STABILITY AND WELL-POSEDNESS IN INTEGRABLE NONLINEAR EVOLUTION EQUATIONS

STABILITY AND WELL-POSEDNESS IN INTEGRABLE EVOLUTION EQUATIONS

By YUSUKE SHIMABUKURO, B.Sc., M.Sc.

A Thesis Submitted to the Department of Mathematics and Statistics and the School of Graduate Studies of McMaster University in Partial Fulfilment of the Requirement for the Degree of

Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2016)	McMaster University
Department of Mathematics and Statistics	Hamilton, Ontario, Canada

TITLE:	Stability and well-posedness in integrable evolution equations
AUTHOR:	Yusuke Shimabukuro B.Sc. (University of Arizona) M.Sc. (University of Arizona)
SUPERVISOR:	Dr. Dmitry Pelinovsky
NUMBER OF PAGES:	IX, 153

Abstract

This dissertation is concerned with analysis of orbital stability of solitary waves and well-posedness of the Cauchy problem in the integrable evolution equations. The analysis is developed by using tools from integrable systems, such as higher-order conserved quantities, Bäcklund transformation, and inverse scattering transform. The main results are obtained for the massive Thirring model, which is an integrable nonlinear Dirac equation, and for the derivative NLS equation. Both equations are related with the same Kaup-Newell spectral problem. Our studies rely on the spectral properties of the Kaup-Newell spectral problem, which convey key information about solution behavior of the nonlinear evolution equations.

Acknowledgements

First and foremost I would like to express my deepest gratitude to my advisor, Professor Dmitry Pelinovsky. It has been truly an exciting journey under his guidance. His vast knowledge, ideas, and insights have inspired and helped me to complete this dissertation. I cannot thank him enough for his constant support and patience throughout these four years of my PhD study. I also owe my gratitude to his family for having us for dinner several times. I always enjoyed the Russian tea time after dinner.

I would like to thank Professor Jean-Pierre Gabardo. He was generous enough to meet me and discuss math problems every week for past three years. It was a wonderful opportunity to learn and explore various areas in analysis. I enjoyed every minute of it and feel very lucky to have met him.

I would like to thank Professor Lia Bronsard. She has given me great guidance in academics and personal matters. Her summer funding greatly helped me during the prime time of my PhD study. She took us to the cherry tree garden that was my first discovery of a little of Japanese in Hamilton. She took us ice skating on Lake Ontario, which was my first experience to walk out onto a lake. She gave me wonderful escapes from studies as well as great encouragement for my studies.

I would like to express my gratitude to Professor Walter Craig. His keen comments and insights when I gave talks (or just when I was in the same seminar room) helped me understand about the subjects better. Also, his course on special topics in mathematical physics during my first semester at McMaster was my thrill, listening about Hamiltonian PDEs, the water wave equations and so on.

My thanks extend to professors at external universities. I would like to thank Professor Tetsu Mizumachi for his hospitality during three weeks of my stay at Hiroshima University. He introduced me to the project on the 2D Benney-Luke equation which I truly enjoy working on. It is an honor to collaborate with him.

I would like to thank Professor Catherine Sulem for introducing us the problem on the derivative NLS. Without our communication with her, one of Chapters in this dissertation would not be present.

I would like to thank Professor Vladimir Zakharov for introducing me to integrable systems.

For personal support during my graduate study, I would like to thank all my colleagues at McMaster for their support, and in particular, Alessandro for fun math discussions. I would like to thank Gleb, Toby, Austin, Diogo, Chengwei, Nadeem and Adrien for their friendship and uncountable good memories from hiking in Arizona's canyons, dancing to country music, staying at Coffee Exchange all night with me, enjoying temporary arts in Art Crawl in Hamilton, playing card games, and having the warmest times at Portuguese bar. I thank Wesley and Peter for being always cheerful housemates.

I am thankful to my oldest friends, Atsushi and Takahiro, in Okinawa who have given me the warmest friendship and support.

My special thanks are to my family, Masao, Hisako, Taeko, and my love, Hannah, for their unconditional love and support.

I would not be here where I am today without all of your support.

To my family and Hannah

Table of Contents

Chapter 1

Introduction

1.1 General Background

The Hamiltonian H with n degrees of freedom is completely integrable in the Liouville sense if there exist n independent first integrals $I_1 = H,\ I_2, \cdots, I_n$ in involution, i.e., $\{I_i, I_j\} = 0$. These integrals are used as new coordinates in which corresponding dynamics is linear in time. Concept of Liouville integrability can be extended to an infinite dimensional Hamiltonian with a countable set of first integrals in involution.

A new theory of completely integrable Hamiltonian systems was stimulated by Gardner, Kruskal and Miura [40] who found that the eigenvalues of the Schrödinger operator

$$L = -\partial_x^2 + u(x,t)$$

are invariant with respect to t if $u(x,t)$ evolves according to the KdV equation

$$u_t - 6uu_x + u_{xxx} = 0. \tag{1.1}$$

Peter Lax [65] formulated a Lax representation of the KdV equation in the form:

$$L_t = [A, L], \tag{1.2}$$

where $[\cdot,\cdot]$ is a Lie bracket and A is a skew symmetric operator which is given by

$$A = -4\frac{d^3}{dx^3} + 6u\frac{d}{dx} + 3u_x.$$

The KdV equation (1.1) is associated with the linear equations defined by the operators L and A,

$$L\phi = \lambda\phi, \quad A\phi = \phi_t.$$

If a spectral parameter λ is independent of space x and time t, a simple computation shows

$$(L\phi)_t = L_t\phi + L\phi_t = (L_t + LA)\phi, \quad (L\phi)_t = (\lambda\phi)_t = \lambda A\phi = AL\phi$$

from which the Lax equation (1.2) is derived.

From the spectral problem $L\phi = \lambda\phi$, solution behavior of the KdV equation can be studied in great detail by the inverse scattering transform [40] where the Gelfand-Levitan-Marchenko equation

$$K(x,y,t) + F(x+y,t) + \int_x^\infty K(x,z,t)F(z+y,t)dz = 0 \tag{1.3}$$

is crucial to express the KdV solution $u(x,t)$ as

$$u(x,t) = -2\frac{d}{dx}K(x,x,t).$$

The inhomogeneous part $F(x)$ distinguishes two important parts of the KdV solution,

$$F(x) = \sum_{n=1}^{N} c_n e^{-\kappa_n x + 8\kappa_n^3 t} + \frac{1}{2\pi}\int_{\mathbb{R}} r(k)e^{8ik^3t+ikx}dk, \tag{1.4}$$

where the first term is related to N solitons and the second term is related to dispersive wave packets. The explicit pure N-solitons are derived by setting $r(k) = 0$ in (1.4).

The study of linear differential equations dates back to the nineteenth century when Sturm and Louville studied spectral property for the second-order ordinary differential equations. At the same time, transformation methods, related to linear and nonlinear equations, were investigated by Darboux and Backlund. For example, Darboux [25] showed that the linear equation

$$y'' = my + f(x)y, \quad m = \text{constant} \tag{1.5}$$

is related to another linear equation

$$w'' = mw + \theta\frac{d^2}{dx^2}\left(\frac{1}{\theta}\right)w \tag{1.6}$$

through $w = y' - \frac{\theta'}{\theta}y$, where $\theta'' = f(x)\theta$. The potentials in (1.5) and (1.6) are related by

$$f(x) \mapsto \theta\frac{d^2}{dx^2}\left(\frac{1}{\theta}\right),$$

whereas the structure of the linear equations (1.5) and (1.6) are invariant.

The same idea was applied to the linear spectral problems whose potentials correspond to solutions of nonlinear PDEs. Thanks to Zakharov and Shabat [118], the cubic NLS equation

$$iu_t + u_{xx} + 2|u|^2u = 0 \tag{1.7}$$

can be associated with the linear systems

$$\partial_x \phi = \begin{bmatrix} \lambda & u \\ -\overline{u} & -\lambda \end{bmatrix} \phi, \quad \partial_t \phi = i \begin{bmatrix} 2\lambda^2 + |u|^2 & \partial_x u + 2\lambda u \\ \partial_x \overline{u} - 2\lambda \overline{u} & -2\lambda^2 - |u|^2 \end{bmatrix} \phi.$$

If λ is independent of x and t, $u(x,t)$ must be the solution of the cubic NLS equation (1.7) that can be represented as $\partial_x \partial_t \phi = \partial_t \partial_x \phi$. There exists a transformation [79] for the above linear systems whose structure remains invariant and the potential u is transformed as

$$u \mapsto -u - \frac{4\mathrm{Re}(\lambda_0)\phi_1\overline{\phi_2}}{|\phi_1|^2 + |\phi_2|^2}, \tag{1.8}$$

where $\phi = (\phi_1, \phi_2)^t$ is a solution of the same linear systems for a particular value λ_0. If one starts with the zero solution $u = 0$ with $k = 2\lambda_0 \in \mathbb{R}$, one obtains a pure one soliton $u = k\,\mathrm{sech}(kx)e^{ik^2t}$. i.e.,

$$0 \mapsto k\,\mathrm{sech}(kx)e^{ik^2t}.$$

The transformation (1.8) can be iterated to generate, for example, N solitons:

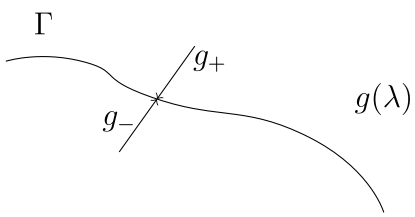
$$0 \mapsto 1 \text{ soliton} \mapsto 2 \text{ solitons} \mapsto 3 \text{ solitons} \mapsto \cdots \mapsto \text{N solitons}.$$

The procedure can be made more efficient by transforming zero solution to N soliton at once [96], called the n-fold Bäcklund/Darboux transformation, due to the fact that permuting the order of inserting solitons does not affect the final N soliton state. This is called Bianchi's permutability.

In the construction of the inverse scattering transform, the solution to a completely integrable system is expressed in terms of the one to the Riemann-Hilbert problem. In its simplest case, the Riemann-Hilbert problem is set to find sectionally analytic function $g(\lambda)$ in $\mathbb{C} \setminus \Gamma$ satisfying the jump condition

$$g_+(\lambda) + \alpha(\lambda)g_-(\lambda) = \beta(\lambda), \quad \lambda \in \Gamma$$

for a given contour Γ in $\mathbb{C}$, and given functions α, β on Γ. The functions g_+ and g_- are the non-tangential limits of g from the two sides of the contour Γ. The contour Γ can be closed or open, bounded or unbounded, as in the figure below:



A function $g(\lambda)$ is closely related to fundamental solutions of the Lax system with a spectral parameter λ. This beautiful aspect of complex analysis, seen in the inverse scattering transform, implicates connection to other branches of science that can be formulated through the Riemann-Hilbert problem. Random matrix theory, for example, has been known for its connection to the Riemann-Hilbert problem and the integrable systems [28]. A hermitian $N \times N$ random matrix, as an example, gives a probability distribution ρ of eigenvalues λ in the form of a Vandermonde determinant with a Gaussian weight and the normalized constant c_n

$$\rho = c_n e^{-\sum_{j=1}^{n} \lambda_j^2} \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j)^2$$

which can be re-expressed as orthogonal polynomials, such as the Hermite polynomials. The orthogonal polynomials can be formulated in the Riemann-Hilbert problem [39]. An eigenvalue behavior as limit $n \to \infty$ exhibits mysterious connection to the integrable systems, for example, to the fifth Painlevé equation [51].

Deift and Zhou discovered an advantage of the Riemann-Hilbert formulation for analytical treatment of integrable PDEs. For example, they developed the steepest decent method to study decay estimate of an oscillatory solution [30]. This led to a number of applications, in particular, to decay estimates in integrable PDEs as well as to orthogonal polynomials. More recently, this type of technique was extended to studying stability problem. Pelinovsky and Cuccagna have studied an asymptotic stability of the NLS soliton using the steepest decent method [24]. Along the same line, the Miura transformation [72], the dressing method [22], and the Bäcklund transformation [46, 79], just to list a few, have been used to treat the stability of solitons in the integrable systems.

This dissertation intertwines analysis of PDEs and beautiful methods from integrable systems. The main goal is to construct a mathematical proof for orbital stability of solitary waves and well-posedness of the Cauchy problem associated to integrable PDEs. It presents novel ways to treat solution of completely integrable systems in a defined function space. The corresponding results are formed in Chapters 2, 3, and 4.

Chapter 5 is concerned with line soliton in the 2D Dirac system. Chapter 5 shows that line soliton of the 2D Dirac system which corresponds to exactly one soliton of the 1D Dirac system is not spectrally stable with respect to transverse perturbations. This adds the first instability result of the Dirac line soliton with respect to transverse perturbations. Instability of line soliton is common in many equations reported in literature, since instability behavior is geometrically richer in 2D than in 1D.

The following sections overview results and techniques obtained in this dissertation and explained in details in the subsequent chapters.

1.2 Orbital stability of Dirac soliton by energy method (Chapter 2)

Chapter 2 is based on our published paper:
D. E. Pelinovsky and Y. Shimabukuro, *Orbital stability of Dirac solitons*, Lett. Math. Phys. **104** (2014), 21-41.

Here, the massive Thirring model (MTM) is an integrable version of the nonlinear Dirac equations written in the form,

$$\begin{cases} i(u_t + u_x) + v = 2|v|^2 u, \\ i(v_t - v_x) + u = 2|u|^2 v. \end{cases} \tag{1.9}$$

Solution of the MTM is denoted as $\mathbf{s}(t) = (u, v)^t$ with an initial data $\mathbf{s}_0 := \mathbf{s}(0)$. The explicit Schwartz function

$$s_\omega(x) := \frac{\sqrt{1-\omega^2}}{\sqrt{1+\omega}\cosh(\sqrt{1-\omega^2}x) + i\sqrt{1-\omega}\sinh(\sqrt{1-\omega^2}x)}$$

for $\omega \in (-1, 1)$ gives the stationary one soliton $\mathbf{s}_\omega e^{i\omega t}$ which solves (1.9), where $\mathbf{s}_\omega = (s_\omega, \overline{s}_\omega)^t$. The *orbit* of Dirac solitons is defined as, for a fixed $\omega \in (-1, 1)$,

$$\Sigma_\omega(t) := \{\mathbf{s}_\omega(\cdot + x_0)e^{i\omega t + i\alpha} | (x_0, \alpha) \in \mathbb{R}^2\},$$

for every $t \in \mathbb{R}$.

Definition 1. *Fix some $\omega \in (-1, 1)$. We say that the orbit Σ_ω is stable in a Hilbert space X if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that, if $\text{dist}_X(\mathbf{s}_0, \Sigma_\omega(0)) < \delta$, then $\text{dist}_X(\mathbf{s}(t), \Sigma_\omega(t)) < \epsilon$ for every $t \in \mathbb{R}$.*

The distance metric dist_X is defined as $\text{dist}_X(f, \Sigma_\omega(t)) := \inf_{g\in\Sigma_\omega(t)} \|f - g\|_X$ for some Hilbert space X equipped with the norm $\|\cdot\|_X$. We essentially use Grillakis-Shatah-Strauss orbital stability theory [41], which says that the orbital stability of Σ_ω holds if $\mathbf{s}_\omega$ is a local minimizer of the energy functional that is constant with respect to the time evolution of the MTM, under some constraint. However, the Dirac Hamiltonian H

$$H = \frac{i}{2}\int_{\mathbb{R}} (u\bar{u}_x - u_x\bar{u} - v\bar{v}_x + v_x\bar{v})\, dx + \int_{\mathbb{R}} \left(-v\bar{u} - u\bar{v} + 2|u|^2|v|^2\right) dx.$$

is sign-indefinite, i.e., there exist infinite-dimensional subspaces $Y_\pm \subset H^{1/2}(\mathbb{R})$ such that

$$H(\mathbf{s} + \mathbf{s}_\omega) - H(\mathbf{s}_\omega) \gtrless 0$$

for every $\mathbf{s} \in Y_\pm$. Therefore, the Dirac Hamiltonian H is not suitable for every $\mathbf{s} \in Y_\pm$ to prove orbital stability of $\mathbf{s}_\omega$. Nevertheless, the MTM is an integrable nonlinear PDE, which possesses arbitrarily many conserved quantities. The main idea is to find a higher conserved quantity on $H^1(\mathbb{R})$ that has a coercive structure. Section 2.5 gives derivation of the conserved quantity R given as

$$\begin{aligned} R(\mathbf{s}) = \textstyle\int_{\mathbb{R}} \big[&|u_x|^2 + |v_x|^2 - \tfrac{i}{2}(u_x\bar{u} - \bar{u}_x u)(|u|^2 + 2|v|^2) + \tfrac{i}{2}(v_x\bar{v} - \bar{v}_x v)(2|u|^2 + |v|^2) \\ &-(u\bar{v} + \bar{u}v)(|u|^2 + |v|^2) + 2|u|^2|v|^2(|u|^2 + |v|^2)\big]\, dx \end{aligned}$$

defined on $H^1(\mathbb{R})$. This energy functional exhibits much nicer structure due to $|u_x|^2 + |v_x|^2$ which gives an elliptic operator. This term is not present in the Dirac Hamiltonian H. The key ingredients of orbital stability of one soliton in the MTM is that one soliton $\mathbf{s}_\omega e^{i\omega t}$ is a critical point and a local minimizer of the functional

$$\Lambda(\mathbf{s}) := R(\mathbf{s}) + (1-\omega^2)\|\mathbf{s}\|_{L^2}$$

under some constraint. The coercivity of R is given in Lemmas 3, 4, 5, depending on different values of ω. Coercivity is used to provide a global bound on $\text{dist}_{H^1}(\mathbf{s}(t), \Sigma(t))$ in the time evolution of the MTM under constraint of either fixed mass or momentum, given in Lemma 7. This leads to proving orbital stability (Theorem 1) by a contradiction argument, given in the end of Section 2.4.

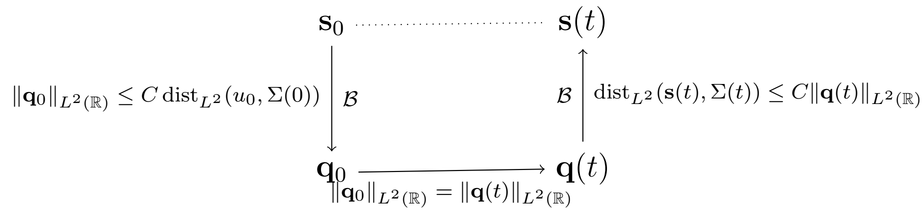
1.3 Orbital stability of Dirac soliton by Bäcklund transform (Chapter 3)

Chapter 3 is based on our published paper:
A. Contreras, D. E. Pelinovsky, and Y. Shimabukuro, L^2 *orbital stability of Dirac solitons in the massive Thirring model*, Comm. PDEs **41** (2016), 227-255

Here, orbital stability of one soliton in the MTM is considered by using the Bäcklund transformation. The transformation can be used to relate N soliton solution and $(N-1)$ soliton solution of the same equation.

