

$$z^l(x) := i\lambda x + i\int_{-\infty}^{x} (\sqrt{A(t)^2+\lambda^2} - \lambda)dt,$$

which are both primitives of $i\sqrt{A(x)^2+\lambda^2}$.

As base point of the symbol $\mathbf{w}^{\pm}_r(x,\varepsilon)$ and $\mathbf{w}^{\pm}_l(x,\varepsilon)$, we choose $e^{\pm i\theta_0}\infty$ and $e^{i(\pi\mp\theta_0)}\infty$ respectively. We recall that θ_0 is the positive angle of the sector at infinity of the domain D (see the assumption (A1)). More precisely, we take, as the contour for the integral operators $I_{\pm}, J$ the image by the map $x \mapsto z^r(x)$ (resp. $x \mapsto z^l(x)$) of a curve from $e^{\pm i\theta_0}\infty$ (rest. $e^{i(\pi\mp\theta_0)}\infty$) to x, which is transverse to the Stokes curves. This is possible for any $x \in D(\mu, R, \theta_0)$ (by which we actually denote $D(\mu,\theta_0)$ where $R = \frac{\mu}{tan\theta_0}$ is the value of the real part of x for which the horizontal part of the upper boundary changes to an inclined line; see the definition of the domain D in assumption (A1)), if μ is sufficiently small and R is sufficiently large, because then $D(\mu, R, \theta_0)$ contains no turning point, the Stokes curves are asymptotic to horizontal lines as $\mathrm{Re} x \to \pm\infty$ and the real axis is itself a Stokes curve. We take a branch for the functions $(A(x) \pm i\lambda)^{1/2}$ and $(A(x) \pm i\lambda)^{1/4}$ in such a way that the argument of these functions tends to 0 as $\lambda \to 0$ (recall that $A(x)$ is positive). Then the real part of the phase $z^r(x), z^l(x)$ or $z(x;\alpha)$ for any α increases as $\operatorname{Im} x$ decreases. Remark also that, by this determination, one has

$$H(x) \to e^{-\frac{\pi}{4} i \operatorname{sgn}\lambda} \quad \text{as } |x| \to \infty.$$

Hence, for $\lambda > 0$, we have

$$Q(x) \to 2e^{\frac{\pi}{4} i} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{as } |x| \to \infty.$$

These exact WKB solutions have the following trivial relations with the Jost solutions:

$$\mathbf{f}_{\pm}^{r} = \mp 2e^{\pi i/4} \widetilde{\mathbf{v}}_{\pm}^{r} \quad \mathbf{f}_{\pm}^{l} = \mp 2e^{\pi i/4} \widetilde{\mathbf{v}}_{\pm}^{l}. \tag{4.3}$$

We further modify slightly our exact WKB solutions. Let $\mathbf{v}_{\pm}^{r}$, $\mathbf{v}_{\pm}^{l}$ be the exact WKB solutions defined just like $\widetilde{\mathbf{v}}_{\pm}^{r}$, $\widetilde{\mathbf{v}}_{\pm}^{l}$ but with $z(x;0)$ for the phase. Then we obviously have

$$\widetilde{\mathbf{v}}_{\pm}^{r} = e^{\pm z_r(0)/\varepsilon} \mathbf{v}_{\pm}^{r}, \quad \widetilde{\mathbf{v}}_{\pm}^{l} = e^{\pm z_l(0)/\varepsilon} \mathbf{v}_{\pm}^{l}. \tag{4.4}$$

In terms of these WKB solutions, the reflection coefficient is expressed by

$$R(\lambda, \varepsilon) = -\frac{\mathcal{W}(\mathbf{v}_{+}^{r}, \mathbf{v}_{+}^{l})}{\mathcal{W}(\mathbf{v}_{+}^{l}, \mathbf{v}_{-}^{r})} e^{2z^{r}(0)/\varepsilon}, \tag{4.5}$$

where we recall that

$$z^{r}(0) = -i \int_{0}^{+\infty} (\sqrt{A(t)^2 + \lambda^2} - \lambda) dt. \tag{4.6}$$

The wronskians appearing in (4.5) can be expressed by the functions $w_{even}^{\pm}$ and $w_{odd}^{\pm}$ using Theorem 3.1.

First, $\mathcal{W}(\mathbf{v}_{+}^{l}, \mathbf{v}_{-}^{r})$ is given by

$$\mathcal{W}(\mathbf{v}_{+}^{l}, \mathbf{v}_{-}^{r}) = 4iw_{even}^{+}(e^{-i\theta_0}\infty; e^{i(\pi-\theta_0)}\infty) =: 4iw_{even}^{+}(z(e^{-i\theta_0}\infty; 0); z(e^{i(\pi-\theta_0)}\infty; 0)).$$

On the other hand, we should express the wronskian $\mathcal{W}(\mathbf{v}_{+}^{r}, \mathbf{v}_{+}^{l})$ via another basis of solutions since there is no progressive path between $e^{i\theta_0}\infty$ and $e^{i(\pi-\theta_0)}\infty$. We take exact WKB solutions defined with the phase (3.2)

$$\mathbf{v}_{+}^{0}(x, \varepsilon) := \mathbf{u}_{+}(x, \varepsilon; 0, x_{+}),$$

$$\mathbf{v}_{-}^{0}(x, \varepsilon) := \mathbf{u}_{-}(x, \varepsilon; 0, x_{-}),$$

where we take $x_{\pm} \in D \cap \mathbb{C}_{\pm}$ near the origin. We can write

$$\begin{cases} \mathbf{v}_{+}^{r} = c_{+}^{r} \mathbf{v}_{+}^{0} + c_{-}^{r} \mathbf{v}_{-}^{0}, \\ \mathbf{v}_{+}^{l} = c_{+}^{l} \mathbf{v}_{+}^{0} + c_{-}^{l} \mathbf{v}_{-}^{0}, \end{cases}$$

with

$$c_{+}^{r} = \frac{\mathcal{W}(\mathbf{v}_{+}^{r}, \mathbf{v}_{-}^{0})}{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{-}^{0})}, \quad c_{-}^{r} = \frac{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{+}^{r})}{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{-}^{0})},$$

$$c_{+}^{l} = \frac{\mathcal{W}(\mathbf{v}_{+}^{l}, \mathbf{v}_{-}^{0})}{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{-}^{0})}, \quad c_{-}^{l} = \frac{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{+}^{l})}{\mathcal{W}(\mathbf{v}_{+}^{0}, \mathbf{v}_{-}^{0})}.$$

Since

$$\mathcal{W}(\mathbf{v}_+^r, \mathbf{v}_+^l) = (c_+^r c_-^l - c_-^r c_+^l)\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_-^0),$$

we have

$$\mathcal{W}(\mathbf{v}_+^r, \mathbf{v}_+^l) = \frac{\mathcal{W}(\mathbf{v}_+^r, \mathbf{v}_-^0)\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_+^l) - \mathcal{W}(\mathbf{v}_+^l, \mathbf{v}_-^0)\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_+^r)}{\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_-^0)}.$$

The wronskian formulae of Theorem 3.1 give us the following expressions.

$$\begin{aligned}
\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_-^0) &= 4iw_{even}^+(x_-; x_+),\\
\mathcal{W}(\mathbf{v}_+^r, \mathbf{v}_-^0) &= 4iw_{even}^+(x_-; e^{i\theta_0}\infty),\\
\mathcal{W}(\mathbf{v}_+^l, \mathbf{v}_-^0) &= 4iw_{even}^+(x_-; e^{i(\pi-\theta_0)}\infty),\\
\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_+^r) &= 4iw_{odd}^+(x_+; e^{i\theta_0}\infty)e^{2\sigma/\varepsilon},\\
\mathcal{W}(\mathbf{v}_+^0, \mathbf{v}_+^l) &= 4iw_{odd}^+(x_+; e^{i(\pi-\theta_0)}\infty)e^{2\sigma/\varepsilon},
\end{aligned}$$

where

$$\sigma = z(x_+; 0) = i\int_0^{x_+} \sqrt{A(t)^2 + \lambda^2}dt. \tag{4.7}$$

Notice that σ has a negative real part since $\operatorname{Im} x_+ > 0$.

Summing up, we arrive at the following formula for the reflection coefficient.

Proposition 4.1. *Let $\lambda > 0$ and σ defined by (4.7). Then one has*

$$R(\lambda, \varepsilon) = -e^{2(\sigma + z_r(0))/\varepsilon} \times$$
$$\frac{w_{even}^+(x_-; e^{i\theta_0}\infty)w_{odd}^+(x_+; e^{i(\pi-\theta_0)}\infty) - w_{even}^+(x_-; e^{i(\pi-\theta_0)}\infty)w_{odd}^+(x_+; e^{i\theta_0}\infty)}{w_{even}^+(x_-; x_+)w_{even}^+(e^{-i\theta_0}\infty; e^{i(\pi-\theta_0)}\infty)}$$

4.1. **Proof of Theorem 2.1.** In this theorem, it is assumed that $\lambda > \delta$ for some positive δ independent of ε.

First recall that $z_r(0)$ is purely imaginary (see (4.6)) and does not affect the absolute value of the reflection coefficient.

Next, as long as $\operatorname{Im} x_+$ is positive and small enough, we have $\operatorname{Re}\sigma < 0$ and we can find a progressive path for each couple of points in w_{even}^+ and w_{odd}^+ of the formula of Proposition 4.1, which means that these quantities are $1 + \mathcal{O}(\varepsilon)$ and $\mathcal{O}(\varepsilon)$ respectively as $\varepsilon \to 0$.

This gives the proof of Theorem 2.1.