The underlying idea of this Chapter is to relate stability of solution around one soliton to stability of solution around zero. Solutions to MTM are stable in L^2 norm, thanks to the mass conservation and L^2 global well-posedness [15].

Let the Bäcklund transformation be denoted as $\mathcal{B}$. If $\mathbf{s} := (u, v)$ is a MTM solution, then $\mathbf{q} = \mathcal{B}[\mathbf{s}]$ is again a solution of the MTM. It is schematically clear that if $\text{dist}_{L^2}(\mathbf{s}_0, \Sigma(0))$ is sufficiently small and $\mathbf{s} = (u, v)$ is a solution to (1.9) with initial data $\mathbf{s}_0 = (u_0, v_0)$, then there is a constant $C > 0$ such that



The main idea presented above is that the global bound on $\text{dist}_{L^2}(\mathbf{s}(t), \Sigma(t))$ comes from the L^2 conservation of $\mathbf{q}$, whose size is controlled by the initial condition.

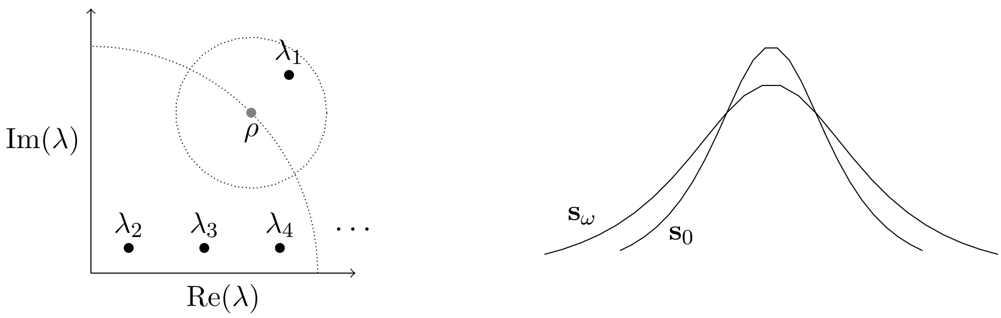
The Bäcklund transformation $\mathcal{B}$, spectrally speaking, removes or adds an eigenvalue of the spectral problem $\partial_x \phi = L_{MTM}\phi$ with the Lax operator L_{MTM}

$$L_{MTM} = \frac{i}{4}(|u|^2 - |v|^2)\sigma_3 - \frac{i\lambda}{2}\sigma_1 W(v) + \frac{i}{2\lambda}\sigma_1 W(u) + \frac{i}{4}\left(\lambda^2 - \frac{1}{\lambda^2}\right)\sigma_3, \quad (1.10)$$

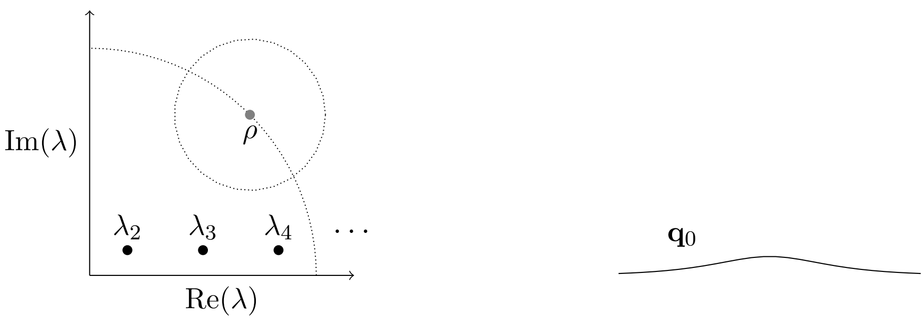
where $W(f) = \begin{pmatrix} f & 0 \\ 0 & \bar{f} \end{pmatrix}$, and σ_1, σ_3 are the Pauli matrices.

Let $\{\lambda_j\}_{j=1}^N$ be a set of eigenvalues of L_{MTM} with the potential $\mathbf{s}_0 = (u_0, v_0)$.

Let ρ be an eigenvalue of L_{MTM} with the pure one soliton $\mathbf{s}_\omega$. If $\text{dist}_{L^2}(\mathbf{s}_0, \Sigma(0))$ is sufficiently small, then the first step is to locate a unique eigenvalue λ_1 that is close to ρ.



The eigenvalue λ_1 contributes to the largest soliton in the initial data $\mathbf{s}_0$. After this eigenvalue is removed, the eigenvalue picture of L_{MTM} with a potential $\mathbf{q}_0 := \mathcal{B}[\mathbf{s}_0]$ may look like



Possible eigenvalues $\lambda_2, \lambda_3, \cdots$ contribute as *small solitary waves.* These eigenvalues do not affect orbital stability theory. If, on the other hand, one asks for asymptotic stability of a soliton, it is important to rule out all eigenvalues in $\mathbf{q}_0$ and to obtain dispersive estimates of the remaining wave packet in a suitable norm.

1.4 Global well-posedness of the derivative NLS equation (Chapter 4)

Chapter 4 is based on our submitted paper:
D. E. Pelinovsky and Y. Shimabukuro, *Existence of global solutions to the derivative NLS equation with the inverse scattering transform method*, arXiv:1602.02118

The Cauchy problem of the derivative NLS equation is given as

$$\left\{ \begin{array}{l} iu_t + u_{xx} + i(|u|^2 u)_x = 0, \quad t > 0, \ x \in \mathbb{R}, \\ u|_{t=0} = u_0. \end{array} \right. \tag{1.11}$$

An interesting open problem concerns with global well-posedness of the Cauchy problem (1.11) with a large initial data u_0, see introduction of Chapter 4.

Definition 2. *The Cauchy problem* (1.11) *is globally well-posed in a Banach space X if a solution $u(t) \in X$ is unique and depends continuously with respect to $u_0 \in X$ for every $t \in [0, \infty)$. We say that a solution map $X \ni u_0 \mapsto u(t) \in X$ is globally well-posed.*

Global well-posedness comes naturally from a local well-posedness in X and a global bound on solution in X. However, for the derivative NLS, in order to obtain such uniform bound by its energy, smallness condition on an initial data u_0 in L^2-norm must be imposed. The inverse scattering transform, instead, constructs a global solution map in Definition 2 without taking use of conservation laws. Chapter 4 is devoted to proving solvability of (1.11) by the inverse scattering transform $\mathcal{R}$ that is bijective and Lipschitz between weighted Sobolev spaces. A global solution map is obtained through the following sequence of maps:

$$u_0 \mapsto \mathcal{R}(u_0) \mapsto \mathcal{R}(u_0)e^{2i\lambda^4 t} \mapsto \mathcal{R}^{-1}(\mathcal{R}(u_0)e^{2i\lambda^4 t}) = u(t)$$

where an important assumption on u_0 is that the spectrum problem of $\partial_x \phi = L_{dNLS}\phi$ with the Lax operator L_{dNLS}

$$L_{dNLS} = -i\lambda^2\sigma_3 + \lambda\sigma_1 \begin{pmatrix} \overline{u}_0 & 0 \\ 0 & u_0 \end{pmatrix} \tag{1.12}$$

does not admit any eigenvalue, i.e., u_0 does not support any soliton. While a sufficiently small initial data satisfies such condition, it is not yet known if large initial data satisfy this condition.

In order to conclude the global well-posedness of the derivative NLS with a large initial data, the case of N solition solution must be considered. We do not include this in this dissertation, due to the lengthy algebraic computations. In fact, once pure dispersion case is established, then adding solitons is more like an algebraic operation. The n-fold Darboux transformations for the Kaup-Newell spectral problem are found in [50, 100]. An excellent exposition of deriving various families of solutions is given in [114].

Let us denote X_N as a function space such that L_{dNLS} admits N simple eigenvalues in the first quadrant. Combining with the result from Chapter 4, global well-posedness of the derivative NLS would follow from the following scheme:

$$\begin{array}{ccc} X_N \ni u_0 & \dashrightarrow & u \in X_N \\ \text{Darboux} \downarrow & & \uparrow \text{Darboux} \\ X_0 \ni u_0 & \xrightarrow[\text{Globally wellposed via IST}]{} & u \in X_0 \end{array}$$

This scheme defines a global solution map $X_N \ni u_0 \mapsto u(t) \in X_N$, which supports N solitons. Such initial data is taken to be arbitraly large. Continuity and bijectivity of the above map are, yet, to be proven on defined function spaces.

1.5 Transverse instability of Dirac line soliton (Chapter 5)

Chapter 5 is based on our published paper:
D. Pelinovsky and Y. Shimabukuro, *Transverse instability of line solitary waves in massive Dirac equations*, J. Nonlinear. Sci. **26** (2016), 365-403

This problem arises in the Gross-Pitaevskii equation with a periodic potential $V(x, y)$, i.e.,

$$i\psi_t = -\Delta\psi + V(x, y)\psi + N(\psi) \quad (x, y) \in \mathbb{R}^2, \tag{1.13}$$

where $N(\psi)$ is a nonlinear term. When $V(x, y)$ is periodic in (x, y), a solution in the equation (1.13) with $N(\psi) = 0$ is expressed in terms of Bloch functions by the Floquet theory. If a nonlinear term is considered, one may make an ansatz where coefficients of Bloch functions now vary in space and time. Direct subsitution may lead to finding that these coefficients satisfy evolution equations of the Dirac type.

Chapter 5 presents the 2D massive Thirring model and the 2D Gross-Neveu model that both can be formally derived from the Gross-Pitaeviskii equation (1.13). The former is the case of the waveguide grating, e.g., $V(x, y) = \epsilon\cos(x)$, and the latter is the case of the honeycomb lattice. Here, line soliton is considered to be a trapped wave in waveguides. In Chapter 5, line soliton is defined to be exactly one soliton solution for 1D case which is independent of y and decays exponentially in x. Due to the fact that line soliton is independent of y, the eigenvalue problem, after linearization around line soliton and the Fourier transform in y, takes the form of

$$i\lambda\mathbf{F} = (Dirac_x + Potential_x + Parameter_p)\mathbf{F},$$

where $Dirac_x$ is the linear Dirac operator in x, $Potential_x$ is the potential term in x, and $Parameter_p$ contains a Fourier variable p that comes from y-derivatives ∂_y, i.e., $\hat{f}(p) = \int_{\mathbb{R}} f(y)e^{ipy}dy$. The first two terms correspond to exactly the 1D case. We prove that for small $|p| > 0$ in $Parameter_p$, the eigenvalue problem above is spectrally unstable, i.e., there exists some $\lambda \in \mathbb{C}$ with an L^2 eigenvector $\mathbf{F}$ such that $\mathrm{Re}\lambda > 0$. The proof is based on locating an unstable bifurcation of zero eigenvalue. Our result indicates spectral instability of line solitons in the limit of long-period transverse perturbations, since a small number $|p|$ corresponds to long-periodicity. For a larger value of $|p|$, we give numerical results, which indicate:

- Spectral instability of the MTM line soliton persists for all transverse wave number p
- Spectral instability of the Gross-Neveu line soliton occurs only in a finite interval of transverse wave number p.

The latter observation is particularly interesting due to the possibility that spectral stability could be observed in a narrow wavequide in the y-direction.

Chapter 2

Orbital Stability of Dirac Soliton by Energy Method

2.1 MTM orbital stability result

We consider the massive Thirring model (MTM)

$$\begin{cases} i(u_t + u_x) + v = 2|v|^2 u, \\ i(v_t - v_x) + u = 2|u|^2 v, \end{cases} \tag{2.1}$$

where $(u, v)(x, t) : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{C}^2$. We denote an initial condition $(u, v)|_{t=0} = (u_0, v_0)$. It has been proven that the MTM is globally well-posed with $(u_0, v_0) \in H^m(\mathbb{R})$ for an integer $m \geq 0$ [15]. The stationary MTM solitons are known in the exact analytical form:

$$\begin{cases} u = U_\omega(x + x_0) e^{i\omega t + i\alpha}, \\ v = \bar{U}_\omega(x + x_0) e^{i\omega t + i\alpha}, \end{cases} \tag{2.2}$$

with

$$U_\omega(x) = \frac{\sqrt{1-\omega^2}}{\sqrt{1+\omega}\cosh\left(\sqrt{1-\omega^2}x\right) + i\sqrt{1-\omega}\sinh\left(\sqrt{1-\omega^2}x\right)}, \tag{2.3}$$

where α and x_0 are real parameters related to the gauge and space translations, whereas $\omega \in (-1, 1)$ is a parameter that determines the frequency of the MTM solitons inside the gap between two branches of the continuous spectrum of the linearized problem at the zero solution. For the MTM (2.1), three conserved quantities are referred to as the charge Q, momentum P, and Hamiltonian H functionals:

$$Q = \int_{\mathbb{R}} \left(|u|^2 + |v|^2\right) dx,$$

$$P = \frac{i}{2} \int_{\mathbb{R}} (u\bar{u}_x - u_x\bar{u} + v\bar{v}_x - v_x\bar{v})\, dx,$$

and

$$H = \frac{i}{2} \int_{\mathbb{R}} (u\bar{u}_x - u_x\bar{u} - v\bar{v}_x + v_x\bar{v})\, dx + \int_{\mathbb{R}} \left(-v\bar{u} - u\bar{v} + 2|u|^2|v|^2\right) dx.$$

To work with a vector function $\mathbf{u} = (u, v, \overline{u}, \overline{v})^t$, we shall work in the function space $X = H^1(\mathbb{R}) \times H^1(\mathbb{R}) \times H^1(\mathbb{R}) \times H^1(\mathbb{R})$, equipped with the inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle_X = \int_{\mathbb{R}} (\mathbf{u}_x \cdot \overline{\mathbf{v}}_x + \mathbf{u} \cdot \overline{\mathbf{v}}) dx,$$

where $\mathbf{u}, \mathbf{v} \in X$ are four component vector, and $\mathbf{u} \cdot \mathbf{v}$ denotes the usual dot product. We define the norm on X as

$$\|\mathbf{u}\|_X = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_X}.$$

We denote the L^2 inner product as

$$\langle \mathbf{u}, \mathbf{v} \rangle_{L^2} = \int_{\mathbb{R}} \mathbf{u} \cdot \overline{\mathbf{v}} dx.$$

We use a notation $T(\theta, s)$ for a two-parameter group of unitary operators on X for each $(\theta, s) \in \mathbb{R}^2$, i.e., $T(\theta, s)\mathbf{u}$ for $\mathbf{u} \in X$ is defined as

$$T(\theta, s)\mathbf{u}(x) := (e^{i\theta} u_1(x+s), e^{i\theta} u_2(x+s), e^{-i\theta} u_3(x+s), e^{-i\theta} u_4(x+s))^t.$$

For a fixed $\omega \in (-1, 1)$, we shall introduce the orbit

$$\{T(\theta, s)\mathbf{u}_\omega : (\theta, s) \in \mathbb{R}^2\}$$

and a small neighborhood around the orbit

$$\Phi_\epsilon = \{\mathbf{u} \in X : \inf_{(\theta,s)\in\mathbb{R}^2} \|\mathbf{u} - T(\theta, s)\mathbf{u}_\omega\|_X < \epsilon\}.$$

From now, $\mathbf{u}$ denotes a solution to the MTM subject to natural constraint in the last two components of the vector, i.e., $\mathbf{u} = (u, v, \overline{u}, \overline{v})^t$. The inner product $\langle \mathbf{u}, \mathbf{v} \rangle_{L^2}$ is always real. The following Theorem presents the main result of this Chapter.

Theorem 1. *There is $\omega_0 \in (0, 1]$ such that for any $\omega \in (-\omega_0, \omega_0)$ and any $\epsilon > 0$ there exists $\delta > 0$ such that if $\|\mathbf{u}_0 - \mathbf{u}_\omega\|_X < \delta$, then the corresponding MTM solution $\mathbf{u}(t)$ satisfies $\mathbf{u}(t) \in \Phi_\epsilon$ for every $t \in \mathbb{R}$.*

Here, one can easily construct a good candidate of a functional used for orbital stability theory, that is,

$$E(\mathbf{u}) = H(\mathbf{u}) - \omega Q(\mathbf{u}),$$

which satisfies $E'(\mathbf{u}_\omega) = 0$, where a functional derivative $E'(\mathbf{u})$ is determined from the Fréchet derivative:

$$\forall \mathbf{v} \in X : \quad \frac{d}{d\epsilon} E(\mathbf{u} + \epsilon \mathbf{v}) \Big|_{\epsilon=0} = \langle E'(\mathbf{u}), \mathbf{v} \rangle_{L^2}. \tag{2.4}$$

Critical points of $E'(\mathbf{u}) = 0$ satisfy the system of first-order differential equations

$$\begin{cases} +i\frac{du}{dx} - \omega u + v = 2|v|^2 u, \\ -i\frac{dv}{dx} - \omega v + u = 2|u|^2 v. \end{cases} \tag{2.5}$$

The stationary MTM solitons (2.2) correspond to the reduction $u = U_\omega$ and $v = \bar{U}_\omega$, where U_ω is a solution of the first-order differential equation

$$i\frac{dU}{dx} - \omega U + \bar{U} = 2|U|^2 U. \tag{2.6}$$

Using definition (2.4) above, we shall write the Taylor expansion of the functional $E(\mathbf{u})$ around $\mathbf{u}_\omega$:

$$E(\mathbf{u}_\omega + \mathbf{u}) = E(\mathbf{u}_\omega) + \langle E'(\mathbf{u}_\omega), \mathbf{u}\rangle_{L^2} + \frac{1}{2}\langle H_\omega \mathbf{u}, \mathbf{u}\rangle_{L^2} + \cdots , \tag{2.7}$$

where $\langle E'(\mathbf{u}_\omega), \mathbf{u}\rangle_{L^2} = 0$ and

$$H_\omega = D_\omega + W_\omega,$$

where

$$D_\omega = \begin{bmatrix} -i\partial_x + \omega & -1 & 0 & 0 \\ -1 & i\partial_x + \omega & 0 & 0 \\ 0 & 0 & i\partial_x + \omega & -1 \\ 0 & 0 & -1 & -i\partial_x + \omega \end{bmatrix}, \quad W_\omega = 4\begin{bmatrix} |v|^2 & \overline{u}v & 0 & \overline{vu} \\ u\overline{v} & |u|^2 & \overline{uv} & 0 \\ 0 & vu & |v|^2 & \overline{v}u \\ uv & 0 & \overline{u}v & |u|^2 \end{bmatrix}$$

where entries of the 4×4 matrix W_ω are all smooth and rapidly decaying at infinity.

The essential spectrum of the non-potential part D_ω coincides with the one of H_ω. As a consequence of Weyl Theorem, see [95, Theorem XIII.14] and [83, B.15], one can show this explicitly by constructing an approximating sequence as given below. We denote $\lambda = \omega - \sqrt{1+k^2} < 0$ for $k \in \mathbb{R}$. We introduce the following sequence:

$$\psi_{n,k} = n^{-1/2}\phi_k\left(\frac{x}{n}\right)e^{ikx}(1, k+\sqrt{1+k^2}, 1, -k+\sqrt{1+k^2})^t,$$

where some smooth and rapidly decaying function $\phi_k(x)$ is suitably normalized so that $\|\psi_{n,k}\|_{L^2} = 1$ for every $n \in \mathbb{N}$ for each fixed $k \in \mathbb{R}$. Using this sequence, we can show that

$$\lim_{n\to\infty} \|(H_\omega - \lambda I)\psi_{n,k}\|_{L^2} = 0 \quad \forall k \in \mathbb{R}.$$

One can show the same result, in the similar way, $\lambda = \omega + \sqrt{1+k^2} > 0$, $k \in \mathbb{R}$.

It follows from the result above that the essential spectra of H_ω is unbounded both above and below:

$$\sigma_{ess}(H_\omega) = \mathbb{R} \setminus (-1+\omega, 1+\omega). \tag{2.8}$$

The essential spectrum (2.8) signifies the sign-infinite property of the energy functional $E(\mathbf{u})$ since the Taylor expansion (2.7) gives

$$E(\mathbf{u}_\omega + \mathbf{u}) - E(\mathbf{u}) = \frac{1}{2}\langle H_\omega \mathbf{u}, \mathbf{u}\rangle_{L^2} + \cdots ,$$

where the sign of the difference depends on the spectral property of the Hessian operatoar H_ω.

In order to use the Grillakis-Shatah-Strauss theory [41], it is necessary to have the condition that the negative spectrum of H_ω is finite, which is not our case.

We shall consider the higher conserved quantity, denoted as R, whose derivation is given in Section 2.5:

$$R = \int_{\mathbb{R}} \Big[|u_x|^2 + |v_x|^2 - \frac{i}{2}(u_x\overline{u} - \overline{u}_x u)(|u|^2 + 2|v|^2) + \frac{i}{2}(v_x\overline{v} - \overline{v}_x v)(2|u|^2 + |v|^2) \\ -(u\overline{v} + \overline{u}v)(|u|^2 + |v|^2) + 2|u|^2|v|^2(|u|^2 + |v|^2) \Big] \, dx. \tag{2.9}$$

With the new energy functional $\Lambda(\mathbf{u}) := R(\mathbf{u}) + \Omega Q(\mathbf{u})$, $\Omega \in \mathbb{R}$, we consider the critical point, $\Lambda'(\mathbf{u}) = 0$, whose first two components are given by

$$\frac{d^2u}{dx^2} + 2i(|u|^2 + |v|^2)\frac{du}{dx} + 2iuv\frac{d\overline{v}}{dx} - 2|v|^2(2|u|^2 + |v|^2)u + (2|u|^2 + |v|^2)v + u^2\overline{v} = \Omega u,$$
$$\frac{d^2v}{dx^2} - 2i(|u|^2 + |v|^2)\frac{dv}{dx} - 2iuv\frac{d\overline{u}}{dx} - 2|u|^2(|u|^2 + 2|v|^2)v + (|u|^2 + 2|v|^2)u + v^2\overline{u} = \Omega v,$$

and the last two elements of $\Lambda'(\mathbf{u}) = 0$ are conjugates of those. Using the reduction $u = U$ and $v = \bar{U}$, we obtain a second-order differential equation

$$\frac{d^2U}{dx^2} + 6i|U|^2\frac{dU}{dx} - 6|U|^4U + 3|U|^2\bar{U} + U^3 = \Omega U. \tag{2.10}$$

Substituting the first-order equation (2.6) to the second-order equation (2.10) yields the constraint

$$(1-\omega^2)U + \left(2|U|^4 + 2\omega|U|^2 - U^2 - \bar{U}^2\right) U = \Omega U,$$

which is satisfied by the MTM soliton $U = U_\omega$ in the explicit form (2.3) if $\Omega = 1 - \omega^2$. Therefore, the MTM soliton (2.3) is a critical point of the modified energy functional

$$\boxed{\Lambda := R + (1-\omega^2)Q, \quad \omega \in (-1,1)} \tag{2.11}$$

2.2 Spectrum of the linearized operator

From the Taypor expansion (2.7), we first see that the Taylor expansion of Λ_ω around $\mathbf{u}_\omega$ is given as

$$\Lambda(\mathbf{u} + \mathbf{u}_\omega) = \Lambda(\mathbf{u}_\omega) + \frac{1}{2}\langle L\mathbf{u}, \mathbf{u}\rangle_{L^2} + \cdots,$$

for $\mathbf{u} \in X$, where L is the Hessian operator from the energy functional Λ_ω around $\mathbf{u}_\omega$. This operator L is explicitly found as

$$L = \begin{bmatrix} L_1 & 2L_2 & L_2 & L_3 \\ 2\bar{L}_2 & \bar{L}_1 & \bar{L}_3 & \bar{L}_2 \\ \bar{L}_2 & \bar{L}_3 & \bar{L}_1 & 2\bar{L}_2 \\ L_3 & L_2 & 2L_2 & L_1 \end{bmatrix}, \tag{2.12}$$

where

$$\begin{aligned} L_1 &= -\frac{d^2}{dx^2} - 4i|U_\omega|^2\frac{d}{dx} - 4i\bar{U}_\omega\frac{dU_\omega}{dx} + 10|U_\omega|^4 - 2U_\omega^2 - 2\bar{U}_\omega^2 + 1 - \omega^2, \\ L_2 &= -2iU_\omega\frac{dU_\omega}{dx} + 4U_\omega^2|U_\omega|^2 - 2|U_\omega|^2, \\ L_3 &= -2i|U_\omega|^2\frac{d}{dx} - 2i\bar{U}_\omega\frac{dU_\omega}{dx} + 8|U_\omega|^4 - U_\omega^2 - \bar{U}_\omega^2. \end{aligned}$$

By taking derivative of the stationary equation $\Lambda'(T(\theta,s)\mathbf{u}_\omega) = 0$ in θ or s and setting $\theta = s = 0$, we find that the kernel vectors $iJ\mathbf{u}_\omega$ and $\partial_x\mathbf{u}_\omega$ are in the kernels of L, i.e.,

$$LiJ\mathbf{u}_\omega = 0, \quad L\partial_x\mathbf{u}_\omega = 0, \tag{2.13}$$

where $J = \text{diag}(1,1,-1,-1)$ is a diagonal matrix. By the Weyl's theorem, we see that the continuous spectrum of L is a semi-infinite strip $[1-\omega^2,\infty)$, which corresponds to the essential spectrum of the linear operator $-\frac{d^2}{dx^2} + 1 - \omega^2$. The 4×4 matrix operator L is diagonalized into two 2×2 matrix operators $L_\pm$ by means of the self-similarity transformation

$$S^tLS = \begin{bmatrix} L_+ & 0 \\ 0 & L_- \end{bmatrix}, \quad \text{where} \quad S = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

is the orthgonal matrix, i.e., $S^t = S^{-1}$. The matrix operators $L_\pm$ are found from this block-diagonalization in the explicit form:

$$L_+ = \begin{pmatrix} \ell_+ & -6\omega U_\omega^2 \\ -6\omega\bar{U}_\omega^2 & \bar{\ell}_+ \end{pmatrix}, \quad L_- = \begin{pmatrix} \ell_- & 2\omega U_\omega^2 \\ 2\omega\bar{U}_\omega^2 & \bar{\ell}_- \end{pmatrix}, \tag{2.14}$$

where

$$\begin{aligned} \ell_+ &= -\frac{d^2}{dx^2} - 6i|U_\omega|^2\frac{d}{dx} + 6|U_\omega|^4 - 3U_\omega^2 + 3\bar{U}_\omega^2 - 6\omega|U_\omega|^2 + 1 - \omega^2, \\ \ell_- &= -\frac{d^2}{dx^2} - 2i|U_\omega|^2\frac{d}{dx} - 2|U_\omega|^4 - U_\omega^2 + \bar{U}_\omega^2 - 2\omega|U_\omega|^2 + 1 - \omega^2. \end{aligned}$$

The continuous spectrum of $L_\pm$, by the Weyl theorem, is $[1-\omega^2,\infty)$. By applying self-similarity transformation S^t to kernel vectors $iJ\mathbf{u}_\omega$ and $\partial_x\mathbf{u}_\omega$ in (2.13), we obtain

$$S^t iJ\mathbf{u}_\omega = -i\sqrt{2}(0,0,U_\omega,-\overline{U}_\omega)^t, \quad S^t\partial_x\mathbf{u}_\omega = \sqrt{2}(U'_\omega,\overline{U}'_\omega,0,0)^t.$$

Therefore, vectors $(U'_\omega,\overline{U}'_\omega)^t$ and $(U_\omega,-\overline{U}_\omega)^t$ are in the kernels of L_+ and L_- for any $\omega \in (-1,1)$, i.e.,

$$L_+(U'_\omega,\overline{U}'_\omega)^t = 0, \quad L_-(U_\omega,-\overline{U}_\omega)^t = 0. \tag{2.15}$$

In addition, for $\omega = 0$, operators $L_\pm$ are diagonal, and we can explicitly find that

$$\omega = 0 : L_+(U'_0,-\overline{U}'_0)^t = (0,0)^t, \quad L_-(U_0,\overline{U}_0)^t = (0,0)^t. \tag{2.16}$$

Next, we count discrete eigenvalues of $L_\pm$ in (2.14).

Lemma 1. *For any $\omega \in (-1,1)$, operator L_- has exactly two eigenvalues below the continuous spectrum. Besides the zero eigenvalue associated with the eigenvector in (2.15), L_- has a positive eigenvalue for $\omega \in (0,1)$ and a negative eigenvalue for $\omega \in (-1,0)$.*

Proof. Let us consider the eigenvalue problem $L_-\mathbf{u} = \mu\mathbf{u}$, where $\mathbf{u} = (u,\bar{u})$ is an eigenvector and μ is the spectral parameter. Using the transformation

$$u(x) = \varphi(x)e^{-i\int_0^x |U_\omega(x')|^2dx'}$$

where φ is a new eigenfunction, we obtain an equivalent spectral problem:

$$\left(\mathbf{s}I_{2\times 2} + 2\omega|U_\omega|^2\sigma_1\right)\begin{bmatrix}\varphi\\ \overline{\varphi}\end{bmatrix} = \mu\begin{bmatrix}\varphi\\ \overline{\varphi}\end{bmatrix},$$

where $\mathbf{s} = -\partial_x^2 + 1 - \omega^2 - 2\omega|U_\omega|^2 - 3|U_\omega|^4$, thanks to the fact that

$$U_\omega^2 e^{2i\int_0^x |U_\omega(x')|^2dx'} = \frac{1-\omega^2}{\omega + \cosh(2\sqrt{1-\omega^2}x)} = |U_\omega|^2.$$

Because the off-diagonal entries are real, we set

$$\psi_\pm := \varphi(x) \pm \bar{\varphi}(x), \quad z := \sqrt{1-\omega^2}x, \quad \mu := (1-\omega^2)\lambda$$

to diagonalize the spectral problem into two uncoupled spectral problems associated with the linear Schrödinger operators:

$$-\frac{d^2\psi_+}{dz^2} + \left[1 - \frac{3(1-\omega^2)}{(\omega+\cosh(2z))^2}\right]\psi_+ = \lambda\psi_+ \tag{2.17}$$

and

$$-\frac{d^2\psi_-}{dz^2} + \left[1 - \frac{3(1-\omega^2)}{(\omega+\cosh(2z))^2} - \frac{4\omega}{\omega+\cosh(2z)}\right]\psi_- = \lambda\psi_-. \tag{2.18}$$

The eigenvector (2.15) in the kernel of L_- yields the eigenfunction

$$\psi_0(z) = \frac{1}{(\omega + \cosh(2z))^{1/2}}$$

of the spectral problem (2.18) for $\lambda = 0$. Because the eigenfunction ψ_0 is positive definite, the simple zero eigenvalue of the spectral problem (2.18) is at the bottom of the Schrödinger spectral problem for any $\omega \in (-1, 1)$, by Sturm's Nodal Theorem [83, Lemma 4.2]. Furthermore, the function

$$\psi_c(z) = \frac{\sinh(2z)}{\omega + \cosh(2z)}$$

corresponds to the end-point resonance at $\lambda = 1$ for the spectral problem

$$-\frac{d^2\psi}{dz^2} + \left[1 - \frac{8(1-\omega^2)}{(\omega + \cosh(2z))^2} - \frac{4\omega}{\omega + \cosh(2z)}\right]\psi = \lambda\psi. \tag{2.19}$$

Because the function ψ_c has exactly one zero, there is only one isolated eigenvalue below the continuous spectrum for the spectral problem (2.19) by Sturm's Nodal Theorem. Now the difference between the potentials of the spectral problems (2.18) and (2.19) is

$$\Delta V(z) = \frac{5(1-\omega^2)}{(\omega + \cosh(2z))^2},$$

where $\Delta V > 0$ for all $z \in \mathbb{R}$ and $\omega \in (-1, 1)$. By Sturm's Comparison Theorem [83, Theorem B.10], a solution of the spectral problem (2.18) for $\lambda = 1$, which is bounded as $z \to -\infty$, has exactly one zero. Therefore, the spectral problem (2.18) has exactly one isolated eigenvalue λ for all $\omega \in (-1, 1)$ and this is the zero eigenvalue with the eigenfunction ψ_0.

The difference between the potentials of the spectral problems (2.17) and (2.18) is given by

$$\Delta V(z) = \frac{4\omega}{\omega + \cosh(2z)}.$$

If $\omega = 0$, $\Delta V = 0$, so that the spectral problem has only one isolated eigenvalue and it is located at $\lambda = 0$. Since $\Delta V > 0$ for $\omega \in (0, 1)$, the spectral problem (2.17) has precisely one isolated eigenvalue for $\omega \in (0, 1)$ by Sturm's Comparison Theorem and this eigenvalue is positive [54, Section I.6.10], i.e., $\lambda > 0$. On the other hand, since $\Delta V < 0$ for $\omega \in (-1, 0)$ and $\psi_0 > 0$ is an eigenfunction of the spectral problem (2.18) for $\lambda = 0$, the spectral problem (2.17) has at least one negative eigenvalue $\lambda < 0$ for $\omega \in (-1, 0)$ [54, Section I.6.10]. To show that this negative eigenvalue is the only isolated eigenvalue of the spectral problem (2.17), we note that

$$\omega + \cosh(2z) \geq \omega + 1 + 2z^2, \quad z \in \mathbb{R}$$

and consider the spectral problem

$$-\frac{d^2\psi}{dz^2} + \left[1 - \frac{3(1-\omega^2)}{(\omega+1+2z^2)^2}\right]\psi = \lambda\psi. \tag{2.20}$$

Rescaling the independent variable $z := \frac{\sqrt{1+\omega}}{\sqrt{2}}y$ and denoting $\psi(z) := \tilde{\psi}(y)$, we rewrite (2.20) in the equivalent form

$$-\frac{d^2\tilde{\psi}}{dy^2} - \frac{3}{(1+y^2)^2}\left(1 - \frac{1+\omega}{2}\right)\tilde{\psi} = \frac{(\lambda-1)(1+\omega)}{2}\tilde{\psi}. \tag{2.21}$$

It follows that the function

$$\tilde{\psi}_c(y) = \frac{y}{\sqrt{1+y^2}}$$

corresponds to the end-point resonance at $\lambda = 1$ for the spectral problem

$$-\frac{d^2\tilde{\psi}}{dy^2} - \frac{3}{(1+y^2)^2}\tilde{\psi} = \frac{(\lambda-1)(1+\omega)}{2}\tilde{\psi}. \tag{2.22}$$

Because the function $\tilde{\psi}_c$ has exactly one zero, there is only one isolated eigenvalue below the continuous spectrum for the spectral problem (2.22). Because the difference between potentials of the spectral problems (2.21) and (2.22) as well as those of the spectral problems (2.17) and (2.20) is strictly positive for all $\omega \in (-1,1)$, by Sturm Comparison Theorem, the spectral problem (2.17) has exactly one isolated eigenvalue λ for all $\omega \in (-1,1)$ and this eigenvalue is negative for $\omega \in (-1,0)$, zero for $\omega = 0$, and positive for $\omega \in (0,1)$. □

For the operator L_+, we can only prove the statement for small ω due to the technical reason.

Lemma 2. *There is $\omega_0 \in (0,1]$ such that for any fixed $\omega \in (-\omega_0,\omega_0)$, operator L_+ has exactly two eigenvalues below the continuous spectrum. Besides the zero eigenvalue associated with the eigenvector in (2.15), L_+ also has a negative eigenvalue for $\omega \in (0,\omega_0)$ and a positive eigenvalue for $\omega \in (-\omega_0, 0)$.*

Proof of Lemma 2. Because the double zero eigenvalue of L_+ at $\omega = 0$ is isolated from the continuous spectrum located for $[1,\infty)$, the assertion of the lemma will follow by the Kato's perturbation theory [54] if we can show that the zero eigenvalue is the lowest eigenvalue of L_+ at $\omega = 0$ and the end-point of the continuous spectrum does not admit a resonance.