4.2. **Proof of Theorem 2.4.** The assumption (A4) together with the obvious fact that $z(x)$ maps a λ-independent neighborhood of $x = 0$ to a λ-independent neighborhood of $z = 0$ imply that the image of $D(\rho(\lambda), \theta(\lambda))$ includes a domain of the form

$$\left\{ z \in \mathbb{C}; |\operatorname{Im} z| > \frac{2|\operatorname{Re} z|}{a(\lambda)} - c \right\}$$

for some positive constant c.

In the computation of the asymptotic behavior of the wronskians appearing in Proposition 4.1, for example $w^+_{even}(x_-; e^{i\theta_0}\infty)$ and $w^+_{odd}(x_+; e^{i\theta_0}\infty)$, we take, as the integral contour for (3.10) and (3.11), the half lines $\{\operatorname{Im} z = -\frac{2\operatorname{Re} z}{a(\lambda)} + \frac{c}{2}, \operatorname{Re} z \leq \frac{ca(\lambda)}{4}\}$ and $\{\operatorname{Im} z = -\frac{2\operatorname{Re} z}{a(\lambda)} - \frac{c}{2}, \operatorname{Re} z \leq -\frac{ca(\lambda)}{4}\}$ respectively, oriented in such a way that $\operatorname{Re} z$ increases (we take $x_\pm$ so that $z(x_\pm) = \mp\frac{ca(\lambda)}{4}$). Then the quantity $\rho(\lambda)$ in Remark 3.2, which measures the distance from the contour to the nearest turning points, is estimated from below by a constant multiple of $a(\lambda)$. This proves Theorem 2.4.

5. Eigenvalues

In this section, we study the eigenvalue problem of the operator L. It is known ([21], [19]) that for our kind of potential $A(x)$ the eigenvalues are all simple and purely imaginary with imaginary part in $[-A(0), A(0)]$.

For $\lambda = i\mu$, $\mu \in (0, A(0))$, there are exactly two simple turning points $x^*(\mu) > 0$ and $-x^*(\mu)$ on the real axis. There is no other turning point in the complex domain $D(\mu_0, R_0, \theta_0)$ if we take μ_0 sufficiently small, R_0 sufficiently large and θ_0 sufficiently small depending on each $\mu > 0$.

The interval $[-x^*(\mu), x^*(\mu)]$ is a Stokes curve on which $A(x)^2 - \mu^2 \geq 0$. There are two other Stokes curves emanating from each of these turning points.

We take two branch cuts along Stokes curves, one from $x^*(\mu)$ with angle $\pi/3$ and the other from $-x^*(\mu)$ with angle $4\pi/3$, and determine the branch of $(A(x)^2 - \mu^2)^{1/2}$ and $(A(x) \pm \mu)^{1/4}$ so that they are all real and positive on the interval $[-x^*(\mu), x^*(\mu)]$. Then automatically

$$(A(x)^2 - \mu^2)^{1/2} \in i\mathbb{R}_+ \text{ in } (-\infty, -x^*(\mu)] \cup [x^*(\mu), \infty),$$

$$(A(x) - \mu)^{1/4} \in e^{i\pi/4}\mathbb{R}_+ \text{ in } (-\infty, -x^*(\mu)] \cup [x^*(\mu), \infty),$$

$$(A(x) + \mu)^{1/4} \in \mathbb{R}_+ \text{ on } \mathbb{R}.$$

Now we define several exact WKB solutions. The 5 Stokes curves in D divide the domain D into 4 connected regions D_r, D_l, D_u and D_d (if D is chosen sufficiently small as mentioned above). The regions D_r and D_l include $(x^*(\mu), +\infty)$ and $(-\infty, -x^*(\mu))$ respectively, and $D_u \subset \mathbb{C}_+ = \{x \in \mathbb{C}; \operatorname{Im} x > 0\}$ and $D_d \in \mathbb{C}_-$ share $(-x^*(\mu), x^*(\mu))$ as a part of their boundary. We take four base points $x_1 \in D_l$, $x_2 \in D_u$, $x_3 \in D_d$, $x_4 \in D_r$. With the above determination, the real part of $z(x)$ increases along curves from x_2 to x_1, from x_2 to x_3, and from x_4 to x_3. Taking this into account, we define six exact WKB solutions.

$$\mathbf{v}_1(x, \varepsilon, \mu) := \mathbf{u}^-(x, \varepsilon; x^*(\mu), x_1),$$
$$\mathbf{v}_2(x, \varepsilon, \mu) := \mathbf{u}^+(x, \varepsilon; x^*(\mu), x_2), \quad \widetilde{\mathbf{v}}_2(x, \varepsilon, \mu) := \mathbf{u}^+(x, \varepsilon; -x^*(\mu), x_2),$$
$$\mathbf{v}_3(x, \varepsilon, \mu) := \mathbf{u}^-(x, \varepsilon; x^*(\mu), x_3), \quad \widetilde{\mathbf{v}}_3(x, \varepsilon, \mu) := \mathbf{u}^-(x, \varepsilon; -x^*(\mu), x_3),$$
$$\mathbf{v}_4(x, \varepsilon, \mu) := \mathbf{u}^+(x, \varepsilon; -x^*(\mu), x_4).$$

Exactly as in [10] in the Schrödinger case, we know

Lemma 5.1. *For each* $\varepsilon > 0$, $\mathbf{v}_1(x,\varepsilon,\mu) \in (L^2(\mathbb{R}_+))^2$, $\mathbf{v}_4(x,\varepsilon,\mu) \in (L^2(\mathbb{R}_-))^2$.

Remark 5.2. *We could have chosen the base point* $-\infty$ *for* $\mathbf{v}_1$ *instead of* x_1, *and* $+\infty$ *for* $\mathbf{v}_4$ *instead of* x_4, *as in the study of the reflection coefficient.*

This lemma immediately implies

Proposition 5.3. $\lambda = i\mu$ *is an eigenvalue if and only if the wronskian between* $\mathbf{v}_1(x,\varepsilon,\mu)$ *and* $\mathbf{v}_4(x,\varepsilon,\mu)$ *vanishes.*

The pairs $(\mathbf{v}_2, \mathbf{v}_3)$ and $(\widetilde{\mathbf{v}}_2, \widetilde{\mathbf{v}}_3)$ of solutions, form bases as seen by Theorem 3.1 (iii) (3.12):

$$\mathcal{W}(\mathbf{v}_2, \mathbf{v}_3) = \mathcal{W}(\widetilde{\mathbf{v}}_2, \widetilde{\mathbf{v}}_3) = 4iw^+_{even}(x_3, \varepsilon; x_2) = 4i + \mathcal{O}(\varepsilon) \neq 0. \tag{5.1}$$

These two bases have the following trivial relations.

Lemma 5.4. *Let* $S(\mu)$ *be the action integral defined by* (2.5). *Then one has*

$$\widetilde{\mathbf{v}}_2 = e^{-iS(\mu)/\varepsilon}\mathbf{v}_2, \;\; \widetilde{\mathbf{v}}_3 = e^{iS(\mu)/\varepsilon}\mathbf{v}_3.$$

In order to compute the asymptotic behavior of $\mathcal{W}(\mathbf{v}_1, \mathbf{v}_4)$, we need to express $\mathbf{v}_1$ and $\mathbf{v}_4$ in terms of $\mathbf{v}_2$ and $\mathbf{v}_3$ or in terms of $\widetilde{\mathbf{v}}_2$ and $\widetilde{\mathbf{v}}_3$:

$$\mathbf{v}_1 = c_2(\varepsilon,\mu)\mathbf{v}_2 + c_3(\varepsilon,\mu)\mathbf{v}_3, \;\; \mathbf{v}_4 = \widetilde{c}_2(\varepsilon,\mu)\widetilde{\mathbf{v}}_2 + \widetilde{c}_3(\varepsilon,\mu)\widetilde{\mathbf{v}}_3.$$

The coefficients are written in terms of wronskians:

$$c_2 = \frac{\mathcal{W}(\mathbf{v}_1, \mathbf{v}_3)}{\mathcal{W}(\mathbf{v}_2, \mathbf{v}_3)}, \;\; c_3 = \frac{\mathcal{W}(\mathbf{v}_1, \mathbf{v}_2)}{\mathcal{W}(\mathbf{v}_3, \mathbf{v}_2)},$$
$$\widetilde{c}_2 = \frac{\mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_3)}{\mathcal{W}(\widetilde{\mathbf{v}}_2, \widetilde{\mathbf{v}}_3)}, \;\; \widetilde{c}_3 = \frac{\mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_2)}{\mathcal{W}(\widetilde{\mathbf{v}}_3, \widetilde{\mathbf{v}}_2)}.$$

Thus a computation of $\mathcal{W}(\mathbf{v}_1, \mathbf{v}_4)$ leads us to the following quantization condition of eigenvalues in terms of wronskians of exact WKB solutions:

Proposition 5.5. *Let* $m(\mu,\varepsilon)$ *be a function defined by*

$$m(\mu,\varepsilon) = \frac{\mathcal{W}(\mathbf{v}_1, \mathbf{v}_3)\mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_2)}{\mathcal{W}(\mathbf{v}_1, \mathbf{v}_2)\mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_3)}.$$

Then $\lambda = i\mu$ *is an eigenvalue if and only if*

$$m(\mu,\varepsilon)e^{2iS(\mu)/\varepsilon} = 1.$$

Now we study the asymptotic behavior of the wronskians appearing in Proposition 5.5 as $\varepsilon \to 0$ using Theorem 3.1.

Lemma 5.6. *It holds that*

$$\mathcal{W}(\mathbf{v}_1, \mathbf{v}_2) = -4iw^+_{even}(x_1, \varepsilon; x_2), \quad \mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_3) = 4iw^+_{even}(x_3, \varepsilon; x_4),$$
$$\mathcal{W}(\mathbf{v}_1, \mathbf{v}_3) = 4w^+_{even}(x_3, \varepsilon; \hat{x}_1), \quad \mathcal{W}(\mathbf{v}_4, \widetilde{\mathbf{v}}_2) = -4w^+_{even}(\hat{x}_4, \varepsilon; x_2),$$

where $\hat{x}_1$ *is the same point as* x_1 *but on the Riemann sheet continued from* x_3 *crossing the branch cut from* $-x^*(\mu)$ *and similarly* $\hat{x}_4$ *is the same point* *as x_4 but on the Riemann sheet continued from x_2 crossing the branch cut from $x^*(\mu)$. Consequently, we have*

$$m(\mu,h) = -\frac{w^+_{even}(x_3,\varepsilon;\hat{x}_1)w^+_{even}(\hat{x}_4,\varepsilon;x_2)}{w^+_{even}(x_1,\varepsilon;x_2)w^+_{even}(x_3,\varepsilon;x_4)}. \tag{5.2}$$

Proof. The formulas for $\mathcal{W}(\mathbf{v}_1,\mathbf{v}_2)$ and $\mathcal{W}(\mathbf{v}_4,\widetilde{\mathbf{v}}_3)$ follow difrectly from (3.12).

For the computation of $\mathcal{W}(\mathbf{v}_1,\mathbf{v}_3)$, we should be careful with the branch cut lying between x_1 and x_3. In order to compute the wronskian on a curve along which $\mathrm{Re}z$ is strictly increasing, we have to rewrite $\mathbf{v}_1$, say, on the Riemann sheet continued from x_3 along this curve crossing the cut.

Let x be a point near x_1 and $\hat{x}$ the same point as x but on the Riemann sheet mentioned above. Then we have, writing $\alpha = x^*(\mu)$ for simplicity,

$$z(x;\alpha) = -z(\hat{x};\alpha),\ H(x) = iH(\hat{x}),\ \mathbf{w}^{\pm}(x,\varepsilon;x_1) = \mathbf{w}^{\mp}(\hat{x},\varepsilon;x_1).$$

In fact, the first identity is obvious. For the second one, remark that the turning point $\beta := -x^*(\mu)$ is a zero of $A(x)-\mu$. The third one can be seen from the first one and (3.6). It follows that

$$\begin{aligned}\mathbf{v}_1 = \mathbf{u}^-(x,\varepsilon;\alpha,x_1) &= e^{-z(x;\alpha)/\varepsilon}Q(x)\mathbf{w}^-(x,\varepsilon;x_1)\\ &= e^{+z(\hat{x};\alpha)/\varepsilon}\begin{pmatrix}-i & 0\\ 0 & i\end{pmatrix}Q(\hat{x})\mathbf{w}^+(\hat{x},\varepsilon;x_1)\\ &= -ie^{+z(\hat{x};\alpha)/\varepsilon}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}Q(\hat{x})\mathbf{w}^+(\hat{x},\varepsilon;x_1),\end{aligned}$$

Since $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}Q(\hat{x}) = Q(\hat{x})\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$, we finally get

$$\mathbf{v}_1 = -i\mathbf{u}^+(\hat{x},\varepsilon;\alpha,\hat{x}_1).$$

Now we can compute the wronskian $\mathcal{W}(\mathbf{v}_1,\mathbf{v}_3)$ by the formula (3.12) and

$$\mathcal{W}(\mathbf{v}_1,\mathbf{v}_3) = \mathcal{W}\left(-i\mathbf{u}^+(\hat{x},\varepsilon;\alpha,\hat{x}_1),\mathbf{u}^-(\hat{x},\varepsilon;\alpha,x_3)\right) = 4w^+_{even}(x_3,\varepsilon;\hat{x}_1).$$

In the same way, we rewrite $\mathbf{v}_4$ using the branch continued from x_2 to x_4 crossing the cut emanating from $x^*(\mu)$;

$$\mathbf{v}_4 = -i\mathbf{u}^-(\hat{x},\varepsilon;\beta,\hat{x}_4).$$

This enables to compute the wronskian

$$\mathcal{W}(\mathbf{v}_4,\widetilde{\mathbf{v}}_2) = -4w^+_{even}(\hat{x}_4,\varepsilon;x_2).$$

□

5.1. **Proof of Theorem 2.2.** For the time let us ignore the assumption (A3) and simply assume that μ stays away from $A(0)$; $\mu\in[\delta, A(0)-\delta]$ for some ε-independent positive constant δ. Then the configuration of Stokes curves in D is as explained at the beginning of this section if μ_0, R_0, θ_0 are suitably chosen.

Notice that the asymptotic behavior of the quantities w^+_{even} in the right hand side of (5.2) are all well known by Theorem 3.1 (ii) to be $1+\mathcal{O}(\varepsilon)$

uniformly with respect to μ in the interval $[\delta, A(0)-\delta]$. In fact, there exist progressive paths from x_2 to x_1, from x_4 to x_3, from $\hat{x}_1$ to x_3 and from x_2 to $\hat{x}_4$. Hence we obtain

$$m(\mu,\varepsilon) = -1 + \mathcal{O}(\varepsilon), \tag{5.3}$$

Next we consider the case where $\mu \to A(0)$. In this limit, the two turning points $x^*(\mu)$ and $-x^*(\mu)$ coalesce at the origin $x = 0$ and they become a double turning point.

Under the additional condition (A3), however, there is no other complex turning point converging to the origin and the Stokes geometry does not change in $D(\mu_0, R_0, \theta_0)$ with small enough ε-independent μ_0 except that the two turning points get closer and closer.

Moreover, it is important to notice that the four paths from x_2 to x_1, from x_4 to x_3, from $\hat{x}_1$ to x_3 and from x_2 to $\hat{x}_4$ are not *pinched* by these two turning points. This fact implies that the asymptotic formula (5.3) holds true also for such energies μ.

5.2. **Proof of Theorem 2.5.** Here we take $x_1 = -\infty$ and $x_4 = +\infty$ (see Remark 5.2).

For the computation of the asymptotic behavior of the wronskians appearing in (5.2), say $w^+_{even}(\hat{x}_4, \varepsilon; x_2)$, we take, as the integral contour for (3.10) and (3.11), a curve from $z(x_2)$ (x_2 should be taken so that $z(x_2) \in G(b(\mu))$) passing inside the tube $G(b(\mu))$ with increasing real part to $z(\hat{x}_4)$ (i.e. going to ∞ passing through the tube in the upper half plane of $G(b(\mu))$).

Then the quantity $\rho(\lambda)$ in Remark 3.2, which measures the distance from the contour to the nearest turning points, is estimated from below by a constant multiple of $b(\lambda)$. This proves Theorem 2.5.

6. Application of the exact WKB results to the focusing nonlinear Schrödinger equation

We are now ready to give a rigorous justification of the asymptotics (1.4) stated in the introduction for a general class of initial data, *without* having to replace them by their WKB approximation. The proof is a variant of Chapter 4 in [17] (see also the first appendix below for the definition of the Riemann-Hilbert problem and the second appendix for the deformations implemented involving a well-chosen phase function g), as improved in [22]. Having estimated the error of the WKB approximation at the level of the scattering data, this error can be built into the Riemann-Hilbert analysis as another layer of approximtion. [1]

In [17] we considered separately two complementary sets for λ: a disc centered at 0 with radius of order ϵ^δ, for some $\delta \in (1/2, 1)$ and the complement of that disc. In the complement of the disc we needed some delicate rigorous estimates while inside the disc we only relied on some symmetry properties.

[1] This is the strategy we explicitly proposed in [17].

This was enough to be able to approximate the "given" Riemann-Hilbert problem (for the WKB-approximated pure soliton data) by an approximate Riemann-Hilbert problem which we could eventually analyse. The approximation was good enough away from zero (and this is all we are interested in because the solution of the NLS equation only depends on the behaviour of the solution of the Riemann-Hilbert problem near infinity).

Now, it turns out that the restriction that $\delta \in (1/2, 1)$ is too strong if we want to use the results of the previous sections. Conveniently, there is an improvement of our argument in [22], based on the observation that the approximation of the Blaschke product (see below) involving the (WKB approximations of the) eigenvalues by a logarithmic integral is better done separately, with slightly different choices, in different sides of the segment $[-iA(0), iA(0)]$. [2]

Consequently, the choice of the small circle separating the two cases for λ above can be allowed to have a radius independent of ϵ as long as it is small enough in the sense of meeting some requirement spelled out in [22] (which in turn is imposed by the asymptotic analysis of the approximate Riemann-Hilbert problem). [3]§ This is certainly good enough for our purposes here.

This point in [22] is only explicitly detailed for the very specific case where the distribution of the eigenvalues is uniform (the Satsuma-Yajima case) and where there is no reflection coefficient, but it is clear that after an obvious modification it can apply to our general case. [4]

Under the assumptions (A1), (A2), (A3), (A5), with $b(\mu) \geq c\mu^{\beta}$ for some $\beta > 0$ and assuming that the action integral S satisfies $|S'(\mu)| \geq c\mu^{\gamma}$ for some $c > 0$, with $\beta + \gamma < 2$, we have the main result of this section.

Proposition 6.1. *Assuming the existence of the finite gap ansatz and also that the density of eigenvalues admits an analytic extension in the upper half-plane, the asymptotics* (1.4) *stated in the introduction are valid.*

Proof. The proof is a variant of Chapter 4 in [17], as improved in [22]. There are two modifications.