To develop the perturbation theory, we consider the eigenvalue problem $L_+\mathbf{u} = \mu\mathbf{u}$, where $\mathbf{u} = (u,\bar{u})$ is an eigenvector and μ is the spectral parameter. Using the transformation

$$u(x) = \varphi(x)e^{-3i\int_0^x |U_\omega(x')|^2 dx'}$$

where φ is a new eigenfunction, we obtain an equivalent spectral problem:

$$(\mathbf{s}I_{2\times 2} - 6\omega W\sigma_1)\begin{bmatrix}\varphi \\ \bar{\varphi}\end{bmatrix} = \mu\begin{bmatrix}\varphi \\ \bar{\varphi}\end{bmatrix},$$

where $\mathbf{s} = -\partial_x^2 + 1 - \omega^2 - 6\omega|U_\omega|^2 - 3|U_\omega|^4$ and

$$\begin{aligned} W &= U_\omega^2 e^{6i\int_0^x |U_\omega(x')|^2 dx'} \\ &= (1-\omega^2)\frac{\left(1+\omega\cosh\left(2\sqrt{1-\omega^2}x\right) + i\sqrt{1-\omega^2}\sinh\left(2\sqrt{1-\omega^2}x\right)\right)^2}{\left(\omega+\cosh\left(2\sqrt{1-\omega^2}x\right)\right)^3}. \end{aligned}$$

Setting now $z := \sqrt{1-\omega^2}x$ and $\mu := (1-\omega^2)\lambda$, we rewrite the spectral problem in the form

$$\begin{bmatrix} -\partial_z^2 + 1 + V_1(z) & V_2(z) \\ \bar{V}_2(z) & -\partial_z^2 + 1 + V_1(z) \end{bmatrix} \begin{bmatrix} \varphi \\ \bar{\varphi} \end{bmatrix} = \lambda \begin{bmatrix} \varphi \\ \bar{\varphi} \end{bmatrix}, \tag{2.23}$$

where

$$V_1(z) := -\frac{3(1-\omega^2)}{(\omega+\cosh(2z))^2} - \frac{6\omega}{\omega+\cosh(2z)}$$

and

$$V_2(z) := -6\omega\frac{\left(1+\omega\cosh(2z) + i\sqrt{1-\omega^2}\sinh(2z)\right)^2}{\left(\omega+\cosh(2z)\right)^3}.$$

The eigenvector (2.15) in the kernel of L_+ yields the eigenvector $(\varphi_\omega, \bar{\varphi}_\omega)$ with

$$\varphi_\omega(z) = \frac{\omega\sinh(2z) + i\sqrt{1-\omega^2}\cosh(2z)}{(\omega+\cosh(2z))^{3/2}},$$

which exists in the spectral problem (2.23) with $\lambda = 0$ for all $\omega \in (-1,1)$. Now, for $\omega = 0$, $\lambda = 0$ is a double zero eigenvalue of the spectral problem (2.23). The other eigenvector is $(\varphi_0, -\bar{\varphi}_0)$ and it corresponds to the eigenvector in (2.16). The end-point $\lambda = 1$ of the continuous spectrum of the spectral problem (2.23) does not admit a resonance for $\omega = 0$, which follows from the comparison results in Lemma 1. No other eigenvalues exist for $\omega = 0$.

To study the splitting of the double zero eigenvalue if $\omega \neq 0$, we compute the quadratic form of the operator on the left-hand side of the spectral problem (2.23) at the vector $(\varphi_0, -\bar{\varphi}_0)$ to obtain

$$-2\int_{\mathbb{R}} (V_2\overline{\varphi}_0^2 + \overline{V}_2\varphi_0^2)dz = -12\omega\int_{\mathbb{R}} \frac{3-2\omega^2-\cosh(4z)}{(\omega+\cosh(2z))^4}dz.$$

Since the integral is positive for $\omega = 0$, Kato's perturbation theory [54, Section VII.4.6] implies that the zero eigenvalue of the spectral problem (2.23) becomes negative for $\omega > 0$ and positive for $\omega < 0$ with sufficiently small $|\omega|$. □

Conjecture 1. *The spectral problem* (2.23) *has exactly two isolated eigenvalues and no endpoint resonances for all* $\omega \in (-1,1)$. *The nonzero eigenvalue is positive for all* $\omega \in (-1,0)$ *and negative for all* $\omega \in (0,1)$.

To illustrate Conjecture 1, we approximate eigenvalues of the spectral problem (2.23) numerically. We use the second-order central difference scheme for the second derivative and the periodic boundary conditions. Figure 2.1 shows the only two isolated eigenvalues of the spectral problem (2.23) (asterisks) and the edge of the continuous spectrum at $\lambda = 1$ (dashed line) versus parameter $\omega \in (-1, 1)$. The nonzero eigenvalue is positive for all $\omega \in (-1, 0)$ and negative for all $\omega \in (0, 1)$.

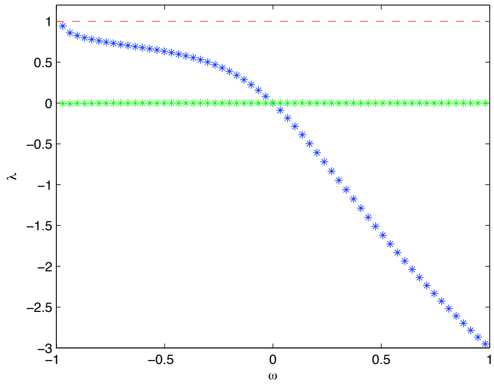


Figure 2.1: Isolated eigenvalue λ (asterisks) and the edge of the continuous spectrum $\lambda = 1$ (dashed line) versus parameter ω in the spectral problem (2.23).

From (2.16), by applying S to vectors $(U_0', -\overline{U}_0', 0, 0)^t$ and $(0, 0, U_0, \overline{U}_0)^t$, we deduce that

$$\omega = 0: \quad L(U_0', -\overline{U}_0', -\overline{U}_0', U_0')^t = 0, \quad L(-U_0, \overline{U}_0, -\overline{U}_0, U_0)^t = 0. \tag{2.24}$$

Therefore, for $\omega = 0$, operator L has four zero eigenvalues with eigenvectors from (2.24) and (2.13). Now, when $\omega \neq 0$, we have proved the following eigenvalue bifurcations:

Corollary 1. *There exists a $\omega_0 \in (0, 1]$ such that for any $\omega \in (-\omega_0, \omega_0) \setminus \{0\}$ the operator L has exactly four eigenvalues below the continuous spectrum. Besides the zero eigenvales associated wth eigenvectors in* (2.13)*, the other two eigenvalues are nonzero with different signs when $\omega \in (-\omega_0, \omega_0) \setminus \{0\}$.*

2.3 Positivity of the Hessian operator

We denote the positive subspace of the operator L in space X by P. The positive spectrum of L is bounded away from zero. By spectral theorem, there exists a positive constant $c > 0$ such that

$$\langle L\mathbf{u}, \mathbf{u} \rangle_{L^2} \geq c \|\mathbf{u}\|_X^2,$$

for every $\mathbf{u} \in P \subset X$.

We denote the eigenvector of L for the only negative eigenvalue of L by $\mathbf{n}$ (if$\omega \neq 0$):

$$L\mathbf{n} = -\lambda^2 \mathbf{n}, \quad \|\mathbf{n}\|_{L^2} = 1.$$

We have additionally the two dimensional kernel of L for $\omega \neq 0$.

We first start with showing positivity of operator L for the case of $\omega \in (0, 1)$.

Lemma 3. *There exists a* $\omega_0 \in (0,1]$ *such that for any* $\omega \in (0,\omega_0)$*, if* $0 = \langle Q'(\mathbf{u}_\omega), \mathbf{y}\rangle_{L^2} = \langle iJ\mathbf{u}_\omega, \mathbf{y}\rangle_{L^2} = \langle \partial_x \mathbf{u}_\omega, \mathbf{y}\rangle_{L^2}$*, then there exists a constant* $k > 0$ *such that*

$$\langle L\mathbf{y}, \mathbf{y}\rangle_{L^2} \geq k\|\mathbf{y}\|_X^2,$$

for $\mathbf{y} \in X$*.*

Proof. Differentiating $\Lambda(\mathbf{u}_\omega)$ twice in $\Omega = 1-\omega^2$ yields $\partial_\Omega \Lambda = Q$ and

$$\partial_\Omega^2 \Lambda(\mathbf{u}_\omega) = \langle Q'(\mathbf{u}_\omega), \partial_\Omega \mathbf{u}_\omega\rangle_{L^2} = -\frac{1}{\omega}\partial_\omega \int_{\mathbb{R}} |U_\omega|^2 dx = \frac{1}{\omega\sqrt{1-\omega^2}}. \tag{2.25}$$

We see that $\partial_\Omega^2 \Lambda(\mathbf{u}_\omega) > 0$ for $\omega \in (0,1)$. Differentiating the stationary equation $R'(\mathbf{u}_\omega) + \Omega Q'(\mathbf{u}_\omega) = 0$ in Ω gives

$$L\partial_\Omega \mathbf{u}_\omega = -Q'(\mathbf{u}_\omega).$$

We find for $\omega \in (0,1)$

$$0 < \langle Q'(\mathbf{u}_\omega), \partial_\Omega \mathbf{u}_\omega\rangle_{L^2} = -\langle L\partial_\Omega \mathbf{u}_\omega, \partial_\Omega \mathbf{u}_\omega\rangle_{L^2}.$$

This implies that a vector $\partial_\Omega \mathbf{u}_\omega$ is in a negative direction of L. We make the spectral decomposition of $\partial_\Omega \mathbf{u}_\omega$ with respect to the spectrum of L:

$$\partial_\Omega \mathbf{u}_\omega = a_0 \mathbf{n} + b_0 iJ\mathbf{u}_\omega + c_0 \partial_x \mathbf{u}_\omega + \mathbf{p}_0, \quad \mathbf{p}_0 \in P,$$

where some $a_0, b_0, c_0 \in \mathbb{C}$. We find that

$$0 > \langle L\partial_\Omega \mathbf{u}_\omega, \partial_\Omega \mathbf{u}_\omega\rangle_{L^2} = -|a_0|^2\lambda^2 + \langle L\mathbf{p}_0, \mathbf{p}_0\rangle_{L^2} \tag{2.26}$$

For any $\mathbf{y} \in X$ with $0 = \langle Q'(\mathbf{u}_\omega), \mathbf{y}\rangle_{L^2} = \langle iJ\mathbf{u}_\omega, \mathbf{y}\rangle_{L^2} = \langle \partial_x \mathbf{u}_\omega, \mathbf{y}\rangle_{L^2}$, we have the decomposition

$$\mathbf{y} = a\mathbf{n} + \mathbf{p}, \quad \mathbf{p} \in P,$$

where some $a \in \mathbb{C}$ and since $0 = \langle Q'(\mathbf{u}_\omega), \mathbf{y}\rangle_{L^2}$ we have

$$0 = \langle Q'(\mathbf{u}_\omega), \mathbf{y}\rangle_{L^2} = -\langle L\partial_\Omega \mathbf{u}_\omega, \mathbf{y}\rangle_{L^2} = a_0\bar{a}\lambda^2 - \langle L\mathbf{p}_0, \mathbf{p}\rangle_{L^2}. \tag{2.27}$$

Therefore, by the Schwarz inequality $\langle L\mathbf{p}, \mathbf{p}\rangle_{L^2}\langle L\mathbf{p}_0, \mathbf{p}_0\rangle_{L^2} \geq |\langle L\mathbf{p}, \mathbf{p}_0\rangle_{L^2}|^2$ which follows from

$$\langle L(\mathbf{p}_0 - \lambda\mathbf{p}), (\mathbf{p}_0 - \lambda\mathbf{p})\rangle_{L^2} \geq 0 \quad \text{with } \lambda = \sqrt{\frac{\langle L\mathbf{p}_0, \mathbf{p}_0\rangle_{L^2}}{\langle L\mathbf{p}, \mathbf{p}\rangle_{L^2}}},$$

we find

$$\langle L\mathbf{y}, \mathbf{y}\rangle_{L^2} = -|a|^2\lambda^2 + \langle L\mathbf{p}, \mathbf{p}\rangle_{L^2} \geq -|a|^2\lambda^2 + \frac{|\langle L\mathbf{p}, \mathbf{p}_0\rangle_{L^2}|^2}{\langle L\mathbf{p}_0, \mathbf{p}_0\rangle_{L^2}} > 0. \tag{2.28}$$

The last strict inequality is due to (2.26) and (2.27).

From (2.28), we see that the orthogonal subspace to which $\mathbf{y}$ belongs to P. Since the spectrum of L is bounded away from zero, it follows that there exists a constant $k > 0$ such that

$$\langle L\mathbf{y}, \mathbf{y}\rangle_{L^2} \geq k.$$

When we replace $\mathbf{y}$ with $\mathbf{y}/\|\mathbf{y}\|_X$ in the above inequality, we attain the assertion of the Lemma. □

We have seen that the vector orthogonal to the change in mass, $Q'(\mathbf{u}_\omega)$, satisfies coercivity of L. Since the solution stays on the manifold of constant Q thanks to the mass conservation, then, intuitively, the negativity of L that comes from the change of mass does not contribute to instability.

To deal with the case of $\omega \in (-\omega_0, 0)$, on the other hand, we find that this comes from the change in momentum,

$$P'(\mathbf{u}_\omega) = i(-U'_\omega, -\overline{U}'_\omega, \overline{U}'_\omega, U'_\omega)^t.$$

Proposition 1. *The vector* $\mathbf{g} := \frac{i}{2}xJ\mathbf{u}_\omega + \frac{1}{4\omega}\gamma_2\mathbf{u}_\omega$, *where* $\gamma_2 = \mathrm{diag}(-1, 1, -1, 1)$, *satisfies*

$$L\mathbf{g} = P'(\mathbf{u}_\omega) \tag{2.29}$$

and, furthermore,

$$\langle L\mathbf{g}, \mathbf{g}\rangle_{L^2} = \frac{\sqrt{1-\omega^2}}{2\omega}$$

negative for $\omega \in (-1, 0)$.

Proof of Proposition 1. In order to find $\mathbf{g}$ that satisfies (2.29), it is convenient to carry out the block-diagonalization:

$$(S^tLS)S^t\mathbf{g} = S^tP'(\mathbf{u}_\omega) = i\sqrt{2}(0, 0, U'_\omega, -\overline{U}'_\omega)^t.$$

Since the first two components of $S^tP'(\mathbf{u}_\omega)$ are zero, we deduce that the first two components of $S^t\mathbf{g}$ are zero as well since the kernel of L_+ is already found in (2.15). Now, the last two components of $S^t\mathbf{g}$ is found by using the differential equations (2.6) and (2.10):

$$L_-\left(-\frac{i\sqrt{2}}{2}x\begin{bmatrix} U_\omega \\ -\bar{U}_\omega \end{bmatrix} + \frac{i\sqrt{2}}{4i\omega}\begin{bmatrix} U_\omega \\ \bar{U}_\omega \end{bmatrix}\right) = i\sqrt{2}\begin{bmatrix} U'_\omega \\ -\bar{U}'_\omega \end{bmatrix}, \tag{2.30}$$

from which we can explicitly find $\mathbf{g}$ as well as we can easily compute the following

$$\begin{aligned}
\langle L\mathbf{g},\mathbf{g}\rangle_{L^2} &= \langle (S^tLS)S^t\mathbf{g}, S^t\mathbf{g}\rangle_{L^2} \\
&= \int_{\mathbb{R}} \left(|U_\omega|^2 - \frac{1}{2i\omega}\left(\bar{U}_\omega U'_\omega - U_\omega \bar{U}'_\omega\right)\right) dx \\
&= \frac{1}{2\omega}\int_{\mathbb{R}} \left(4|U_\omega|^4 - U_\omega^2 - \bar{U}_\omega^2 + 4\omega|U_\omega|^2\right) dx \\
&= \frac{1-\omega^2}{\omega}\int_{\mathbb{R}} \frac{1+\omega\cosh(2\sqrt{1-\omega^2}x)}{(\omega+\cosh(2\sqrt{1-\omega^2}x))^2} dx \\
&= \frac{\sqrt{1-\omega^2}}{\omega}.
\end{aligned}$$

□

Thanks to Proposition 1, we can repeat the same proof in Lemma 3 to prove the following:

Lemma 4. *There exists a* $\omega_0 \in (0,1]$ *such that for any* $\omega \in (-\omega_0, 0)$, *if* $0 = \langle P'(\mathbf{u}_\omega), \mathbf{y}\rangle_{L^2} = \langle iJ\mathbf{u}_\omega, \mathbf{y}\rangle_{L^2} = \langle \partial_x \mathbf{u}_\omega, \mathbf{y}\rangle_{L^2}$, *then there exists a constant* $k > 0$ *such that*

$$\langle L\mathbf{y},\mathbf{y}\rangle_{L^2} \geq k\|\mathbf{y}\|_X^2,$$

for $\mathbf{y} \in X$.

Here, we denote $\gamma_1 = \text{diag}(1,-1,-1,1)$ and $\gamma_2 = \text{diag}(-1,1,-1,1)$ so that vectors $\gamma_1\partial_x\mathbf{u}_0$ and $\gamma_2\mathbf{u}_0$ are kernel vectors of L at $\omega = 0$ from (2.24), as well as $iJ\mathbf{u}_0$ and $\partial_x\mathbf{u}_0$.

Lemma 5. *For* $\omega = 0$, *if* $0 = \langle \gamma_1\partial_x\mathbf{u}_0, \mathbf{y}\rangle_{L^2} = \langle \gamma_2\mathbf{u}_0, \mathbf{y}\rangle_{L^2}$ *and* $0 = \langle iJ\mathbf{u}_0, \mathbf{y}\rangle_{L^2} = \langle \partial_x\mathbf{u}_0, \mathbf{y}\rangle_{L^2}$, *then there exists a constant* $k > 0$ *such that*

$$\langle L\mathbf{y},\mathbf{y}\rangle_{L^2} \geq k\|\mathbf{y}\|_X^2,$$

for $\mathbf{y} \in X$.

Proof. This follows from the fact that operator L has exactly four zero eigenvalues below the positive continuous spectrum that is bounded away from zero. □

2.4 Proof of orbital stability

Before giving a proof of Theorem 1, we will collect final key ingredients.

Lemma 6. *There exist* $\epsilon > 0$ *and a differentiable map*

$$(\theta, s) : \Phi_\epsilon \to \mathbb{R}^2$$

such that, for all $\mathbf{u} \in \Phi_\epsilon$*, the following is true:*

$$\langle T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u}, iJ\mathbf{u}_\omega\rangle_{L^2} = \langle T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u}, \partial_x \mathbf{u}_\omega\rangle_{L^2} = 0 \tag{2.31}$$

and the function $\rho(\theta, s) := \|T(\theta, s)\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2$ *has a local minimum at* $(\theta, s) = (\theta(\mathbf{u}), s(\mathbf{u}))$.

Proof. We define a function $\rho(\theta, s) = \|T(\theta, s)\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2$ at a fixed ω for every $\mathbf{u} \in \Phi_\epsilon$. Derivatives of ρ in θ and s yield

$$\partial_\theta \rho = 2\langle T(\theta, s)\mathbf{u}, iJ\mathbf{u}_\omega\rangle_{L^2}, \quad \partial_s \rho = 2\langle T(\theta, s)\mathbf{u}, \partial_x \mathbf{u}_\omega\rangle_{L^2},$$

$$\partial_\theta^2 \rho = 2\langle T(\theta, s)\mathbf{u}, \mathbf{u}_\omega\rangle_{L^2}, \quad \partial_s^2 \rho = 2\langle T(\theta, s)\partial_x\mathbf{u}, \partial_x \mathbf{u}_\omega\rangle_{L^2}$$

and

$$\partial_s\partial_\theta \rho = \partial_\theta\partial_s \rho = 2\langle T(\theta, s)\partial_x \mathbf{u}, iJ\mathbf{u}_\omega\rangle_{L^2}.$$

We find that $\partial_\theta \rho = \partial_s \rho = 0$ at $\theta = s = 0$ and $\mathbf{u} = \mathbf{u}_\omega$, but $\partial_\theta^2 \rho = 2\|\mathbf{u}_\omega\|_{L^2}^2$, $\partial_s^2 \rho = 2\|\partial_s \mathbf{u}_\omega\|_{L^2}^2$, and $\partial_s\partial_\theta \rho = 0$ at $\theta = s = 0$ and $\mathbf{u} = \mathbf{u}_\omega$. The determinant of $\begin{bmatrix} \partial_\theta^2 \rho & \partial_s\partial_\theta\rho \\ \partial_\theta\partial_s\rho & \partial_s^2\rho \end{bmatrix}$ is strictly positive at $\theta = s = 0$ and $\mathbf{u} = \mathbf{u}_\omega$. In order to prove Lemma, we want to run the implicit function theorem on $F(\theta, s) := (\partial_\theta\rho, \partial_s\rho)$, since $F(\theta, s) = 0$ implies the orthogonality (2.31) and the positivity of the Jacobian of F together with $F(\theta, s) = 0$ implies a local minimum of $\rho(\theta, s)$ at (θ, s).

The implicit function theorem tells that there exist $\epsilon > 0$ and the neighborhood $I \subset \mathbb{R}$ around $(\theta, s) = (0, 0)$ such that for every $\mathbf{u} \in \Phi_\epsilon$ there exists a unique solution (θ, s) of $F(\theta, s) = 0$. Furthermore, a map $(\theta, s) : \Phi_\epsilon \to I$ is a C^1 map. □

Thanks to the above Lemma, we can make the following decomposition in space X:

Corollary 2. *Let* $\theta(\mathbf{u})$ *and* $s(\mathbf{u})$ *be the ones determined in Lemma 6. Then,*

$$\mathbf{z} := T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega$$

is orthogonal to the kernel vectors of L *in space* X*, i.e.,*

$$\langle \mathbf{z}, iJ\mathbf{u}_\omega\rangle_{L^2} = \langle \mathbf{z}, \partial_x \mathbf{u}_\omega\rangle_{L^2} = 0. \tag{2.32}$$

Proof. This follows from (2.31), because $\mathbf{u}_\omega$ satisfies (2.32). □

The following Lemma states that $\mathbf{u}_\omega$ is a local constrained minimizer of the energy functional R.

Lemma 7. *There exists a* $\omega_0 \in (0, 1]$ *such that for any* $\omega \in (-\omega_0, \omega_0)$*, there exist* $c > 0$ *and* $\epsilon > 0$ *such that*

$$R(\mathbf{u}) - R(\mathbf{u}_\omega) \geq c\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|^2$$

for every $\mathbf{u} \in \Phi_\epsilon$ *with a fixed mass* $Q(\mathbf{u}) = Q(\mathbf{u}_\omega)$ *and* $P(\mathbf{u}) = P(\mathbf{u}_\omega)$.

Proof. First, we consider the case of $\omega \in (0, \omega_0)$. We begin by decomposing $T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega$ as

$$T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega = a\mathbf{q} + \mathbf{y}, \quad \langle \mathbf{q}, \mathbf{y} \rangle_{L^2} = 0, \tag{2.33}$$

where $\mathbf{q} = Q'(\mathbf{u}_\omega)$ and some $a \in \mathbb{C}$. Since $Q'(\mathbf{u}_\omega)$ is orthogonal to the kernel of L, from Corollary 2, $\mathbf{y}$ is also orthogonal to the kernel of L. Since $Q(\mathbf{u}) = Q(\mathbf{u}_\omega)$, we find

$$\begin{aligned}
Q(\mathbf{u}_\omega) = Q(\mathbf{u}) &= Q(T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u}) \\
&= Q(\mathbf{u}_\omega + a\mathbf{q} + \mathbf{y}) \\
&= Q(\mathbf{u}_\omega) + \langle \mathbf{q}, a\mathbf{q} + \mathbf{y} \rangle_{L^2} + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2) \\
&= Q(\mathbf{u}_\omega) + \overline{a}\|\mathbf{q}\|_{L^2}^2 + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2),
\end{aligned}$$

that is, $a = \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2)$. Next, thanks to smallness of constant a, we can show

$$\begin{aligned}
\Lambda(\mathbf{u}) - \Lambda(\mathbf{u}_\omega) &= \frac{1}{2}\langle L(a\mathbf{q} + \mathbf{y}), a\mathbf{q} + \mathbf{y} \rangle_{L^2} + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{H^1}^3) \\
&= \frac{1}{2}\langle L\mathbf{y}, \mathbf{y} \rangle_{L^2} + \mathcal{O}(|a|^2) + \mathcal{O}(a\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2) \\
&\qquad + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{H^1}^3) \\
&= \frac{1}{2}\langle L\mathbf{y}, \mathbf{y} \rangle_{L^2} + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{H^1}^3).
\end{aligned}$$

We obtained

$$R(\mathbf{u}) - R(\mathbf{u}_\omega) = \frac{1}{2}\langle L\mathbf{y}, \mathbf{y} \rangle_{L^2} + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{H^1}^3) \tag{2.34}$$

Therefore, by Lemma 3, inequality (2.34) becomes

$$R(\mathbf{u}) - R(\mathbf{u}_\omega) \geq \frac{1}{2}c\|\mathbf{y}\|_X^2 + \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{H^1}^3). \tag{2.35}$$

Since $\|\mathbf{y}\|_X = \|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega - a\mathbf{q}\|_X \geq \|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_X - |a|\|\mathbf{q}\|_X$ and $a = \mathcal{O}(\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_{L^2}^2)$, for sufficiently small $\epsilon > 0$ for Φ_ϵ, inequality (2.35) yields

$$R(\mathbf{u}) - R(\mathbf{u}_\omega) \geq \frac{1}{4}c\|T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_\omega\|_X^2.$$

The other cases can be shown with slight modifications. For the case of $\omega \in (-\omega_0, 0)$, we replace $\mathbf{q} = Q'(\mathbf{u}_\omega)$ with $\mathbf{q} = P'(\mathbf{u}_\omega)$ in (2.33) and use $P(\mathbf{u}) = P(\mathbf{u}_\omega)$ to show smallness of a constant a, and use Lemma 4 to obtain (2.35) since $P'(\mathbf{u}_\omega)$ is orthogonal to the kernel of L, that is, $\mathbf{y}$ is also orthogonal to the kernel of L due to Corollary 2.

For the case $\omega = 0$, we use the decomposition $T(\theta(\mathbf{u}), s(\mathbf{u}))\mathbf{u} - \mathbf{u}_0 = \mathbf{q} + \mathbf{y}$ with $\mathbf{q} = a\gamma_1\partial_x\mathbf{u}_0 + b\gamma_2\mathbf{u}_0$ and $\langle \mathbf{q}, \mathbf{y} \rangle = 0$ in (2.33). One can easily verify that $\gamma_1\partial_x\mathbf{u}_0$ and $\gamma_2\mathbf{u}_0$ are orthogonal to the kernel vectors of L, and so is $\mathbf{y}$. Lemma 5 is used to obtain (2.35). □

Finally, we give a proof of Theorem 1. This is, in fact, a straightforward consequence of Lemma 7 by a contradiction argument. We will denote $\{\mathbf{u}_n(0)\}$ as a sequence of initial data and $\{\mathbf{u}_n(t)\}$ as a sequence of corresponding solutions.

Proof of Theorem 1. We only consider the case of $\omega \in (0, \omega_0)$ since the other cases follow in the same way.
Suppose that Theorem 1 does not hold. For every $\epsilon > 0$, there exist $N, \delta > 0$ and a sequence $\{\mathbf{u}_n(0)\}$ such that if $n > N$

$$\inf_{\theta, s \in \mathbb{R}} \|\mathbf{u}_n(0) - T(\theta, s)\mathbf{u}_\omega\|_X < \epsilon$$

and

$$\sup_{t>0} \inf_{\theta, s \in \mathbb{R}} \|\mathbf{u}_n(t) - T(\theta, s)\mathbf{u}_\omega\|_X \geq \delta.$$

Since $\mathbf{u}_n(t)$ depends continuously on time t, we can pick t_n so that $\inf_{\theta, s\in\mathbb{R}} \|\mathbf{u}_n(t_n) - T(\theta, s)\mathbf{u}_\omega\|_{H^1} = \delta$. By continuity of functionals R and Q on H^1 space,

$$R(\mathbf{u}_n(t_n)) = R(\mathbf{u}_n(0)) \to R(\mathbf{u}_\omega)$$
$$Q(\mathbf{u}_n(t_n)) = Q(\mathbf{u}_n(0)) \to Q(\mathbf{u}_\omega).$$

We make decomposition:

$$\mathbf{v}_n = \mathbf{u}_n(t_n) + \mathbf{r}_n$$

for each n such that $Q(\mathbf{v}_n) = Q(\mathbf{u}_\omega)$ and a remainder $\|\mathbf{r}_n\|_{H^1} \to 0$. By continuity of R, we have $R(\mathbf{v}_n) \to R(\mathbf{u}_\omega)$. Choosing ϵ sufficiently small, we apply Lemma 7 to obtain

$$R(\mathbf{v}_n) - R(\mathbf{u}_\omega) \geq c\|T(\theta(\mathbf{v}_n), s(\mathbf{v}_n))\mathbf{v}_n - \mathbf{u}_\omega\|_X^2,$$

for $\mathbf{v}_n \in \Phi_\epsilon$, where the left hand side goes to zero. Hence

$$\|\mathbf{u}_n(t_n) - T(-\theta(\mathbf{v}_n), -s(\mathbf{v}_n))\mathbf{u}_\omega\|_X \leq \|T(\theta(\mathbf{v}_n), s(\mathbf{v}_n))\mathbf{v}_n - \mathbf{u}_\omega\|_X + \|\mathbf{r}_n\|_X < \delta,$$

for n large enough. This contradicts our assumption. □

2.5 Conserved quantities by the inverse scattering method

The MTM (2.1) is a compatibility condition of the Lax system

$$\frac{\partial}{\partial x}\phi = L\phi, \quad \frac{\partial}{\partial t}\phi = A\phi, \tag{2.36}$$

where $\vec{\phi}(x,t) : \mathbb{R} \times \mathbb{R} \to \mathbb{C}^2$ and L and A are given by

$$L = \frac{i}{2}(|v|^2 - |u|^2)\sigma_3 - \frac{i\lambda}{\sqrt{2}}\begin{pmatrix} 0 & \overline{v} \\ v & 0 \end{pmatrix} - \frac{i}{\sqrt{2}\lambda}\begin{pmatrix} 0 & \overline{u} \\ u & 0 \end{pmatrix} + \frac{i}{4}\left(\frac{1}{\lambda^2} - \lambda^2\right)\sigma_3, \tag{2.37}$$

$$A = \frac{i}{2}(|v|^2 + |u|^2)\sigma_3 - \frac{i\lambda}{\sqrt{2}}\begin{pmatrix} 0 & \overline{v} \\ v & 0 \end{pmatrix} + \frac{i}{\sqrt{2}\lambda}\begin{pmatrix} 0 & \overline{u} \\ u & 0 \end{pmatrix} - \frac{i}{4}\left(\frac{1}{\lambda^2} + \lambda^2\right)\sigma_3.$$

The MTM (2.1) is equivalent to the expression

$$L_t - A_x + [L, A] = 0.$$

When the potential (u, v) is sufficiently smooth in x and t, existence of fundamental solutions in (2.36) can be approached by the standard ODE theory. Here, we give a formal argument. As $|x| \to \infty$, the Lax operator L has the expression

$$\lim_{|x|\to\infty} L = \frac{i}{4}\left(\frac{1}{\lambda^2} - \lambda^2\right)\sigma_3$$

which implies that solutions of $\phi_x = L\phi$, denoted as $\varphi_\pm$ and $\phi_\pm$ respectively, have limits

$$\lim_{x\to\pm\infty} e^{-ik(\lambda)x}\varphi_\pm(x) = (1, 0)^t, \quad \lim_{x\to\pm\infty} e^{ik(\lambda)x}\phi_\pm(x) = (0, 1)^t$$

where $k(\lambda) = \frac{1}{4}(\lambda^{-2} - \lambda^2) \in \mathbb{R}$ if $\lambda^2 \in \mathbb{R}$.

Since $\phi_x = L\phi$ is the first order 2×2 system, a solution is spanned by two independent ones, e.g,

$$\varphi_- = a(\lambda)\varphi_+ + b(\lambda)\phi_+,$$

for $\lambda^2 \in \mathbb{R}$, where $a(\lambda)$ and $b(\lambda)$ are coefficients, given as

$$a(\lambda) = W(\varphi_-, \phi_+), \quad b(\lambda) = W(\varphi_-, \varphi_+), \tag{2.38}$$

where W is the Wronskian determinant.

Now, looking at the time evolution system $\phi_t = A\phi$, we notice that since

$$\lim_{|x|\to\infty} A = -\frac{i}{4}\left(\frac{1}{\lambda^2} + \lambda^2\right)\sigma_3,$$

solutions $\varphi_\pm$ and $\phi_\pm$ must be modified as $e^{-id(\lambda)t}\varphi_\pm$ and $e^{id(\lambda)t}\phi_\pm$ to incorporate the boundary condition for $\phi_t = A\phi$, where $d(\lambda) = \frac{1}{4}(\lambda^{-2} + \lambda^2)$. In order to find the time evolution of coefficients $a(\lambda)$ and $b(\lambda)$ according to the time evolution of the Lax system, we write (2.38) as

$$a(\lambda) = W(e^{-id(\lambda)t}\varphi_-, e^{id(\lambda)t}\phi_+), \quad b(\lambda) = e^{2id(\lambda)t}W(e^{-id(\lambda)t}\varphi_-, e^{-id(\lambda)t}\varphi_+). \tag{2.39}$$

The Wronskians of solutions are independent of x and t since trances of L and A are zero. It follows that $a(\lambda)$ is independent of time.

We make the following ansatz:

$$\varphi_-(x, t; \lambda) = \begin{bmatrix} 1 \\ \nu(x, t; \lambda) \end{bmatrix} \exp\left(ik(\lambda)x + \int_{-\infty}^{x} \chi(x', t; \lambda)dx'\right) \tag{2.40}$$

for some suitable functions ν and χ. By substituting (2.40) into (2.38) for $a(\lambda)$

and taking $x \to \infty$, we find

$$a(\lambda) = \exp\left(\int_{-\infty}^{\infty} \chi(x;\lambda)dx\right) \quad \Rightarrow \quad \log a(\lambda) = \int_{-\infty}^{\infty} \chi(x;\lambda)dx. \tag{2.41}$$

Since the scattering coefficient $a(\lambda)$ does not depend on time t, expansion of $\int_{-\infty}^{\infty} \chi(x;\lambda)dx$ in powers of λ yields conserved quantities with respect to t [64]. Substituting equation (2.40) into the x-derivative part of the Lax system (2.36), we find that functions ν and χ must satisfy

$$\chi = \frac{i}{2}(|v|^2 - |u|^2) - \frac{i}{\sqrt{2}}\left(\lambda\overline{v} + \frac{1}{\lambda}\overline{u}\right)\nu, \tag{2.42}$$

where ν satisfies a Ricatti equation

$$\nu_x + i\left(2k(\lambda) + |v|^2 - |u|^2\right)\nu - \frac{i}{\sqrt{2}}\left(\lambda\overline{v} + \frac{1}{\lambda}\overline{u}\right)\nu^2 + \frac{i}{\sqrt{2}}\left(\lambda v + \frac{1}{\lambda}u\right) = 0. \tag{2.43}$$

We consider the formal asymptotic expansion of $\chi(x;\lambda)$ in powers for sufficiently small λ

$$\chi(x;\lambda) = \sum_{n=0}^{N} \lambda^n \chi_n(x) + o(\lambda^N), \quad \nu(x;\lambda) = \sum_{n=1}^{N} \lambda^n \nu_n(x) + o(\lambda^N) \tag{2.44}$$

and in inverse powers for sufficiently large λ

$$\chi(x;\lambda) = \sum_{n=0}^{N} \frac{1}{\lambda^n}\widetilde{\chi}_n(x) + o(\lambda^{-N}), \quad \nu(x;\lambda) = \sum_{n=1}^{N} \frac{1}{\lambda^n}\widetilde{\nu}_n(x) + o(\lambda^{-N}). \tag{2.45}$$

From (3.67), (2.43), (2.44) and (2.45), we can determine χ_n and $\widetilde{\chi}_n$ from which we define

$$I_n := \int_{-\infty}^{\infty} \chi_n(x)dx, \quad I_{-n} := \int_{-\infty}^{\infty} \widetilde{\chi}_n(x)dx. \tag{2.46}$$

Using expansions (2.44) and (2.45) for (2.41), we find the important expressions:

$$\log a(\lambda) = \sum_{even}^{N} \lambda^n I_n + o(\lambda^N), \quad \log a(1/\lambda) = \sum_{even}^{N} \lambda^n I_{-n} + o(\lambda^N)$$

for sufficiently small λ. Finally, we arrive the formulas:

$$\lim_{\lambda\to 0}\left[\frac{d^{2n}}{d\lambda^{2n}} \log a(\lambda)a(1/\lambda)\right] = I_{2n} + I_{-2n}$$

$$\lim_{\lambda\to 0}\left[\frac{d^{2n}}{d\lambda^{2n}} \log \frac{a(\lambda)}{a(1/\lambda)}\right] = I_{2n} - I_{-2n}$$

for $n \geq 0$. Let us explicitly write out first conserved quantities

$$I_0 = \int_{\mathbb{R}} (|u|^2 + |v|^2) dx,$$

$$I_2 = \int_{\mathbb{R}} (-2u_x\overline{u} + i\overline{v}u + i\overline{u}v - 2i|u|^2|v|^2) dx,$$
$$I_{-2} = \int_{\mathbb{R}} (-2v_x\overline{v} - i\overline{v}u - i\overline{u}v + 2i|u|^2|v|^2) dx,$$

$$\begin{aligned} I_4 &= \int_{\mathbb{R}} [-4i\overline{u}u_{xx} - 2(u_x\overline{v} + \overline{u}v_x) + 4\overline{u}(u|v|^2)_x + 4u_x\overline{u}(|u|^2 + |v|^2) + i(|u|^2 + |v|^2) \\ &\quad - 2iu\overline{v}(|u|^2 + |v|^2) - 2iv\overline{u}(|u|^2 + |v|^2) + 4i|u|^2|v|^2(|u|^2 + |v|^2)] dx, \end{aligned}$$

and

$$\begin{aligned} I_{-4} &= \int_{\mathbb{R}} [4i\overline{v}v_{xx} - 2(u_x\overline{v} + \overline{u}v_x) + 4\overline{v}(v|u|^2)_x + 4v_x\overline{v}(|u|^2 + |v|^2) - i(|u|^2 + |v|^2) \\ &\quad + 2iu\overline{v}(|u|^2 + |v|^2) + 2iv\overline{u}(|u|^2 + |v|^2) - 4i|u|^2|v|^2(|u|^2 + |v|^2)] dx. \end{aligned}$$

Successively, we find the following:

$$\begin{aligned} I_0 &= Q \\ \mathrm{Re}\left[\frac{1}{2} i(I_2 + I_{-2})\right] &= P \\ \mathrm{Re}\left[\frac{1}{2} i(I_2 - I_{-2})\right] &= H, \end{aligned}$$

where Q, P, and H are mass, momentum, and Hamiltonian of the MTM.