First, because of the non-triviality of the reflection coefficient, the Riemann-Hilbert contour is augmented by the real line (oriented from left to right so that the "+"-side is on top). Still, we use exactly the same g as in [17]. The important condition

$$g(\lambda^*) + g(\lambda)^* = 0 \text{ for all } \lambda \in \mathbb{C} \setminus (C \cup C^*),$$

[2] As we explain in [15] and [16] this corresponds to different sheets of the logarithmic kernel in this integral. Different approximations are required in different sheets for best results.

[3] We could still ignore the improvement in [22] and give a different argument involving different circles in different steps of the Riemann-Hilbert sequence. We feel that the argument would become a bit more cumbersome.

[4] The function $\theta^0(\lambda) = i\pi\lambda + \pi A$ of [22] has to be replaced by the integral of the eigenvalue density in the general case.

(recall that C is a contour lying in the closed upper half lane, including 0 and encricling the segment $(0, iA_0]$ and approaching 0 at an angle strictly between 0 and $\pi/2$ from the first quadrant) implies that $g(\lambda)$ is imaginary on the real line. Here * denotes complex conjugation.

As a consequence, the fact that the reflection coefficient is exponentially small outside of a small open disc D with center 0 and radius small enough for the asymptotic analysis to go through means that the jump matrix on $\mathbb{R} \setminus D$ is of the form $I +$ *exponentially small* uniformly.

Within the disc the important observation is that the (no more trivial) jump matrix still respects the Schwarz reflection symmetry conditions and the positive definiteness condition needed for the application of the results in the Appendix A of [17].

The second modification comes from the eigenvalues. We first note ([21], [19]) that the eigenvalues are simple, imaginary, their total number $2N$ is finite and

$$N = \left[1/2 + \frac{1}{\epsilon\pi}||A||_{L^1}\right]$$

where [.] here denotes the integer part of a real. [5]

As a result, the number of the actual eigenvalues is *equal* to the number of the "Bohr-Sommerfeld-approximate eigenvalues", i.e. the *exact* roots of $e^{2iS(\mu)/\varepsilon} = -1$.

Now, our rigorous estimates in the previous section give a 1-1 correspondence between the eigenvalues λ_n and the Bohr-Sommerfeld approximations λ_n^{WKB}, as long as $|\lambda_n^{\mathrm{WKB}}|$ are greater than $G\epsilon^\alpha$, for some constant $G > 0$ independent of ϵ and for some $\alpha \in (0,1)$ and in fact $\lambda_n - \lambda_n^{\mathrm{WKB}} = o(\epsilon)$ uniformly in that set.

It follows that there is also a 1-1 correspondence between the rest of the eigenvalues λ_n and the Bohr-Sommerfeld approximations λ_n^{WKB}, in which case λ_n^{WKB} are in the closed disc of radius $G\epsilon^\alpha$ and thus λ_n have to be in a disc of radius of order $O(\epsilon^\alpha)$ (possibly somewhat larger than $G\epsilon^\alpha$). Clearly then $\lambda_n - \lambda_n^{\mathrm{WKB}} = O(\epsilon^\alpha)$.

The crucial quantities to consider are the two "Blaschke" products (see the first section of the appendix)

$$\left(\prod_{n=0}^{N-1} \frac{\lambda - \lambda_n^*}{\lambda - \lambda_n}\right)$$

one for the actual eigenvalues, say P, and one for their WKB approximaton say P_{WKB}.

---

[5] In [21] the exact estimate is stated in the abstract but a proof is only presented for the case of A with compact support. Still, the proof presented easily generalises for the case of non-compact support. In fact, the crucial integral in (2.15) of [21] is positive *also* in the non-compact support case (as is proved for example in [19]) and this in turn implies the exact estimate for the number of eigenvalues.

Let the open disc D^{ϵ^α} have center 0 and radius $G\epsilon^\alpha$. Outside the disc D^{ϵ^α} we have shown (in Corollary 2.9) that the difference between the actual eigenvalues and their formal WKB approximation (which was used in [17]) is of order $o(\epsilon)$ uniformly. Let P^{ϵ^α} and $P^{\epsilon^\alpha}_{\mathrm{WKB}}$ be the "Blaschke" products as above, but excluding the eigenvalues λ_n lying in D^{ϵ^α}.

A short calculation gives $P^{\epsilon^\alpha} = P^{\epsilon^\alpha}_{\mathrm{WKB}}(1+o(1))$ uniformly in $\lambda \in C \setminus D^{\epsilon^\alpha}$, since the number of terms in the product is $O(\epsilon)$ and each ratio $\frac{\lambda_n - \lambda_n^{\mathrm{WKB}}}{\lambda}$ is $o(\epsilon)$.

A similar calculation relates the product over eigenvalues that lie in D^{ϵ^α}. The corresponding ratio between the Blaschke products is $(1 + O(\epsilon^{2\alpha-1}))$. We then simply choose $\alpha > 1/2$.

By splitting each product into two products accordingly depending on whether $|\lambda_n^{\mathrm{WKB}}|$ are greater than $G\epsilon^\alpha$ or not, we see that $P = P_{\mathrm{WKB}}(1 + o(1))$. These estimates are pointwise, for any fixed λ in the complex plane. But clearly they are also uniform in any set consisting of the complement of a disc centered in 0 with radius independent of ϵ.

The rest of the argument is the same as for the reflection coefficients. For $\lambda \in C \setminus D$ where D is the small but ϵ-independent disc mentioned above we have a uniform $o(1)$ approximation. Inside the small set D we have the right symmetry and positive definiteness conditions. Then an appropriate local parametrix exists inside D according to the results in the Appendix A of [17], giving rise to an appropriate global approximative solution of the Riemann-Hilbert problem defined in terms of theta functions near $\lambda = \infty$ and thus leading to the formulae stated in the introduction.

□

Remark 6.2. *It can happen (non-generically, for isolated values of ϵ) that the reflection coefficent actually has a pole singularity at* 0*. In other words there is a spectral singularity at* 0*. In such a case one can amend the analysis by considering a very small circle around* 0 *say of radius* $O(\epsilon)$*, and removing the singularity exactly in the same way we have removed the poles due to the eignevalues in* [17]*. The reflection coefficient of course is not analytically extensible in general but one can simply extract the singular part of the reflection coefficent which is of course rational. The main result is not affected.*

In [18] we have studied the energy equilibrium problem that underlies the function g appearing in the change of variables of Chapter 4 in [17] which expresses the finite gap ansatz. We have been able to show that the equilibrium measure μ exists for a particular contour and hence that the right g exists so long as the support of μ does not touch the segment $[0, iA]$ at more than a finite number of points. This is referred to as Assumption (A) in [18].

The next proposition follows.

Proposition 6.3. *Under the assumptions before the statement of Proposition 6.1, under assumption (A) in* [18] *and also the assumption that the density of eigenvalues admits an analytic extension in the upper half-plane, the asymptotics* (1.4) *stated in the introduction are valid.*

Proof. Follows directly from Proposition 6.1 and Theorem 8 in [18] with ϕ given by formula (C.2) in appendix C. □

It has eventually become clear that both the finite gap ansatz stated in [17] and the assumption (A) in [18] are too restrictive. Hints of this inadequacy were already apparent in [17] and the phenomenon was further explored in [22].

In [15] we have added an amendment to [18] showing how to extend the analysis without the assumption (A). Also, in an unpublished preprint [14], reproduced here in the last section of the appendix, we show how to proceed if the assumption of analyticity for the eigenvalue density is not true, by solving an auxiliary scalar Riemann-Hilbert problem. We end up with the following result.

Proposition 6.4. *The asymptotics* (1.4) *stated in the introduction are always valid under initial data that satisfy the assumptions before the statement of Proposition 6.1 above.*

Proof. Follows from Proposition 6.3 and the Theorem in [15] in view of the discussion in Appendix C. □

ACKNOWLEDGEMENT. Research supported by the ARISTEIA II program of the Greek Secretariat of Research and Technology under Grant No. 3964. The second author also acknowledges the generous support of Ritsumeikan University during three visits in 2015-2018.

Appendix A. A Riemann-Hilbert factorisation problem for the focusing nonlinear Schrödinger equation

We first present some elementary facts about the Riemann-Hilbert factorisation formulation of the inverse scattering problem for the focusing nonlinear Schrödinger equation, as described in [17]. We describe first the case of reflectionless data which has been the main concern in [17]. Then we indicate how the problem changes if we allow the reflection coefficient to be non-zero. We will give some proofs for the reflectionless data but we will not do so for the general case, for which we just refer to standard textbooks like [9]. The reason for this is that the part of the proof that transforms the meromorphic problem to a holomorphic Riemann-Hilbert problem is the same in both cases. Meanwhile the contour C (and its complex conjugate) introduced for this purpose plays a role in the proofs of the main text. On the other hand the general proof is more complicated, since it encapsulates the whole scattering and inverse scattering theory, and it is not neceesary for the purposes of this paper.

The focusing nonlinear Schrödinger equation is "completely integrable". Although there is no precise definition of this notion for infinite dimensional dynamical systems, one thing it always entails is the fact that it admits a "Lax pair". In our context this means that, for arbitrary ϵ, it is represented as the compatibility condition for two systems of linear ordinary differential equations:

$$\epsilon\partial_x \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -i\lambda & \psi \\ -\psi^* & i\lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{A.1}$$

$$i\epsilon\partial_t \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \lambda^2 - |\psi|^2/2 & i\lambda\psi - \epsilon\partial_x\psi/2 \\ -i\lambda\psi^* - \epsilon\partial_x\psi^*/2 & -\lambda^2 + |\psi|^2/2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{A.2}$$

where λ is an arbitrary complex parameter.