The higher conserved quantity R in (2.9) is given as

$$\mathrm{Re}\left[\frac{-1}{4} i(I_4 - I_{-4})\right] = R$$

and we shall also include

$$\begin{aligned} \mathrm{Re}\left[\frac{-1}{4} i(I_4 + I_{-4})\right] &= \int_{\mathbb{R}} \left[|u_x|^2 - |v_x|^2 + \frac{i}{4}(u_x\overline{v} + \overline{u}v_x - \overline{u}_x v - u\overline{v}_x) \right. \\ &\quad \left. -\frac{i}{2}(|u|^2 + 2|v|^2)(u_x\overline{u} - \overline{u}_x u) - \frac{i}{2}(2|u|^2 + |v|^2)(v_x\overline{v} - \overline{v}_x v) \right] dx. \end{aligned}$$

Chapter 3

Orbital Stability Theory by Bäcklund transformation

3.1 Main result

We consider the MTM of the form

$$\begin{cases} i(u_t + u_x) + v + |v|^2 u = 0, \\ i(v_t - v_x) + u + |u|^2 v = 0, \end{cases} \tag{3.1}$$

subject to an initial condition $(u,v)|_{t=0} = (u_0, v_0)$ in $H^s(\mathbb{R})$ for $s \geq 0$..

The Cauchy problem for the MTM system (3.1) is known to be locally well-posed in $H^s(\mathbb{R})$ for $s > 0$ and globally well-posed for $s > \frac{1}{2}$ [98] (see earlier results in [32]). More pertinent to our study is the global well-posedness in $L^2(\mathbb{R})$ proved in the recent works [15, 48]. The next theorem summarizes the global well-posedness result for the scopes needed in our work.

Theorem 2. *[15, 48] Let $(u_0, v_0) \in L^2(\mathbb{R})$. There exists a global solution $(u, v) \in C(\mathbb{R}; L^2(\mathbb{R}))$ to the MTM system* (2.1) *such that the charge is conserved*

$$\|u(\cdot,t)\|_{L^2}^2 + \|v(\cdot,t)\|_{L^2}^2 = \|u_0\|_{L^2}^2 + \|v_0\|_{L^2}^2 \tag{3.2}$$

for every $t \in \mathbb{R}$. Moreover, the solution is unique in a certain subspace of $C(\mathbb{R}; L^2(\mathbb{R}))$ and depends continuously on initial data $(u_0, v_0) \in L^2(\mathbb{R})$.

We are interested in orbital stability of Dirac solitons of the MTM system (3.1) given by the explicit expressions

$$\begin{cases} u_\lambda(x,t) = i\delta^{-1}\sin(\gamma)\,\mathrm{sech}\left[\alpha(x+ct) - i\frac{\gamma}{2}\right]\, e^{-i\beta(t+cx)}, \\ v_\lambda(x,t) = -i\delta\sin(\gamma)\,\mathrm{sech}\left[\alpha(x+ct) + i\frac{\gamma}{2}\right]\, e^{-i\beta(t+cx)}, \end{cases} \tag{3.3}$$

where λ is an arbitrary complex nonzero parameter that determines $\delta = |\lambda|$, $\gamma = 2\mathrm{Arg}(\lambda)$, as well as

$$c = \frac{\delta^2 - \delta^{-2}}{\delta^2 + \delta^{-2}}, \quad \alpha = \frac{1}{2}(\delta^2 + \delta^{-2})\sin\gamma, \quad \beta = \frac{1}{2}(\delta^2 + \delta^{-2})\cos\gamma.$$

Let us now state the main result of our work.

Theorem 3. *Let $(u, v) \in C(\mathbb{R}; L^2(\mathbb{R}))$ be a solution of the MTM system (2.1) in Theorem 2 and λ_0 be a complex non-zero number. There exists a real positive ϵ_0 such that if the initial value $(u_0, v_0) \in L^2(\mathbb{R})$ satisfies*

$$\epsilon := \|u_0 - u_{\lambda_0}(\cdot, 0)\|_{L^2} + \|v_0 - v_{\lambda_0}(\cdot, 0)\|_{L^2} \leq \epsilon_0, \tag{3.4}$$

then for every $t \in \mathbb{R}$, there exists $\lambda \in \mathbb{C}$ such that

$$|\lambda - \lambda_0| \leq C\epsilon, \tag{3.5}$$

$$\inf_{a,\theta\in\mathbb{R}} (\|u(\cdot + a, t) - e^{-i\theta} u_\lambda(\cdot, t)\|_{L^2} + \|v(\cdot + a, t) - e^{-i\theta} v_\lambda(\cdot, t)\|_{L^2}) \leq C\epsilon, \tag{3.6}$$

where the positive constant C is independent of ϵ and t.

Remark 1. *One can expect extension of Theorem 3 to N soliton case by the N-fold Bäcklund transformation.*

3.2 Bäcklund transformation for the MTM system

The formal compatibility condition $\vec{\phi}_{xt} = \vec{\phi}_{tx}$ for the system of linear equations

$$\vec{\phi}_x = L\vec{\phi} \quad \text{and} \quad \vec{\phi}_t = A\vec{\phi} \tag{3.7}$$

yields the MTM system (2.1), where L and A are given by

$$L = \frac{i}{4}(|u|^2 - |v|^2)\sigma_3 - \frac{i\lambda}{2}\begin{pmatrix} 0 & \overline{v} \\ v & 0 \end{pmatrix} + \frac{i}{2\lambda}\begin{pmatrix} 0 & \overline{u} \\ u & 0 \end{pmatrix} + \frac{i}{4}\left(\lambda^2 - \frac{1}{\lambda^2}\right)\sigma_3 \tag{3.8}$$

and

$$A = -\frac{i}{4}(|u|^2 + |v|^2)\sigma_3 - \frac{i\lambda}{2}\begin{pmatrix} 0 & \overline{v} \\ v & 0 \end{pmatrix} - \frac{i}{2\lambda}\begin{pmatrix} 0 & \overline{u} \\ u & 0 \end{pmatrix} + \frac{i}{4}\left(\lambda^2 + \frac{1}{\lambda^2}\right)\sigma_3. \tag{3.9}$$

The auto-Bäcklund transformation relates two solutions of the MTM system (2.1) while preserving the linear system (3.7). Now let us state the auto-Bäcklund transformation.

Proposition 2. *Let (u, v) be a C^1 solution of the MTM system (2.1) and $\vec{\phi} = (\phi_1, \phi_2)^t$ be a C^2 nonzero solution of the linear system (3.7) associated with the potential (u, v) and the spectral parameter $\lambda = \delta e^{i\gamma/2}$. Then, the following transformations*

$$\mathbf{u}(x, t) = -u(x, t)\frac{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2} + \frac{2i\delta^{-1}\sin\gamma\overline{\phi}_1\phi_2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2} \tag{3.10}$$

and

$$\mathbf{v}(x,t) = -v(x,t)\frac{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2}{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2} - \frac{2i\delta \sin\gamma \overline{\phi}_1 \phi_2}{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2}, \tag{3.11}$$

generates a new C^1 solution of the MTM system (2.1). Furthermore, the transformation

$$\psi_1 = \frac{\overline{\phi}_2}{|e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2|}, \quad \psi_2 = \frac{\overline{\phi}_1}{|e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2|} \tag{3.12}$$

yields a new C^2 nonzero solution $\vec{\psi} = (\psi_1, \psi_2)^t$ of the linear system (3.7) *associated with the new potential* $(\mathbf{u}, \mathbf{v})$ *and the same spectral parameter* λ.

Proof. Setting $\Gamma = \phi_1/\phi_2$ in the linear system (3.7) with Lax operators (3.8) and (3.9) yields the Riccati equations

$$\left\{ \begin{array}{l} \Gamma_x = 2i(\rho_2^2 - \rho_1^2)\Gamma + \frac{i}{2}(|u|^2 - |v|^2)\Gamma + i(\rho_2 v - \rho_1 u)\Gamma^2 - i(\rho_2 \overline{v} - \rho_1 \overline{u}), \\ \Gamma_t = 2i(\rho_2^2 + \rho_1^2)\Gamma - \frac{i}{2}(|u|^2 + |v|^2)\Gamma + i(\rho_2 v + \rho_1 u)\Gamma^2 - i(\rho_2 \overline{v} + \rho_1 \overline{u}), \end{array} \right. \tag{3.13}$$

where $\rho_1 = \frac{1}{2\lambda}$ and $\rho_2 = \frac{\lambda}{2}$. If we choose $\Gamma' := \frac{1}{\Gamma}$, $\mathbf{u} := M(\Gamma;\rho_1)f(\Gamma;u,\rho_1)$, and $\mathbf{v} := M(\Gamma;\rho_2)f(\Gamma;v,\rho_2)$ with

$$M(\Gamma;k) = -\frac{k|\Gamma|^2 + \overline{k}}{\overline{k}|\Gamma|^2 + k}, \quad f(\Gamma;q,k) = q + \frac{4i\mathrm{Im}(k^2)\overline{\Gamma}}{k|\Gamma|^2 + \overline{k}},$$

then the Riccati equations (3.13) remain invariant in variables Γ', $\mathbf{u}$, and $\mathbf{v}$. The transformation formulas above yield representation (3.10) and (3.11). Note that if $\vec{\phi} = \vec{0}$ at one point (x_0, t_0), then $\vec{\phi} = \vec{0}$ for all (x,t). If (u,v) is C^1 in (x,t), $\vec{\phi}$ is C^2 in (x,t), and $\vec{\phi} \neq \vec{0}$, then $(\mathbf{u}, \mathbf{v})$ is C^1 for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$.

The validity of (3.12) has been verified with Wolfram's Mathematica. Again, if $\vec{\phi}$ is C^2 in (x,t) and $\vec{\phi} \neq \vec{0}$, then $\vec{\psi}$ is C^2 and $\vec{\psi} \neq \vec{0}$ for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$. □

Let us denote the transformations (3.10)–(3.11) by $\mathcal{B}$, hence

$$\mathcal{B} : (u, v, \vec{\phi}, \lambda) \mapsto (\mathbf{u}, \mathbf{v}),$$

where $\vec{\phi}$ is a corresponding vector of the linear system (3.7) associated with the potential (u,v) and the spectral parameter λ.

In the simplest example, the MTM soliton (3.3) is recovered by the transformations (3.10) and (3.11) from the zero solution $(u,v) = (0,0)$, that is,

$$\mathcal{B} : (0, 0, \vec{\phi}, \lambda) \mapsto (u_\lambda, v_\lambda).$$

Indeed, a solution satisfying the linear system (3.7) with $(u,v) = (0,0)$ is given by

$$\left\{ \begin{array}{l} \phi_1 = e^{\frac{i}{4}(\lambda^2 - \lambda^{-2})x + \frac{i}{4}(\lambda^2 + \lambda^{-2})t}, \\ \phi_2 = e^{-\frac{i}{4}(\lambda^2 - \lambda^{-2})x - \frac{i}{4}(\lambda^2 + \lambda^{-2})t}. \end{array} \right. \tag{3.14}$$

Substituting this expression into (3.10) and (3.11) yields $(\mathbf{u}, \mathbf{v}) = (u_\lambda, v_\lambda)$ given by (3.3).

Another important example is a transformation from the MTM solitons (3.3) to the zero solution. We shall only give the explicit expressions of this transformation for the case $|\lambda| = \delta = 1$. By (3.12) and (3.14), we can find the vector $\vec{\psi}$ solving the linear system (3.7) with (u_λ, v_λ) given by (3.3). When $\lambda = e^{i\gamma/2}$, the vector $\vec{\psi}$ is given by

$$\begin{cases} \psi_1 = e^{\frac{1}{2}x\sin\gamma + \frac{i}{2}t\cos\gamma}\left|\operatorname{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|, \\ \psi_2 = e^{-\frac{1}{2}x\sin\gamma - \frac{i}{2}t\cos\gamma}\left|\operatorname{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|. \end{cases} \tag{3.15}$$

We note that $\vec{\psi}$ has exponential decay as $|x| \to \infty$ and, therefore, it is an eigenvector of the linear system (3.7) for the eigenvalue $\lambda = e^{i\gamma/2}$. Substituting the eigenvector $\vec{\psi}$ into the transformation (3.10) and (3.11), we obtain the zero solution from the MTM soliton, that is,

$$\mathcal{B} : (u_\lambda, v_\lambda, \vec{\psi}, \lambda) \mapsto (0, 0).$$

When $|\lambda| = \delta = 1$ for (u_λ, v_λ) given by (3.3), we realize that $c = 0$ and hence the MTM solitons (3.3) are stationary. Travelling MTM solitons with $c \neq 0$ can be recovered from the stationary MTM solitons with $c = 0$ by the Lorentz transformation. Hence, without loss of generality, we can choose $\lambda_0 = e^{i\gamma_0/2}$ for a fixed $\gamma_0 \in (0, \pi)$ in Theorem 3. Let us state the Lorentz transformation, which can be verified with the direct substitutions.

Proposition 3. *Let (u, v) be a solution of the MTM system (2.1) and let $\vec{\phi}$ be a solution of the linear system (3.7) associated with (u, v) and $\lambda = e^{i\gamma/2}$. Then,*

$$\begin{cases} u'(x,t) := \delta^{-1}u(k_1x + k_2t, k_1t + k_2x), \\ v'(x,t) := \delta v(k_1x + k_2t, k_1t + k_2x), \end{cases} \quad k_1 := \frac{\delta^2 + \delta^{-2}}{2}, \quad k_2 := \frac{\delta^2 - \delta^{-2}}{2}, \tag{3.16}$$

is a new solution of the MTM system (2.1), whereas

$$\vec{\phi}'(x,t) := \vec{\phi}(k_1x + k_2t, k_1t + k_2x), \tag{3.17}$$

is a new solution of the linear system (3.7) associated with (u', v') and $\lambda = \delta e^{i\gamma/2}$.

We shall denote the stationary MTM solitons at $t = 0$ as

$$\begin{cases} u_\gamma(x) = i\sin\gamma\operatorname{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right), \\ v_\gamma(x) = -i\sin\gamma\operatorname{sech}\left(x\sin\gamma + i\frac{\gamma}{2}\right), \end{cases} \tag{3.18}$$

that depend on the parameter $\gamma \in (0, \pi)$.

Let us now describe our method for the proof of Theorem 3. First we clarify some notations: $(u_{\gamma_0}, v_{\gamma_0})$ denotes one-soliton solution given by (3.18) with a fixed $\gamma_0 \in (0, \pi)$, $\vec{\psi}_{\gamma_0}$ denotes the corresponding eigenvector given by (3.15) for $t = 0$, whereas $L(u, v, \lambda)$ and $A(u, v, \lambda)$ denote the Lax operators L and A that contain (u, v) and a spectral parameter λ.

The main steps for the proof of Theorem 3 are the following. First, we fix an initial data $(u_0, v_0) \in H^2(\mathbb{R})$ such that (u_0, v_0) is sufficiently close to $(u_{\gamma_0}, v_{\gamma_0})$ in L^2-norm, according to the bound (3.4).

Step 1: From a perturbed one-soliton solution to a small solution at $t = 0$. In this step, we need to study the vector solution $\vec{\psi}$ of the linear equation

$$\partial_x \vec{\psi} = L(u_0, v_0, \lambda)\vec{\psi} \quad \text{at time } t = 0. \tag{3.19}$$

In addition to proving the existence of an exponentially decaying solution $\vec{\psi}$ of the linear equation (3.19) for an eigenvalue λ, we need to prove that if (u_0, v_0) is close to $(u_{\gamma_0}, v_{\gamma_0})$ in L^2-norm, then $\vec{\psi}$ is close to $\vec{\psi}_{\gamma_0}$ in H^1-norm and λ is close to $e^{i\gamma_0/2}$. Parameter λ in bound (3.5) is now determined by the eigenvalue of the linear equation (3.19).

The earlier example of obtaining the zero solution from the one-soliton solution gives a good insight that the auto-Bäcklund transformation given by Proposition 2 produces a function (p_0, q_0) at $t = 0$,

$$\mathcal{B} : (u_0, v_0, \vec{\psi}, \lambda) \mapsto (p_0, q_0), \tag{3.20}$$

such that (p_0, q_0) is small in L^2-norm. Moreover, if $(u_0, v_0) \in H^2(\mathbb{R})$, then $(p_0, q_0) \in H^2(\mathbb{R})$.

Step 2: Time evolution of the transformed solution. By the standard well-posedness theory for Dirac equations [32, 84, 98], there exists a unique global solution $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ to the MTM system (2.1) such that $(p, q)|_{t=0} = (p_0, q_0)$. Thanks to the L^2-conservation (3.2), the solution $(p(\cdot, t), q(\cdot, t))$ remains small in the L^2-norm for every $t \in \mathbb{R}$.

Step 3: From a small solution to a perturbed one-soliton solution for all times $t \in \mathbb{R}$. In this step, we are interested in the existence problem of the vector function $\vec{\phi}$ that solves the linear system

$$\partial_x \vec{\phi} = L(p, q, \lambda)\vec{\phi}, \quad \partial_t \vec{\phi} = A(p, q, \lambda)\vec{\phi} \tag{3.21}$$

where $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ is the unique global solution to the MTM system (2.1) starting with the initial data $(p, q)|_{t=0} = (p_0, q_0)$ in $H^2(\mathbb{R})$. Using the vector $\vec{\phi}$ and the auto-Bäcklund transformation given by Proposition 2, we obtain a new solution (u, v) to the MTM system (2.1),

$$\mathcal{B} : (p, q, \vec{\phi}, \lambda) \mapsto (u, v). \tag{3.22}$$

Moreover, if $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$, then $(u, v) \in C(\mathbb{R}; H^2(\mathbb{R}))$. Some translational parameter a and θ arise as functions of time t in the construction of the most general solution of the linear equation $\partial_x \vec{\phi} = L(p, q, \lambda)\vec{\phi}$ in the system (3.21). Bound (3.6) on the solution (u, v) is found from the analysis of the auto–Bäcklund transformation (3.22).

To summarize, there are three key ingredients in our method: mapping of an

L^2-neighborhood of the one-soliton solution to that of the zero solution at $t = 0$, the L^2-conservation of the MTM system, and mapping of an L^2-neighborhood of the zero solution to that of the one-soliton solution for every $t \in \mathbb{R}$. As a result, if the initial data is sufficiently close to the one-soliton solution in L^2 according to the initial bound (3.4), then the solution of the MTM system remains close to the one-soliton solution in L^2 for all times according to the final bound (3.6). A schematic picture is as follows:

$$\begin{array}{ccc} (u_0, v_0) & \cdots\cdots\cdots & (u, v) \\ \Big\downarrow \mathcal{B} & & \mathcal{B} \Big\uparrow \\ (p_0, q_0) & \longrightarrow & (p, q) \end{array}$$

Finally, we can remove the technical assumption that $(u_0, v_0) \in H^2(\mathbb{R})$ by an approximation argument in $L^2(\mathbb{R})$. This is possible because the MTM system (2.1) is globally well-posed in $L^2(\mathbb{R})$ by Theorem 2, whereas the bounds (3.5) and (3.6) are found to be uniform for the sequence of approximating solutions of the MTM system (2.1), the initial data of which approximate (u_0, v_0) in $L^2(\mathbb{R})$.

We note that the solution (p, q) to the MTM system (2.1) in a L^2-neighborhood of the zero solution could contain some L^2-small MTM solitons, which are related to the discrete spectrum of the spectral problem (3.19). Sufficient conditions for the absence of the discrete spectrum were derived in [84], and the L^2 smallness of the initial data is not generally sufficient for excluding eigenvalues of the discrete spectrum. If the small solitons occur in the Cauchy problem associated with the MTM system (2.1), asymptotic decay of solutions (u, v) to the MTM solitons given by (3.3) can not be proved, in other words, (p, q) do not decay to $(0, 0)$ in L^∞-norm as $t \to \infty$. Therefore, a more restrictive hypothesis on the initial data is generally needed to establish asymptotic stability of MTM solitons. See [24] for restrictions on initial data of the cubic NLS equation required in the proof of asymptotic stability of NLS solitons.

We also note that modulation equations for parameters a and θ in Theorem 3 are not included in our method. This can be viewed as an advantage of the auto-Bäcklund transformation, which does not rely on the global control of the dynamics of a and θ by means of the modulation equations. Values of a and θ are related to arbitrary constants that appear in the construction of $\vec{\phi}$ as a solution of the linear equation $\partial_x \vec{\phi} = L(p, q, \lambda)\vec{\phi}$ in the system (3.21). These values are eliminated in the infimum norm stated in the orbital stability result (3.6) in Theorem 3.

3.3 From a perturbed one-soliton solution to a small solution

Here we use the auto-Bäcklund transformation given by Proposition 2 to transform a L^2-neighborhood of the one-soliton solution to that of the zero solution at $t = 0$. Let $(u_0, v_0) \in L^2(\mathbb{R})$ be the initial data of the MTM system (2.1) satisfying bound (3.4) for $\lambda_0 = e^{i\gamma_0/2}$. Let $\vec{\psi}$ be a decaying eigenfunction of the spectral problem

$$\partial_x \vec{\psi} = L(u_0, v_0, \lambda)\vec{\psi}, \tag{3.23}$$

for an eigenvalue λ. First, we show that under the condition (3.4), an eigenvector $\vec{\psi}$ always exists and λ is close to λ_0. Then, we write $\lambda = \delta e^{i\gamma/2}$ and define

$$p_0 := -u_0 \frac{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2} + \frac{2i\delta^{-1}\sin\gamma\overline{\psi}_1\psi_2}{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2} \tag{3.24}$$

and

$$q_0 := -v_0 \frac{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2} - \frac{2i\delta\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}. \tag{3.25}$$

We intend to show that (p_0, q_0) is small in L^2 norm.

When $(u_0, v_0) = (u_{\gamma_0}, v_{\gamma_0})$ and $\lambda = \lambda_0 = e^{i\gamma_0/2}$, the spectral problem (3.23) has exactly one decaying eigenvector $\vec{\psi}$ given by

$$\begin{cases} \psi_1 = e^{\frac{1}{2}x\sin\gamma_0}\left|\operatorname{sech}\left(x\sin\gamma_0 - i\frac{\gamma_0}{2}\right)\right|, \\ \psi_2 = e^{-\frac{1}{2}x\sin\gamma_0}\left|\operatorname{sech}\left(x\sin\gamma_0 - i\frac{\gamma_0}{2}\right)\right|. \end{cases} \tag{3.26}$$

The other linearly independent solution $\vec{\xi}$ of the spectral problem (3.23) is given by

$$\begin{cases} \xi_1 = e^{\frac{1}{2}x\sin\gamma_0}(e^{-2x\sin\gamma_0} - x\sin(2\gamma_0))|\operatorname{sech}\left(x\sin\gamma_0 - i\frac{\gamma_0}{2}\right)|, \\ \xi_2 = -e^{-\frac{1}{2}x\sin\gamma_0}(e^{2x\sin\gamma_0} + 2\cos\gamma_0 + x\sin(2\gamma_0))|\operatorname{sech}\left(x\sin\gamma_0 - i\frac{\gamma_0}{2}\right)|. \end{cases} \tag{3.27}$$

This solution grows exponentially as $|x| \to \infty$. Therefore, $\dim\ker(\partial_x - L(u_{\gamma_0}, v_{\gamma_0}, \lambda_0)) = 1$ for the kernel subspace of the L^2 space. For clarity, we denote the decaying eigenvector (3.26) by $\vec{\psi}_{\gamma_0}$.

When (u_0, v_0) is close to $(u_{\gamma_0}, v_{\gamma_0})$ in L^2-norm, we would like to construct a decaying solution $\vec{\psi}$ of the spectral problem (3.23), which is close to the eigenvector $\vec{\psi}_{\gamma_0}$. This is achieved in Lemma 8 below. To simplify analysis, we introduce a unitary transformation in the linear equation (3.23),

$$\vec{\psi} = \begin{bmatrix} f & 0 \\ 0 & \overline{f} \end{bmatrix} \vec{\phi}, \tag{3.28}$$

where $f(x) = e^{\frac{i}{4}\int_0^x (|u_0|^2 - |v_0|^2)dx}$ is well defined for any $(u_0, v_0) \in L^2(\mathbb{R})$. Then, the linear equation (3.23) becomes

$$\partial_x \vec{\phi} = M(u_0, v_0, \lambda)\vec{\phi}, \tag{3.29}$$

where

$$M(u_0, v_0, \lambda) := \frac{i}{4}\begin{bmatrix} \lambda^2 - \lambda^{-2} & 2(\overline{u}_0\lambda^{-1} - \overline{v}_0\lambda)\overline{f}^2 \\ 2(u_0\lambda^{-1} - v_0\lambda)f^2 & \lambda^{-2} - \lambda^2 \end{bmatrix}.$$

The following lemma gives the main result of the perturbation theory. Below, $A \lesssim B$ means that there exists a positive constant C independent of ϵ such that

$A \leq CB$ for all sufficiently small ϵ.

Lemma 8. *For a fixed $\lambda_0 = e^{i\gamma_0/2}$ with $\gamma_0 \in (0,\pi)$, there exist a real positive ϵ such that if*

$$\|u_0 - u_{\gamma_0}\|_{L^2} + \|v_0 - v_{\gamma_0}\|_{L^2} \leq \epsilon, \tag{3.30}$$

then there exists a solution $\vec{\psi} \in H^1(\mathbb{R};\mathbb{C}^2)$ of the linear equation (3.23) for $\lambda \in \mathbb{C}$ satisfying the bound

$$|\lambda - \lambda_0| + \|\vec{\psi} - \vec{\psi}_{\gamma_0}\|_{H^1} \lesssim \|u_0 - u_{\gamma_0}\|_{L^2} + \|v_0 - v_{\gamma_0}\|_{L^2}. \tag{3.31}$$

Proof. Set $u_0 = u_{\gamma_0} + u_s$ and $v_0 = v_{\gamma_0} + v_s$, where $(u_s, v_s) \in L^2(\mathbb{R})$ are remainder terms, which are $\mathcal{O}(\epsilon)$ small in L^2 norm, according to the bound (3.30). We expand $1/\lambda^2$ and $1/\lambda$ around λ_0 and expand $u_0 f^2$ and $v_0 f^2$ in Taylor series, e.g.,

$$u_0 f^2 = u_0 e^{\frac{i}{2}\int_0^x (|u_0|^2 - |v_0|^2)dx} = (u_{\gamma_0} + u_s)\left(1 + g + \frac{1}{2}g^2 + \mathcal{O}(g^3)\right), \tag{3.32}$$

where

$$g := i\int_0^x \mathrm{Re}\,(u_s \overline{u}_{\gamma_0} - v_s \overline{v}_{\gamma_0})\,dx + \frac{i}{2}\int_0^x (|u_s|^2 - |v_s|^2)dx.$$

Note that g is well defined for $(u_s, v_s) \in L^2(\mathbb{R})$. From these expansions, the linear equation (3.29) becomes

$$(\partial_x - M_{\gamma_0})\vec{\phi} = \Delta M \vec{\phi}, \tag{3.33}$$

where

$$M_{\gamma_0} = M(u_{\gamma_0}, v_{\gamma_0}, \lambda_0) = \frac{1}{2}\begin{bmatrix} -\sin\gamma_0 & i(e^{-i\gamma_0/2}\overline{u}_{\gamma_0} - e^{i\gamma_0/2}\overline{v}_{\gamma_0}) \\ i(e^{-i\gamma_0/2}u_{\gamma_0} - e^{i\gamma_0/2}v_{\gamma_0}) & \sin\gamma_0 \end{bmatrix}$$

and the perturbation term ΔM applied to any $\vec{\phi} \in H^1(\mathbb{R})$ satisfies the inequality

$$\|\Delta M \vec{\phi}\|_{L^2} \lesssim (|\lambda - \lambda_0| + \|u_s\|_{L^2} + \|v_s\|_{L^2})\|\phi\|_{H^1}, \tag{3.34}$$

thanks to the embedding of $H^1(\mathbb{R})$ in $L^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$. Note that the bound (3.34) can not be derived in the context of the spectral problem (4.10) without the unitary transformation (3.28), which removes the term $\frac{i}{4}(|u|^2 - |v|^2)\sigma_3$ from the operator L in (1.10). This explains a posteriori why we are using the technical transformation (3.28).

We will later need the explicit computation of the leading order part in the perturbation term ΔM with respect to $(\lambda - \lambda_0)$, that is,

$$\Delta M = \frac{i}{2}(\lambda - \lambda_0)\begin{bmatrix} (\lambda_0 + \lambda_0^{-3}) & -(\overline{u}_{\gamma_0}\lambda_0^{-2} + \overline{v}_{\gamma_0}) \\ -(u_{\gamma_0}\lambda_0^{-2} + v_{\gamma_0}) & -(\lambda_0 + \lambda_0^{-3}) \end{bmatrix} + \mathcal{O}((\lambda - \lambda_0)^2, \|u_s\|_{L^2}, \|v_s\|_{L^2}). \tag{3.35}$$

We aim to construct an appropriate projection operator by which we split the linear equation (3.33) into two parts. Recall that $\dim\ker(\partial_x - M_{\gamma_0}) = 1$ and let $\vec{\phi}_{\gamma_0} \in \ker(\partial_x - M_{\gamma_0})$ and $\vec{\eta}_{\gamma_0} \in \ker(\partial_x + M_{\gamma_0}^*)$. These null vectors can be obtained explicitly:

$$\vec{\phi}_{\gamma_0} = \begin{bmatrix} e^{\frac{x}{2}\sin\gamma_0} \\ e^{-\frac{x}{2}\sin\gamma_0} \end{bmatrix} \left|\operatorname{sech}\left(\sin\gamma_0 x - i\frac{\gamma_0}{2}\right)\right|, \quad \vec{\eta}_{\gamma_0} = \begin{bmatrix} e^{-\frac{x}{2}\sin\gamma_0} \\ -e^{\frac{x}{2}\sin\gamma_0} \end{bmatrix} \left|\operatorname{sech}\left(\sin\gamma_0 x - i\frac{\gamma_0}{2}\right)\right|.$$

We note that $\langle \vec{\eta}_{\gamma_0}, \vec{\phi}_{\gamma_0}\rangle_{L^2} = 0$ but $\langle \sigma_3\vec{\eta}_{\gamma_0}, \vec{\phi}_{\gamma_0}\rangle_{L^2} \neq 0$, where $\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Also note that $\vec{\phi}_{\gamma_0} = \vec{\psi}_{\gamma_0}$ given by (3.26) because $|u_{\gamma_0}| = |v_{\gamma_0}|$. We make the following decomposition:

$$\vec{\phi} = \vec{\phi}_{\gamma_0} + \vec{\phi}_s, \quad \langle \sigma_3\vec{\eta}_{\gamma_0}, \vec{\phi}_s\rangle_{L^2} = 0. \tag{3.36}$$

To deal with the existence of such decomposition (3.36), we introduce the projection operator $P_{\gamma_0} : L^2(\mathbb{R};\mathbb{C}^2) \to L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\sigma_3\vec{\eta}_{\gamma_0}\}^\perp$ defined by

$$P_{\gamma_0}\vec{\phi} = \vec{\phi} - \frac{\langle \sigma_3\vec{\eta}_{\gamma_0}, \vec{\phi}\rangle_{L^2}}{\langle \sigma_3\vec{\eta}_{\gamma_0}, \vec{\phi}_{\gamma_0}\rangle_{L^2}}\vec{\phi}_{\gamma_0}.$$

Note that $P_{\gamma_0}\vec{\phi}_s = \vec{\phi}_s$ and $P_{\gamma_0}\vec{\phi}_{\gamma_0} = \vec{0}$. From equations (3.33) and (3.36), we define the operator equation

$$F(\vec{\phi}_s, u_s, v_s, \lambda) := (\partial_x - M_{\gamma_0})\vec{\phi}_s - \Delta M(\vec{\phi}_{\gamma_0} + \vec{\phi}_s) = 0. \tag{3.37}$$

Clearly, since $\dim\ker(\partial_x - M_{\gamma_0}) = 1 \neq 0$, the Fréchet derivative $D_{\vec{\phi}_s}F(0,0,0,\lambda_0) = \partial_x - M_{\gamma_0}$ has no bounded inverse. Let $\hat{P}_{\gamma_0} = \sigma_3 P_{\gamma_0}\sigma_3$ and notice that $\hat{P}_{\gamma_0} : L^2(\mathbb{R};\mathbb{C}^2) \to L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\vec{\eta}_{\gamma_0}\}^\perp$. We decompose equation (3.37) by the projection $\hat{P}$ into two equations

$$G(\vec{\phi}_s, u_s, v_s, \lambda) := \hat{P}_{\gamma_0}F(\vec{\phi}_s, u_s, v_s, \lambda) = 0, \tag{3.38}$$

$$H(\vec{\phi}_s, u_s, v_s, \lambda) := (I - \hat{P}_{\gamma_0})F(\vec{\phi}_s, u_s, v_s, \lambda) = 0. \tag{3.39}$$

Since $\dim\ker(\partial_x - M_{\gamma_0}) = \dim\ker(\partial_x + M_{\gamma_0}^*) = 1 < \infty$, then $\partial_x - M_{\gamma_0}$ is a Fredholm operator of index zero. Observe that $\operatorname{Range}(G) = L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\vec{\eta}_{\gamma_0}\}^\perp$, where $\vec{\eta}_{\gamma_0} \in \ker\{\partial_x + M_{\gamma_0}^*\}$. By the Fredholm alternative theorem, $\hat{P}_{\gamma_0}(\partial_x - M_{\gamma_0})^{-1}P_{\gamma_0}$ is a bounded operator:

$$P_{\gamma_0}(\partial_x - M_{\gamma_0})^{-1}\hat{P}_{\gamma_0} : L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\vec{\eta}_{\gamma_0}\}^\perp \to H^1(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\sigma_3\vec{\eta}_{\gamma_0}\}^\perp. \tag{3.40}$$

We can write equation (3.38) as

$$(I - P_{\gamma_0}(\partial_x - M_{\gamma_0})^{-1}\hat{P}_{\gamma_0}\Delta M)\vec{\phi}_s = P_{\gamma_0}(\partial_x - M_{\gamma_0})^{-1}\hat{P}_{\gamma_0}\Delta M\vec{\phi}_{\gamma_0}. \tag{3.41}$$

The operator in (3.41) is shown to be invertible by the Neumann series for operator and there exists a unique solution $\vec{\phi}_* = \vec{\phi}_s \in H^1$ for $(u_s, v_s, \lambda) \in U_\epsilon$ with sufficiently small ϵ, and furthermore $\vec{\phi}_*$ is a C^∞ function in λ, where $U_\epsilon = \{(u_s, v_s, \lambda) \in$

$[L^2(\mathbb{R})]^2 \times \mathbb{C} : \|(u_s, v_s)\|_{L^2} + |\lambda - \lambda_0| < \epsilon\}$. Furthermore, $\vec{\phi}_*$ satisfies the estimate

$$\begin{aligned} \|\vec{\phi}_*\|_{H^1} &\lesssim \|P_{\gamma_0}(\partial_x - M_{\gamma_0})^{-1}\hat{P}_{\gamma_0}\Delta M\vec{\phi}_{\gamma_0}\|_{L^2} \lesssim \|\Delta M\vec{\phi}_{\gamma_0}\|_{L^2} \\ &\lesssim |\lambda - \lambda_0| + \|u_s\|_{L^2} + \|v_s\|_{L^2}, \end{aligned} \tag{3.42}$$

if $(u_s, v_s, \lambda) \in U_\epsilon$.