The N-soliton solutions of the nonlinear Schrödinger equation can be thought of as those complex functions $\psi(x,t)$ for which there exist simultaneous column vector solutions of the linear ODEs above in the particularly simple form:

$$\begin{aligned} \mathbf{u}^+(x,t,\lambda) &= \begin{bmatrix} \displaystyle\sum_{p=0}^{N-1} A_p(x,t)\lambda^p \\ \\ \lambda^N + \displaystyle\sum_{p=0}^{N-1} B_p(x,t)\lambda^p \end{bmatrix} \exp(i(\lambda x + \lambda^2 t)/\epsilon), \\ \mathbf{u}^-(x,t,\lambda) &= \begin{bmatrix} \lambda^N + \displaystyle\sum_{p=0}^{N-1} C_p(x,t)\lambda^p \\ \\ \displaystyle\sum_{p=0}^{N-1} D_p(x,t)\lambda^p \end{bmatrix} \exp(-i(\lambda x + \lambda^2 t)/\epsilon), \end{aligned} \tag{A.3}$$

satisfying the relations

$$\begin{aligned} \mathbf{u}^+(x,t,\lambda_k) &= \gamma_k \mathbf{u}^-(x,t,\lambda_k), \\ -\gamma_k^*\mathbf{u}^+(x,t,\lambda_k^*) &= \mathbf{u}^-(x,t,\lambda_k^*), \qquad k = 1,\ldots,N, \end{aligned} \tag{A.4}$$

for some distinct complex numbers $\lambda_0,\ldots,\lambda_{N-1}$ in the upper half-plane and nonzero complex numbers (not necessarily distinct) $\gamma_0,\ldots,\gamma_{N-1}$. It is easy to check that given the numbers $\{\lambda_k\}$ and $\{\gamma_k\}$, the relations determine the coefficient functions $A_p(x,t)$, $B_p(x,t)$, $C_p(x,t)$ and $D_p(x,t)$ in terms of exponentials via the solution of a square inhomogeneous linear algebraic system. In the classic book of Faddeev and Takhtajan [9] it is shown that this linear system is always nonsingular assuming the $\{\lambda_k\}$ are distinct and nonreal and that the $\{\gamma_k\}$ are nonzero. The solution of the nonlinear Schrödinger equation for which the column vectors $\mathbf{u}^\pm(x,t,\lambda)$ are simultaneous solutions of the linear ODEs turns out to be

$$\psi(x,t) = 2iA_{N-1}(x,t). \tag{A.5}$$

A typical initial condition $A(x)$ will not correspond exactly to a multi-soliton solution. As is well-known ([25], [9]) the procedure generally begins with the study the solutions of the linear ODEs for real λ and for $\psi = A(x)$. One obtains from this analysis a complex-valued *transmission coefficient* $T(\lambda) = 1/a(\lambda)$, $\lambda \in \mathbb{R}$. It turns out that the function $a(\lambda)$ has an analytic continuation into the whole upper half-plane, and its zeros occur at values of λ for which there is an $L^2(\mathbb{R})$ eigenfunction. In this sense, the study of the scattering problem for real λ yields results for complex λ by unique analytic continuation. The function $a(\lambda)$ can be interpreted as a Wronskian between two particular solutions that have analytic continuations into the upper half-plane. Thus at each L^2 eigenvalue λ_k, there is a complex number γ_k that is the ratio of these two analytic solutions. In addition to the transmission coefficient, one also finds a complex-valued function $b(\lambda)$ that gives rise to a *reflection coefficient* $R(\lambda) := b(\lambda)/a(\lambda)$, $\lambda \in \mathbb{R}$. Following Zakharov and Shabat [25] we have:

(1) When $\psi(x,t)$ is the solution of the focusing NLS with initial data $A(x)$, then for each $t > 0$ one has different coefficients in the linear problem, and therefore the eigenvalues $\{\lambda_k\}$, proportionality constants $\{\gamma_k\}$ and the function $b(\lambda)$, can be computed independently for each $t > 0$. However, the eigenvalues $\{\lambda_k\}$ (more generally the function $a(\lambda)$) and also $|b(\lambda)|$, $\lambda \in \mathbb{R}$, are independent of t, and the proportionality constants $\{\gamma_k\}$ and $\arg(b(\lambda))$, $\lambda \in \mathbb{R}$ evolve simply in time. Thus, $R(\lambda,t) = R(\lambda,0)\exp(-2i\lambda^2 t/\epsilon)$ and $\gamma_k(t) = \gamma_k(0)\exp(-2i\lambda_k^2 t/\epsilon)$.
(2) The function $\psi(x,t)$ can be reconstructed at later times $t > 0$ in terms of the discrete spectrum $\{\lambda_k\}$, $\{\gamma_k\}$, and the reflection coefficient $R(\lambda)$.

If for the initial condition $A(x)$ we have $b(\lambda) \equiv 0$, then the step of reconstructing the solution of the initial value problem is essentially what we have already described. Namely, one solves the linear equations for the coefficient $A_{N-1}(x,t)$ and then the solution of the initial value problem is given by (A.5). N turns out to be the number of L^2 eigenvalues for $\psi_0(x)$ in the upper half-plane.

In general, the reconstruction of ψ from the scattering data can be recast in terms of the solution of a matrix-valued meromorphic Riemann-Hilbert problem. One seeks (for each x and t, which play the role of parameters) a matrix-valued function $\mathbf{m}(\lambda)$ of λ that is jointly meromorphic in the upper and lower half-planes and for which

(1) $\mathbf{m}(\lambda) \to \mathbb{I}$ in each half-plane as $\lambda \to \infty$.
(2) The singularities of $\mathbf{m}(\lambda)$ are completely specified. There are simple poles at the eigenvalues $\{\lambda_k\}$ and the complex conjugates with residues of a certain specified type (see below).

(3) On the real axis $\lambda \in \mathbb{R}$, there is the *jump relation*

$$\mathbf{m}_+(\lambda) = \mathbf{m}_-(\lambda)\mathbf{v}(\lambda)\,, \qquad \mathbf{m}_\pm(\lambda) := \lim_{\eta \downarrow 0} \mathbf{m}(\lambda \pm i\eta) \tag{A.6}$$

where $\mathbf{v}(\lambda)$ is a certain *jump matrix* built out of $R(\lambda)$ and depending *explicitly* on x and t (and ϵ). The jump matrix becomes the identity matrix for $b(\lambda) \equiv 0$.

If the boundary values $\mathbf{m}_\pm(\lambda)$ are continuous, and if $b(\lambda) \equiv 0$, then it is easy to see that the solution $\mathbf{m}(\lambda)$ must be a rational function of λ. In [17] this is the only case considered. In the current paper however the jump matrix is non-trivial. In fact

$$\mathbf{v}(\lambda) = \begin{bmatrix} 1 & R(\lambda)\exp\left(\frac{1}{\epsilon}(-2i\lambda x - 2i\lambda^2 t)\right) \\ R^*(\lambda)\exp\left(\frac{1}{\epsilon}(2i\lambda x + 2i\lambda^2 t)\right) & 1 + |R(\lambda)|^2 \end{bmatrix} \tag{A.7}$$

The proof of this fact is much more complicated than the reflectionless case and we omit it here. We refer the reader to the comprehensive discussion in Chapter II of the authoritative textbook of Faddeev and Takhtajan [9] of the full Riemann-Hilbert problem (which they call just "Riemann problem").

Continuing with the pure soliton case of $b(\lambda) \equiv 0$, from the column vectors $\mathbf{u}^\pm(x,t,\lambda)$, we build a matrix solution of (A.1)-(A.2):

$$\Psi(\lambda) := [\mathbf{u}^-(x,t,\lambda), \mathbf{u}^+(x,t,\lambda)]\mathrm{diag}\left(\prod_{j=1}^N (\lambda-\lambda_j)^{-1}, \prod_{j=1}^N (\lambda-\lambda_j^*)^{-1}\right)\exp(i\sigma_3\lambda^2 t/\epsilon)\,. \tag{A.8}$$

This special matrix solution is the familiar Jost solution. If we now define a matrix $\mathbf{m}(\lambda)$ by

$$\mathbf{m}(\lambda) := \Psi(\lambda)\exp(i\sigma_3\lambda x/\epsilon)\,, \tag{A.9}$$

then we find using (A.1)-(A.2) that for all fixed complex λ different from the eigenvalues $\{\lambda_k\}$ and their complex conjugates, $\mathbf{m}(\lambda)$ is a uniformly bounded function of x that satisfies $\mathbf{m}(\lambda) \to \mathbb{I}$ as $x \to +\infty$.

We can deduce from the explicit form of the vectors $\mathbf{u}^\pm(x,t,\lambda)$ and from the relations (A.1)-(A.2) that $\mathbf{m}(\lambda)$ solves the following problem.

Given the discrete data $\{\lambda_k\}$ and $\{\gamma_k\}$, find a matrix $\mathbf{m}(\lambda)$ with the following two properties:

(1) $\mathbf{m}(\lambda)$ is a rational function of λ, with simple poles confined to the eigenvalues $\{\lambda_k\}$ and the complex conjugates. At the singularities:

$$\begin{aligned} \mathop{\mathrm{Res}}_{\lambda=\lambda_k} \mathbf{m}(\lambda) &= \lim_{\lambda\to\lambda_k} \mathbf{m}(\lambda)\begin{bmatrix} 0 & 0 \\ c_k(x,t) & 0 \end{bmatrix}, \\ \mathop{\mathrm{Res}}_{\lambda=\lambda_k^*} \mathbf{m}(\lambda) &= \lim_{\lambda\to\lambda_k^*} \mathbf{m}(\lambda)\begin{bmatrix} 0 & -c_k(x,t)^* \\ 0 & 0 \end{bmatrix}, \end{aligned} \tag{A.10}$$

for $k = 0, \ldots, N-1$, with

$$c_k(x,t) := \left(\frac{1}{\gamma_k}\right) \frac{\prod_{n=0}^{N-1} (\lambda_k - \lambda_n^*)}{\prod_{\substack{n=0 \\ n\neq k}}^{N-1} (\lambda_k - \lambda_n)} \exp(2i(\lambda_k x + \lambda_k^2 t)/\epsilon) \,. \tag{A.11}$$

(2)

$$\mathbf{m}(\lambda) \to \mathbb{I}\,, \qquad \text{as} \qquad \lambda \to \infty\,. \tag{A.12}$$

These two properties actually characterise the matrix function $\mathbf{m}(\lambda)$ uniquely. We have ([17])

Proposition A.1. *The meromorphic Riemann-Hilbert Problem corresponding to the discrete data $\{\lambda_k\}$ and $\{\gamma_k\}$ has a unique solution whenever the λ_k are distinct in the upper half-plane and the γ_k are nonzero. The function defined from the solution by*

$$\psi := 2i \lim_{\lambda\to\infty} \lambda m_{12}(\lambda) \tag{A.13}$$

(that this limit exists is part of the proposition) is a nontrivial N-soliton solution of the focusing nonlinear Schrödinger equation.

For an asymptotic analysis it is useful to convert the meromorphic Riemann-Hilbert problem back into a sectionally holomorphic Riemann-Hilbert problem. This can be easily be done by constructing (for example) small circles around the poles and redifining the unkonwn inside those circles accordingly, see [4]. Here, we proceed as follows.

Let C be a simple closed contour that is the boundary of a simply-connected domain D in the upper half-plane that contains *all* of the eigenvalues $\{\lambda_k\}$. We assign to C a counterclockwise orientation. By C^* and D^* we mean the corresponding complex conjugate sets in the lower half-plane, and we assign both loops the same orientation.

It is not hard to see ([17]) that for our symmetric even data $A(x)$ one has $\gamma_k = (-1)^k$. Still, it has proved convenient in the asymptotic analysis of [17] and [18] to interpolate the proportionality constants as follows. One can easily choose a constant Q (always 1 or -1, but depending on x, t) and a function $X(\lambda)$ analytic in D so that

$$\gamma_k = Q \exp(X(\lambda_k)/\epsilon)\,, \qquad k = 0, \ldots, N-1\,. \tag{A.14}$$

In general, $X(\lambda)$ could be systematically constructed as an interpolating polynomial of degree $\sim N$. In our (symmetric) case the phases γ_k are highly correlated so that for very large N one can easily choose for $X(\lambda)$ a polynomial of low degree or another simple expression. Note that the interpolant of the γ_k is not necessarily unique; for each K in some indexing set (an integer) there is a distinct pair $(Q_K, X_K(\lambda))$ such that for all k, $\gamma_k = Q_K \exp(X_K(\lambda_j)/\epsilon)$.

Remark A.2. *In* [15] *we have made use of this freedom. We have found that the best choice depends on the Riemann surface sheet where our contour is allowed to expand. The issue of the right choice is also related to the improvement of the approximation achieved in* [22].

With the help of the interpolant of the proportionality constants, we define a new matrix $\mathbf{M}(\lambda)$ for $\lambda \in \mathbb{C} \setminus (C \cup C^*)$ in the following way. First, for all $\lambda \in D$, set

$$\mathbf{M}(\lambda) := \mathbf{m}(\lambda) \begin{bmatrix} 1 & 0 \\ -\left(\frac{1}{Q_K}\right)\left(\prod_{n=0}^{N-1} \frac{\lambda - \lambda_n^*}{\lambda - \lambda_n}\right) \exp\left(\frac{1}{\epsilon}(2i\lambda x + 2i\lambda^2 t - X_K(\lambda))\right) & 1 \end{bmatrix} . \tag{A.15}$$

Next, for all $\lambda \in D^*$, set

$$\mathbf{M}(\lambda) := \sigma_2 \mathbf{M}(\lambda^*)^* \sigma_2 \, . \tag{A.16}$$

Finally, for all $\lambda \in \mathbb{C} \setminus (\overline{D} \cup \overline{D}^*)$ (*i.e.* in the rest of the complex plane minus $C \cup C^*$) simply set

$$\mathbf{M}(\lambda) := \mathbf{m}(\lambda) \, . \tag{A.17}$$

It is straightforward to verify that by our choice of interpolants, and the "Blaschke" factor appearing in (A.15), that $\mathbf{M}(\lambda)$ *has no poles in* D *or* D^* *and hence is sectionally holomorphic in the complex* λ *plane.* By definition, we have preserved the reflection symmetry of $\mathbf{m}(\lambda)$ so that for all $\lambda \in \mathbb{C} \setminus (C \cup C^*)$ we have:

$$\mathbf{M}(\lambda^*) = \sigma_2 \mathbf{M}(\lambda)^* \sigma_2 \, . \tag{A.18}$$

The matrix $\mathbf{M}(\lambda)$ has continuous boundary values from either side on C and C^*. To describe these, let the left (respectively right) side of the oriented contour $C \cup C^*$ be denoted by "+" (respectively "−"). For $\lambda \in C \cup C^*$ define

$$\mathbf{M}_\pm(\lambda) := \lim_{\substack{\mu \to \lambda \\ \mu \in \pm \text{ side of } C \cup C^*}} \mathbf{M}(\mu) \, , \tag{A.19}$$

that is, the nontangential limits from the left and right sides. Then, using the fact that $\mathbf{m}(\lambda)$ is analytic on $C \cup C^*$ and the piecewise definition of $\mathbf{M}(\lambda)$ given by (A.15), (A.16), and (A.17), we find

$$\begin{aligned} \mathbf{M}_+(\lambda) &= \mathbf{M}_-(\lambda) \mathbf{v}_\mathbf{M}(\lambda) \, , & \lambda \in C_\omega \, , \\ \mathbf{M}_+(\lambda) &= \mathbf{M}_-(\lambda) \sigma_2 \mathbf{v}_\mathbf{M}(\lambda^*)^* \sigma_2 \, , & \lambda \in C^* \, , \end{aligned} \tag{A.20}$$

where for $\lambda \in C$,

$$\mathbf{v}_{\mathbf{M}}(\lambda) := \begin{bmatrix} 1 & 0 \\ -\omega\left(\frac{1}{Q_K}\right)\left(\prod_{n=0}^{N-1}\frac{\lambda-\lambda_n^*}{\lambda-\lambda_n}\right)\exp\left(\frac{1}{\epsilon}(2i\lambda x + 2i\lambda^2 t - X_K(\lambda))\right) & 1 \end{bmatrix}. \tag{A.21}$$

Now, by defining the discrete measure

$$d\mu = \sum_{k=0}^{N-1}\left[\epsilon\delta_{\lambda_k^*} - \epsilon\delta_{\lambda_k}\right], \tag{A.22}$$

we see that for any branch of the logarithm,

$$\prod_{k=0}^{N-1}\frac{\lambda-\lambda_k^*}{\lambda-\lambda_k} = \exp\left(\frac{1}{\epsilon}\int \log(\lambda-\eta)\, d\mu(\eta)\right). \tag{A.23}$$

In the general case (with a nontrivial reflection coefficient R) the same calculation applies and does not affect the jump across the real line. Suppose then the eigenvalues $\{\lambda_k\}$ and proportionality constants $\{\gamma_k\}$ are given along with an appropriate interpolation $Q_K \exp(X_K(\lambda)/\epsilon)$ of the γ_k and a smooth closed oriented contour C enclosing the eigenvalues in the upper half-plane. Suppose also that R, the reflection coefficient corresponding to the initial data is also given. We define a Riemann-Hilbert problem as follows.

Find a matrix function $\mathbf{M}(\lambda)$ that satisfies:

(1) $\mathbf{M}(\lambda)$ is analytic in each component of $\mathbb{C}\setminus(C\cup C^*)$.
(2) $\mathbf{M}(\lambda)$ assumes continuous boundary values on $C\cup C^*$ and the real line.
(3) The boundary values taken on $C\cup C^*$ satisfy the relations (A.20) with $\mathbf{v}_{\mathbf{M}}(\lambda)$ given explicitly by (A.21). In the case where the reflection coefficient is non-trivial, there is a jump condition across the real line (A.7).
(4) $\mathbf{M}(\lambda)$ is normalized at infinity:

$$\mathbf{M}(\lambda) \to \mathbb{I} \text{ as } \lambda \to \infty. \tag{A.24}$$

Proposition A.3. *The holomorphic Riemann-Hilbert Problem has a unique solution $\mathbf{M}(\lambda)$ whenever the λ_k are distinct and nonreal, and the γ_k are nonzero. The function defined by*

$$\psi := 2i \lim_{\lambda\to\infty} \lambda M_{12}(\lambda), \tag{A.25}$$

is independent of the particular choice of loop contour C and interpolant index K, and is the solution of the focusing nonlinear Schrödinger equation corresponding to the reflection coefficient R and the discrete data $\{\lambda_k\}$ and $\{\gamma_k\}$.

Proof. The proof that the inverse scattering problem is equivalent to a meromoprhic Riemann-Hilbert problem can be found in [9]. The proof that the meromorphic Riemann-Hilbert problem is equivalent to a holomorphic Riemann-Hilbert problem under the appropriate choice of the contour C is in [17] in the reflectionless case only, but it works exactly the same in the case of a non-trivial reflection coefficient. □

As shown in [17] it is possible to allow C to meet the real axis at one or more isolated points $u_k \in \mathbb{R}$, as long as at each u_k the incoming and outgoing parts of C make nonzero angles with the real axis and with each other. The contour C should thus meet the axis in "corners" (if at all). The potential non-triviality of the reflection coefficient is irrelevant here.

Appendix B. Asymptotic analysis via a Riemann-Hilbert deformation

In this section we give a very rudimentary description of the analysis of [17] and the philosophy behind it. It is not really essential for the understanding of this paper and it is no substitute for a detailed reading of [17]. A somewhat more comprehensive review can be found in [16].

In view of A.21 and A.23, the Riemann-Hilbert problem we have to analyse asymptotically can be seen as a (nonlinear) analogue of exponential integrals. While in linear problems where the Fourier integral method can be applied we end up with exponential integrals, here we have a *Riemann−Hilbert problem* with exponential phase.

It was first realised back in 1981 by Alexander Its ([12]) that the long time asymptotics for the solution of the initial value problem to the focusing NLS can be extracted by reducing the Riemann-Hilbert problem to a "model" Riemann-Hilbert problem that can be solved explicitly, exactly as one does in the asymptotic analysis of exponential integrals. (Apparently Its was inspired by work of Jimbo, Miwa and Ueno on the isomonodromy method for Painleve equations [13].) The Riemann-Hilbert problem deformation method has been made rigorous and systematic in later work by Deift and Zhou in 1993 ([5]). The basic ideas of the Deift-Zhou method are:

1. Equivalence of the solvability of a matrix Riemann-Hilbert problem to the invertibility of an associated singular integral operator. Expression of the solution of the matrix Riemann-Hilbert problem as a singular integral involving the inverse of the associated singular integral operator.

The idea goes back to the Georgian school of Mushkelishvilli and provides a nice way to show that under some conditions, small changes in the jump data result in small changes in the solution.

2. Appropriate lower/diagonal/upper factorisations of jump matrices.

3. Introduction and solution of auxiliary scalar problems leading to a conjugation of the original problem by an exponential factor (the "g-function").

The semiclassical problem is more complicated and requires two more ideas.

4. An auxiliary variational problem of "electrostatic type" (going back to work of Lax, Levermore and Venakides in the 1980s on the zero dispersion

KdV problem). Solution of the Euler-Lagrange equations for this problem via theta functions.

5. Search for an optimal contour (selection of a contour of steepest descent) where all of the above can be applied. Deformation from one contour to another. This is a feature appearing only in problems with non-self-adjoint Lax operator, where the spectrum is not necessarily real and the whole deformation procedure is conducted fully in the complex plane.

A more detailed expedition of the above ideas appears in [16]. The rigorous implementation of the whole sequence is done in [17] and [18].

Our first step is to employ a "change of variables" $N(z) = M(z)\exp(\frac{g(z)\sigma_3}{\epsilon})$ which will enable us to reduce the given Riemann-Hilbert problem to one that is easier to handle and which asymptotically will be explicitly solvable.

The function g is given as a logarithmic transform of an "equilibrium" measure on a particular smooth curve: $g(z) = \int \log(z-\eta)d\mu(\eta)$ where $d\mu$ is the equilibrium measure corresponding to a particular external field depending on the parameters x, t and the initial data. The particular curve on which the equilibrium problem is defined is chosen such that it maximises the corresponding equilibrium energy. In other words the equilibrium measure solves a max-min type variational problem.

More precisely, let C be a contour enclosing the eigenvalues in the upper half-plane as described in the previous section. A priori we seek a function satisfying

$$
\begin{aligned}
g(\lambda) \text{ is independent of } \hbar. \\
g(\lambda) \text{ is analytic for } \lambda \in \mathbb{C} \setminus (C \cup C^*). \\
g(\lambda) \to 0 \text{ as } \lambda \to \infty. \\
g(\lambda) \text{ assumes continuous boundary values from both sides of } C \cup C^*, \\
\text{denoted by } g_+(g_-) \text{ on the left (right) of } C \cup C^*. \\
g(\lambda^*) + g(\lambda)^* = 0 \text{ for all } \lambda \in \mathbb{C} \setminus (C \cup C^*).
\end{aligned}
$$

The assumptions above permit us to write g in terms of a measure ρ defined on the contour $C \cup C^*$. Indeed

$$
g(\lambda) = \int_{C \cup C^*} \log(\lambda - \eta)\rho(\eta)d\eta,
$$

for an appropriate definition of the logarithm branch.

Further technical conditions are necessary to ensure that the Riemann-Hilbert deformations required can go through. Such conditions characterise g *and* the contour C. In [17] these are given by conditions (4.20) and (4.31).

In [18] we define g in a somewhat different but equivalent way, in terms of an equilibrium energy problem. For any *given* contour C we choose $\rho(\eta)d\eta$ to be the equilibrium measure for a certain external field depending on x, t and $A(x)$; but eventually we choose C that maximizes the equilibrium energy. The extra technical conditions on g are equivalent to the Euler-Lagrange conditions for the energy equilibrium problem.

The Riemann-Hilbert problem is then asymptotically "deformed" to a "model" problem that can be explicitly solved in terms of theta functions. The model problem is an asymptotic semiclassical approximation of the original one. The semiclassical asymptotics of the focusing NLS problem are thus also recovered via A.25.

Appendix C. On the Analyticity of the Spectral Density

It is essential for the proofs in [17] that the "asymptotic density of eigenvalues" $\rho^0(\eta)$ (see (3.2) of [17]) derived by WKB theory, which is a priori defined in the straight line interval connecting 0 to iA, be analytically extensible to the closed upper half-plane $\mathbb{H}$. In section 5.1.1 of [17] we state this as an assumption and we actually make heavy use of this assumption in the discussion and the proofs that follow.

The main issue is whether the function

$$R^0(\eta) = \int_{x_-(\eta)}^{x_+(\eta)} (A(x)^2 + \eta^2)^{1/2} dx,$$

where the points $x_\pm$ are defined by

$$A(x_\pm(\eta)) = -i\eta,\ 0 < -i\eta < A_0,$$
$$-A_0 < x_-(\eta) < 0 < x_+(\eta) < A_0,$$

admits an analytic extension. We note here that we choose the branch of the square root that is positive for $x_- < x < x_+$.

We will show in this appendix that even if R^0 does not admit an analytic extension in $\mathbb{H}$, the analysis of Chapter 5 (paragraph 5.1.1) in [17] can be amended via the solution of a scalar Riemann-Hilbert problem. In particular, formula (C.1) below is an amendment of formula (5.1.11) of [17] which serves the same useful purpose in the ensuing discussion therein.

Indeed, consider the following scalar additive Riemann-Hilbert problem, with jump on the linear segment $\Sigma = [-iA_0, iA_0]$. Let p be a function analytic in $\mathbb{C} \setminus [-iA_0, iA_0]$, such that

$$p_+(\eta) + p_-(\eta) = \rho_0(\eta) = \frac{dR^0}{d\eta},\ \eta \in (-iA_0, iA_0).$$

More explicitly, let

$$p(\eta) = (A_0^2 + \eta^2)^{1/2} \int_{(-iA_0, iA_0)} \frac{\rho_0(s)}{(A_0^2 + s^2)^{1/2}} \frac{ds}{2\pi i(s-\eta)},\ \eta \in \mathbb{C} \setminus (-iA_0, iA_0).$$

Here $R^0(\eta)$ is extended to the lower half of Σ by the relation $R^0(\eta^*) = R^0(\eta)$. The "+" side is to the left of Σ and the "-" side is to the right of Σ.

Now, the analysis of Chapter 5 in [17] can be amended as follows. First, let's amend the definition of X in Chapter 3, which describes the interpolant of the norming constants. We simply set

$$X(\lambda) = i\pi(2K+1) \int_{\lambda}^{iA_0} (p_+(\eta) + p_-(\eta)) d\eta,$$

for λ in the linear segment $[0, iA_0]$. Then, the discussion of Chapter 5 in [17], in particular from relation (5.4) to (5.8), is amended by substitutitng $\bar{\rho}^\sigma = p - \rho$. More precisely, taking $\sigma = 1$,

$$\int_0^{iA_0} L^0_\eta(\lambda) p_-(\eta) d\eta = \int_{C_I} L^C_{\eta -}(\lambda) p(\eta) d\eta,$$

and similarly, by symmetry,

$$\int_{-iA_0}^0 L^0_\eta(\lambda) p_-(\eta^*)^* \, d\eta = \int_{C_I^*} L^C_{\eta -}(\lambda) p(\eta^*)^* \, d\eta.$$

(Recall here that $L^0_\eta(\lambda) = log(\lambda - \eta)$, with a cut along the imaginary axis from η to $-i\infty$. In the above integral we integrate over the "-" side, while in the integral just following we integrate over the "+" side.) Also

$$\int_0^{iA_0} L^0_\eta(\lambda) p_+(\eta) d\eta = \int_{C_F} L^C_{\eta -}(\lambda) p(\eta) d\eta,$$

and similarly, by symmetry,

$$\int_{-iA_0}^0 L^0_\eta(\lambda) p_+(\eta^*)^* d\eta = \int_{C_F^*} L^C_{\eta -}(\lambda) p(\eta^*)^* d\eta.$$

Next, note that $L^C_{\eta +}(\lambda) = L^C_{\eta -}(\lambda)$ for all $\eta \in C_I \cup C_I^*$ "below" $\lambda \in C_I$ and at the same time $L^C_{\eta +}(\lambda) = 2\pi i + L^C_{\eta -}(\lambda)$ for $\eta \in C_I$ "above" λ. This means that for $\lambda \in C$,

$$\int_C L^C_{\eta \pm}(\lambda) p(\eta) d\eta + \int_{C^*} L^C_{\eta \pm}(\lambda) p(\eta^*)^* d\eta =$$

$$\int_C \overline{L^C_\eta}(\lambda) p(\eta) d\eta + \int_{C^*} \overline{L^C_\eta}(\lambda) p(\eta^*)^* d\eta \pm \pi i/2 \int_{C_I} p(\eta) d\eta \pm \pi i/2 \int_{C_F} p(\eta) d\eta,$$

with $\overline{L^C_\eta}(\lambda) = \frac{L^C_{\eta +}(\lambda) + L^C_{\eta -}(\lambda)}{2}$. Assembling these results gives the expression

$$\tilde{\phi}(\lambda) = \int_C \overline{L^C_\eta}(\lambda) \overline{\rho}(\eta) d\eta + \int_{C^*} \overline{L^C_\eta}(\lambda) \overline{\rho}(\eta^*)^* d\eta$$

$$+ J(2i\lambda x + 2i\lambda^2 t) - (J(2K+1)+1) \left(\pm \pi i/2 \int_{C_I} p(\eta) d\eta \pm \pi i/2 \int_{C_F} p(\eta) d\eta\right),$$

valid for $\lambda \in C$, where we have introduced the complementary density for $\eta \in C : \overline{\rho}(\eta) := p(\eta) - \rho(\eta)$. Choosing K so that $J(2K+1)+1 = 0$, the last term vanishes and we simply have

$$\tilde{\phi}(\lambda) = \int_C \overline{L^{C,\sigma}_\eta}(\lambda) \overline{\rho}(\eta) d\eta + \int_{C^*} \overline{L^C_\eta}(\lambda) \overline{\rho}(\eta^*)^* d\eta + J(2i\lambda x + 2i\lambda^2 t). \quad \text{(C.1)}$$

Comparing with (5.11) of [17] this last formula is less awkward, since it does not depend on the a priori constraint that the contour C has to go through iA, a constraint that is eventually suspended anyway.

The rest of the proofs of [17] go through, with p substituting ρ^0. We omit the detailed discussion, but we *do* stress one major point on the variational problem of Chapter 8 of [17].

The contour C and the measure $\rho d\eta$ are characterized by a solution of a Green's variational problem of electrostatic kind. Indeed

$$E_\phi(\rho d\eta) = \max_{C'} \min_{\mu:\mathrm{supp}(\mu)\in C} E_\phi(\mu),$$

where the contours C' are a priori supported in the upper half-plane minus the linear segment $[0, iA_0]$, and E_ϕ is the weighted energy of a measure with respect to the external field given by

$$\phi(z) = \int \log \frac{|z-\eta^*|}{|z-\eta|}\rho^0(\eta)d\eta - \mathrm{Re}(i\pi J \int_z^{iA_0} p(\eta)d\eta + 2iJ(zx+z^2t)).$$

The harmonicity of ϕ is important to the structure of $C, supp(\rho)$. But again, even if ρ^0 is not analytically extended, it can be written as a sum of two terms that *are*.

One could write ϕ as

$$\phi(z) = \int \log \frac{|z-\eta^*|}{|z-\eta|}(p_+ + p_-)(\eta)d\eta - \mathrm{Re}(i\pi J \int_z^{iA_0} p(\eta)d\eta + 2iJ(zx+z^2t)). \tag{C.2}$$

Again, this representation is perhaps more natural, since in setting the variational problem it is more appropriate to think of the "left" and "right" sides of the linear segment $[0, iA_0]$ as distinct.

Remark C.1. *The moral of the story is that if ρ^0 does not admit an entire extension, we can write it as the average of two functions p_-, p_+ that can be extended to the left and right of the segment $[0, iA_0]$ respectively, and proceed as before, with ρ^0 substituted by p.*

Remark C.2. *In* [18] *we assume that the solution of the variational problem does not touch the spike $[0, iA_0]$ except possibly at a finite number of points. As shown in* [15]*, this obstacle can be overcome by setting the variational problem on an infinite sheeted Riemann surface $\mathbb{L}$. For this, we use the analyticity of ρ^0 even across the spike. Here we don't have that (in fact this is the whole point of this appendix). But a careful examination of* [15] *shows that what we actually need is analyticity across all but one liftings of the spike on $\mathbb{L}$. This we can get by simply setting our scalar Riemann-Hilbert problem on $\mathbb{L}$ and letting the jump be a single copy of the spike $[0, iA]$ in $\mathbb{L}$. The scalar Riemann-Hilbert problem on $\mathbb{L}$ can be explicitly solved by mapping conformally $\mathbb{L}$ to $\mathbb{C}$.*

References

[1] T. Brooke Benjamin, J.E. Feir: *The disintegration of wave trains on deep water. Part 1. Theory*, Journal of Fluid Mechanics, **27 (3)**, pp. 417–430 (1967)

[2] G. Biondini, D. Mantzavinos: *Universal Nature of the Nonlinear Stage of Modulational Instability*, Phys. Rev. Lett., **116**, 043902 (2016)

[3] M. Bertola, A. Tovbis: *Universality for the focusing nonlinear Schrödinger equation at the gradient catastrophe point: Rational breathers and poles of the tritronquée solution to Painlevé I*, Communications in Pure and Applied Mathematics, **66**, n.5, pp. 678–752 (2013)

[4] P. Deift, S. Kamvissis, Thomas Kriecherbauer, Xin Zhou: *The Toda Rarefaction Problem*, Communications in Pure and Applied Mathematics, **49**, n.1, pp. 35–83 (1996)

[5] P. Deift, X. Zhou: *A Steepest Descent Method for Oscillatory Riemann-Hilbert Problems; Asymptotics for the MKdV Equation*, Annals of Mathematics, Second Series, **137**, n.2, pp. 295-368 (1993)

[6] J. Ecalle: *Cinq applications des fonctions résurgentes,* prépublication Orsay 1984.

[7] S. Fujiié, C. Lasser, L. Nédélec: *Semiclassical resonances for a two-level Schrödinger operator with a conical intersection*, Asymptotic Analysis, **65**, n.1-2 (2009), pp. 17-58.

[8] S. Fujiié, T. Ramond: *Matrice de scattering et résonances asociées à une orbite hétérocline*, Ann. Inst. H. Poincaré Phys. Théor., **69**, n.1 (1998), pp. 31-82.

[9] L. Faddeev, L. Takhtajan: *Hamiltonian Methods in the Theory of Solitons,* Springer 1987.

[10] C. Gérard, A. Grigis, : *Precise Estimates of Tunneling and Eigenvalues near a Potential Barrier*, J.Differential Equations, **72** (1988), pp.149-177.

[11] K. Hirota, J. Wittsten : *Complex eigenvalue splitting for the Zakharov-Shabat operator*, arXiv:1909.08739.

[12] A.R. Its : *Asymptotics of Solutions of the Nonlinear Schrödinger Equation and Isomonodromic Deformations of Systems of Linear Differential Equations*, Soviet Mathematics Doklady, **24**, n.3 (1982), pp. 14-18.

[13] M. Jimbo, T. Miwa, K. Ueno: *Monodromy preserving deformation of linear ordinary differential equations with rational coefficients. I. General theory and τ-function*, Physica D: Nonlinear Phenomena, **2**, n.2 (1981), pp. 306-352.

[14] S. Kamvissis: *On the Analyticity of the Spectral Density for Semiclassical NLS,* Max Planck Institute preprint 2002-43 (2002).

[15] S. Kamvissis: *Comment on the article "Existence and Regularity for an Energy Maximization Problem in Two Dimensions" by Spyridon Kamvissis, Evguenii A. Rakhmanov*, Journal of Mathematical Physics, **50** , 104101 (2009)

[16] S. Kamvissis: *From Stationary Phase to Steepest Descent,* Contemporary Mathematics, **458**, AMS 2008, pp.145-162.

[17] S. Kamvissis, Kenneth D. T.-R. McLaughlin, P. D. Miller : *Semiclassical Soliton Ensembles for the Focusing Nonlinear Schrödinger Equation*, Annals of Mathematics, **154** (2003), Princeton University Press, Princeton, NJ.

[18] S. Kamvissis, E. A. Rakhmanov: *Existence and Regularity for an Energy Maximization Problem in Two Dimensions*, Journal of Mathematical Physics, **46** , n.8 (2005)

[19] M. Klaus: *Eigenvalue asymptotics for Zakharov-Shabat systems with long-range potentials*, Operators and Matrices, **12**, n.1 (2018), pp. 55-106; also private communication.

[20] M. Klaus, J. K. Shaw : *Purely imaginary eigenvalues of Zakharov-Shabat systems,* Phys. Rev. E **65**, 036607 (2002)

[21] M. Klaus, J. K. Shaw : *On the eigenvalues of Zakharov-Shabat systems,* SIAM J. Math. Anal., **34**, n.4, pp.759-773 (2003)

[22] G. Lyng, P. D. Miller : *The N-soliton of the focusing nonlinear Schrödinger equation for N large,* Comm. Pure Appl. Math., **60** , pp. 951-1026 (2007)

[23] E. Madelung : *Quantentheorie in Hydrodynamischer Form,* Z. Phys., **40** (3-4), pp. 322 - 326 (1927)

[24] A. Voros : *The return of the quartic oscillator. The complex W.K.B. method*, Ann. Inst. H. Poincaré, **29**, pp. 211-338 (1983)

[25] V. E. Zakharov and A. B. Shabat : *Exact Theory of Two-dimensional Self-focusing and One-dimensional Self-modulation of Wave in Nonlinear Media,* Journal of Experimental and Theoretical Physics **34** n.1, pp. 62-69 (1972).

Department of Mathematical Sciences, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga, 525-8577, Japan

Department of Pure and Applied Mathematics, University of Crete, GR–700 13 Voutes Campus, Greece, and Institute of Applied and Computational Mathematics, FORTH, GR–711 10 Voutes Campus, Greece