Lastly we address the bifurcation equation (3.39) to determine $\lambda \in \mathbb{C}$. From equations (3.37) and (3.39), the bifurcation equation can be written explicitly as

$$I(u_s, v_s, \lambda) := \langle \vec{\eta}_{\gamma_0}, \Delta M(\vec{\phi}_{\gamma_0} + \vec{\phi}_*(u_s, v_s, \lambda))\rangle_{L^2} = 0, \tag{3.43}$$

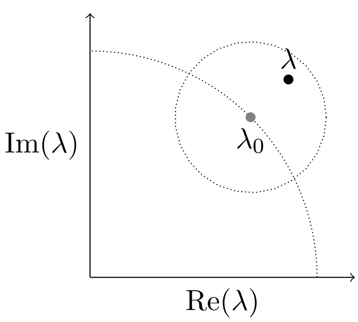
where $\vec{\phi}_*(u_s, v_s, \lambda)$ is uniquely expressed from (3.40) if $(u_s, v_s, \lambda) \in U_\epsilon$ and smooth in λ. By using the explicit expression (3.35),we check that $s := \partial_\lambda I(0, 0, \lambda_0) \neq 0$, where

$$\begin{aligned} s &= \frac{i}{2}\langle \vec{\eta}_{\gamma_0}, \begin{bmatrix} (\lambda_0 + \lambda_0^{-3}) & -(\overline{u}_{\gamma_0}\lambda_0^{-2} + \overline{v}_{\gamma_0}) \\ -(u_{\gamma_0}\lambda_0^{-2} + v_{\gamma_0}) & -(\lambda_0 + \lambda_0^{-3}) \end{bmatrix} \vec{\phi}_{\gamma_0}\rangle_{L^2} \\ &= ie^{-i\gamma_0/2} \int_{\mathbb{R}} \left(2\cos\gamma_0 \, |Q(x)|^2 + \sin^2\gamma_0 \, |Q(x)|^4\right) dx \\ &= 4ie^{-i\gamma_0/2} \int_{\mathbb{R}} \frac{1 + \cos\gamma_0 \cosh(2x\sin\gamma_0)}{(\cosh(2x\sin\gamma_0) + \cos\gamma_0)^2} dx \\ &= \frac{4ie^{-i\gamma_0/2}}{\sin\gamma_0}, \end{aligned}$$

where $Q(x) = \operatorname{sech}\left(x\sin\gamma_0 - i\frac{\gamma_0}{2}\right)$.

As a result, equation (3.43) can be used to uniquely determine the spectral parameter λ if $(u_s, v_s, \lambda) \in U_\epsilon$. From inequalities (3.35), (3.42) and (3.43), we obtain that this λ satisfies the bound $|\lambda - \lambda_0| \lesssim \|u_s\|_{L^2} + \|v_s\|_{L^2}$. □

Remark 2. *A spectral parameter λ in Lemma 8 may not be on the unit circle $|\lambda| = 1$ while $\lambda_0 = e^{i\gamma_0/2}$ is on the unit circle. Thus, a soliton corresponding to a spectral parameter λ may be a moving soliton in* (3.3) *at $t = 0$.*



Remark 3. *In what follows, we develop the theory when λ occurs on the unit circle, hence we write $\lambda = e^{i\gamma/2}$ for some $\gamma \in (0, \pi)$. All results obtained below can be generalized to the case of $|\lambda| \neq 1$ by using the Lorentz transformation in Proposition 3.*

In Lemma 11 below, we will show that a solution $\vec{\phi}$ determined in the proof of Lemma 8 can be written explicitly as the perturbed solution around $\vec{\phi}_\gamma$ in suitable function spaces. Then, in Lemma 12 below, we will use this representation and the auto-Bäcklund transformation (3.24) and (3.25) to show that (p_0, q_0) is small in L^2 norm.

To develop this analysis, we first prove several technical results. We consider the linear inhomogeneous equation

$$\partial_x \vec{w} - M_\gamma \vec{w} = \vec{f}, \tag{3.44}$$

where

$$M_\gamma = \frac{1}{2}\begin{bmatrix} -\sin\gamma & i(e^{-i\gamma/2}\overline{u}_\gamma - e^{i\gamma/2}\overline{v}_\gamma) \\ i(e^{-i\gamma/2}u_\gamma - e^{i\gamma/2}v_\gamma) & \sin\gamma \end{bmatrix}.$$

We introduce Banach spaces $X = X_1 \times X_2$ and $Y = Y_1 \times Y_2$ such that for $\vec{w} = (w_1, w_2)^t \in X$ and $\vec{f} = (f_1, f_2)^t \in Y$, we have

$$\|\vec{w}\|_X := \|w_1\|_{X_1} + \|w_2\|_{X_2}, \quad \|\vec{f}\|_Y := \|f_1\|_{Y_1} + \|f_2\|_{Y_2},$$

where

$$\begin{aligned} \|w_1\|_{X_1} &:= \inf_{w_1 = v_1 + u_1}\left(\left\| v_1 e^{-\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^\infty_x} \right. \\ &\qquad \left. + \left\| u_1 e^{\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x \cap L^\infty_x}\right), \\ \|w_2\|_{X_2} &:= \inf_{w_2 = v_2 + u_2}\left(\left\| v_2 e^{\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^\infty_x} \right. \\ &\qquad \left. + \left\| u_2 e^{-\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x \cap L^\infty_x}\right) \end{aligned}$$

and

$$\begin{aligned} \|f_1\|_{Y_1} &:= \inf_{f_1 = g_1 + h_1}\left(\left\| g_1 e^{\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x} \right. \\ &\qquad \left. + \left\| h_1 e^{-\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x \cap L^1_x}\right), \\ \|f_2\|_{Y_2} &:= \inf_{f_2 = g_2 + h_2}\left(\left\| g_2 e^{-\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x} \right. \\ &\qquad \left. + \left\| h_2 e^{\frac{x}{2}\sin\gamma}\left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|\right\|_{L^2_x \cap L^1_x}\right). \end{aligned}$$

It is obvious that X and Y are continuously embedded into $L^2(\mathbb{R})$. We shall estimate the bound of the operator $P_\gamma(\partial_x - M_\gamma)^{-1}\hat{P}_\gamma : Y \to X$, where projection operators P_γ and $\hat{P}_\gamma$ are defined in the proof of Lemma 8. First, we will obtain an explicit solution $\vec{w} \in H^1(\mathbb{R};\mathbb{C}^2) \cap \text{span}\{\sigma_3\vec{\eta}_\gamma\}^\perp$ for the linear inhomogeneous equation (3.44) when $\vec{f} \in L^2(\mathbb{R};\mathbb{C}^2) \cap \ker(\partial_x + M_\gamma^*)^\perp$. Then, we will prove that the mapping $Y \ni \vec{f} \mapsto \vec{w} \in X$ is bounded. These goals are achieved in the next two lemmas.

Lemma 9. *For any $\vec{f} = (f_1, f_2)^t \in L^2(\mathbb{R};\mathbb{C}^2) \cap \text{span}\{\vec{\eta}_\gamma\}^\perp$, there exists a unique solution $\vec{w} \in H^1(\mathbb{R};\mathbb{C}^2) \cap \text{span}\{\sigma_3\vec{\eta}_\gamma\}^\perp$ of the inhomogeneous equation (3.44) that can be written as*

$$\vec{w}(x) = \frac{1}{4}\vec{\phi}_\gamma(x)\left[k(\vec{f}) + W_-(x) + W_+(x)\right] + \frac{1}{4}\vec{\xi}_\gamma(x)\int_{-\infty}^{x}\vec{\eta}_\gamma(y)\cdot\vec{f}(y)dy, \quad (3.45)$$

where

$$\begin{aligned} W_-(x) &:= \int_{-\infty}^{x} e^{-\frac{1}{2}y\sin\gamma}(e^{2y\sin\gamma} + 2\cos\gamma + y\sin(2\gamma))\left|\text{sech}\left(y\sin\gamma - i\frac{\gamma}{2}\right)\right| f_1(y)dy, \\ W_+(x) &:= \int_{x}^{\infty} e^{-\frac{3}{2}y\sin\gamma}(-1 + e^{2y\sin\gamma}y\sin(2\gamma))\left|\text{sech}\left(y\sin\gamma - i\frac{\gamma}{2}\right)\right| f_2(y)dy, \end{aligned}$$

and $k(\vec{f})$ is a continuous linear functional on $L^2(\mathbb{R};\mathbb{C}^2)$.

Proof. Since $\partial_x - M_\gamma : H^1(\mathbb{R};\mathbb{C}^2) \to L^2(\mathbb{R};\mathbb{C}^2)$ is a Fredholm operator of index zero and $\ker(\partial_x + M_\gamma^*) = \text{span}\{\vec{\eta}_\gamma\}$, the inhomogeneous equation (3.44) has a solution in $H^1(\mathbb{R};\mathbb{C}^2)$ if and only if $\vec{f} \in L^2(\mathbb{R};\mathbb{C}^2) \cap \text{span}\{\vec{\eta}_\gamma\}^\perp$. For uniqueness, we add the constraint $\vec{w} \in \text{span}\{\sigma_3\vec{\eta}_\gamma\}^\perp$.

Recall that $U = [\vec{\phi}_\gamma, \vec{\xi}_\gamma]$ is a fundamental matrix of the homogeneous equation $(\partial_x - M_\gamma)U = 0$ and $\vec{\eta}_\gamma$ is a decaying solution of $(\partial_x + M_\gamma^*)\vec{\eta} = \vec{0}$. All functions are known explicitly as

$$\vec{\phi}_\gamma(x) = \begin{bmatrix} e^{\frac{1}{2}x\sin\gamma} \\ e^{-\frac{1}{2}x\sin\gamma} \end{bmatrix} Q(x), \quad \vec{\eta}_\gamma(x) = \begin{bmatrix} e^{-\frac{1}{2}x\sin\gamma} \\ -e^{\frac{1}{2}x\sin\gamma} \end{bmatrix} Q(x),$$

and

$$\vec{\xi}_\gamma(x) = \begin{bmatrix} e^{\frac{1}{2}x\sin\gamma}(e^{-2x\sin\gamma} - x\sin(2\gamma)) \\ -e^{-\frac{1}{2}x\sin\gamma}(e^{2x\sin\gamma} + 2\cos\gamma + x\sin(2\gamma)) \end{bmatrix} Q(x),$$

where

$$Q(x) := \left|\text{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|.$$

From variation of parameters, we have the explicit representation (3.45), where $k(\vec{f})$ is the constant of integration and the other constant is set to zero to ensure that $\vec{w} \in H^1(\mathbb{R};\mathbb{C}^2)$. It remains to prove that every term in the explicit expression (3.45) belongs to $L^2(\mathbb{R};\mathbb{C}^2)$.

Since $|\vec{\phi}_\gamma(x)| \lesssim e^{-\frac{|x|}{2}\sin\gamma}$ and $|Q(x)| \lesssim e^{-|x|\sin\gamma}$ for all $x \in \mathbb{R}$, we have

$$\begin{aligned}
&\|W_-\vec{\phi}_\gamma\|_{L^2} \\
\lesssim &\left\| e^{-\frac{1}{2}|x|\sin\gamma} \int_{-\infty}^x e^{-\frac{1}{2}y\sin\gamma}(e^{2y\sin\gamma} + 2\cos\gamma + y\sin(2\gamma))Q(y)f_1(y)dy \right\|_{L^2_x} \\
\lesssim &\left\| \int_{-\infty}^x e^{\frac{1}{2}(y-x)\sin\gamma}|f_1(y)|dy \right\|_{L^2_x} + \left\| e^{-\frac{1}{2}|x|\sin\gamma} \int_{-\infty}^x e^{-\frac{1}{2}|y|\sin\gamma}(2+|y|)|f_1(y)|dy \right\|_{L^2_x} \\
\lesssim &\|\vec{f}\|_{L^2}.
\end{aligned}$$

and

$$\begin{aligned}
\|W_+\vec{\phi}_\gamma\|_{L^2} \lesssim &\left\| e^{-\frac{1}{2}|x|\sin\gamma} \int_x^\infty e^{-\frac{3}{2}y\sin\gamma}(-1 + e^{2y\sin\gamma}y\sin(2\gamma))Q(y)f_2(y)dy \right\|_{L^2_x} \\
\lesssim &\left\| \int_x^\infty e^{\frac{1}{2}(x-y)\sin\gamma}|f_2(y)|dy \right\|_{L^2_x} + \left\| e^{-\frac{1}{2}|x|\sin\gamma} \int_x^\infty e^{-\frac{1}{2}|y|\sin\gamma}|y||f_2(y)|dy \right\|_{L^2_x} \\
\lesssim &\|\vec{f}\|_{L^2},
\end{aligned}$$

where notation $\|f(x)\|_{L^2_x}$ is used in place of $\|f(\cdot)\|_{L^2}$. Since $\vec{f} \in L^2(\mathbb{R};\mathbb{C}^2) \cap \text{span}\{\vec{\eta}_\gamma\}^\perp$, then

$$\int_x^\infty \vec{\eta}_\gamma(y) \cdot \vec{f}(y)dy = -\int_{-\infty}^x \vec{\eta}_\gamma(y) \cdot \vec{f}(y)dy.$$

Using this equality, we can estimate the last term in the explicit expression (3.45) as follows

$$\begin{aligned}
&\left\| \vec{\xi}_\gamma(x) \int_{-\infty}^x \vec{\eta}_\gamma(y) \cdot \vec{f}(y)dy \right\|_{L^2_x} \\
\lesssim &\left\| e^{-\frac{1}{2}x\sin\gamma} \int_{-\infty}^x \vec{\eta}_\gamma(y) \cdot \vec{f}(y)dy \right\|_{L^2_x} + \left\| e^{\frac{1}{2}x\sin\gamma} \int_x^\infty \vec{\eta}_\gamma(y) \cdot \vec{f}(y)dy \right\|_{L^2_x} \\
\lesssim &\left\| \int_{-\infty}^x e^{\frac{1}{2}(y-x)\sin\gamma}|\vec{f}(y)|dy \right\|_{L^2_x} + \left\| \int_x^\infty e^{-\frac{1}{2}(y-x)\sin\gamma}|\vec{f}(y)|dy \right\|_{L^2_x} \\
\lesssim &\|\vec{f}\|_{L^2},
\end{aligned}$$

where $|\vec{f}|$ is the vector norm of the 2-vector $\vec{f}$. Since $\langle \sigma_3\vec{\eta}_\gamma, \vec{\phi}_\gamma \rangle_{L^2} \neq 0$, $k(\vec{f})$ is uniquely determined from the orthogonality condition $\langle \sigma_3\vec{\eta}_\gamma, \vec{w} \rangle_{L^2} = 0$. Since all other terms in (3.45) are in $L^2(\mathbb{R};\mathbb{C}^2)$, $k(\vec{f})$ is bounded for all $\vec{f} \in L^2(\mathbb{R};\mathbb{C}^2)$. Therefore, $k(\vec{f})$ is a continuous linear functional on $L^2(\mathbb{R};\mathbb{C}^2)$. □

Lemma 10. *Let $\vec{f} \in Y \cap \text{span}\{\vec{\eta}_{\gamma_0}\}^\perp$ and let $\vec{w}$ be a solution of the inhomogeneous equation (3.44) in Lemma 9. Then there is a $\vec{f}$-independent constant $C > 0$ such that $\|\vec{w}\|_X \leq C\|\vec{f}\|_Y$.*

Proof. The solution $\vec{w}$ is given by the explicit formula (3.45). We assume now that

$\vec{f}$ belongs to the exponentially weighted space Y and prove that $\vec{w}$ belongs to the exponentially weighted space X. Since $\|a\vec{\phi}_\gamma\|_X \leq 2\|a\|_{L^\infty}$ for any $a \in L^\infty(\mathbb{R})$, $k(\vec{f})$ is a continuous linear functional on $L^2(\mathbb{R};\mathbb{C}^2)$, and Y is embedded into $L^2(\mathbb{R};\mathbb{C}^2)$, we have

$$\|k(\vec{f})\vec{\phi}_{\gamma_0}\|_X \lesssim |k(\vec{f})| \lesssim \|\vec{f}\|_{L^2} \lesssim \|\vec{f}\|_Y.$$

The second term in (3.45) is estimated by

$$\begin{aligned}
&\|W_-\vec{\phi}_{\gamma_0}\|_X \\
&\lesssim \left\| \int_{-\infty}^x e^{-\frac{1}{2}y\sin\gamma}(e^{2y\sin\gamma} + 2\cos\gamma + y\sin(2\gamma))Q(y)f_1(y)dy \right\|_{L^\infty_x} \\
&\lesssim \inf_{f_1=g_1+h_1} \left(\|g_1 e^{-\frac{1}{2}x\sin\gamma}|Q(x)|\|_{L^1_x} + \|h_1 e^{\frac{1}{2}x\sin\gamma}|Q(x)|\|_{L^2_x} \right) \\
&\leq \|\vec{f}\|_{Y_1},
\end{aligned}$$

where $Q(x) = \cosh(x\sin\gamma - i\gamma/2)$.

Similarly, the third term in (3.45) is estimated by $\|W_+\vec{\phi}_\gamma\|_X \lesssim \|\vec{f}\|_{Y_2}$. The last term in (3.45) is estimated as follows:

$$\left\| \xi \int_{-\infty}^x \eta(y)\cdot\vec{f}(y)dy \right\|_X \leq N_1 + N_2 + N_3 + N_4,$$

where

$$\begin{aligned}
N_1 &= \left\| e^{-x\sin\gamma} \int_{-\infty}^x \eta(y)\cdot\vec{f}(y)dy \right\|_{L^\infty_x \cap L^2_x}, \\
N_2 &= \left\| x\sin(2\gamma) \int_{-\infty}^x \eta(y)\cdot\vec{f}(y)dy \right\|_{L^\infty_x}, \\
N_3 &= \left\| e^{x\sin\gamma} \int_x^\infty \eta(y)\cdot\vec{f}(y)dy \right\|_{L^\infty_x \cap L^2_x}, \\
N_4 &= \left\| (2\cos\gamma + x\sin(2\gamma)) \int_{-\infty}^x \eta(y)\cdot\vec{f}(y)dy \right\|_{L^\infty_x}.
\end{aligned}$$

Since $|\eta(x)| \lesssim e^{-\frac{|x|}{2}\sin\gamma_0}$ and $\|e^{\frac{|x|}{2}\sin\gamma}\vec{f}\|_{L^2_x} \lesssim \|\vec{f}\|_Y$ for all $x \in \mathbb{R}$, we have

$$\begin{aligned}
N_1 &\lesssim \left\| \int_{-\infty}^x e^{-(y-x)\sin\gamma} e^{\frac{|y|}{2}\sin\gamma}(|f_1| + |f_2|)dy \right\|_{L^\infty_x \cap L^2_x} \\
&\leq \|e^{\frac{|x|}{2}\sin\gamma} f_1\|_{L^2_x} + \|e^{\frac{|x|}{2}\sin\gamma} f_2\|_{L^2_x} \\
&\lesssim \|\vec{f}\|_Y.
\end{aligned}$$

The other terms N_2, N_3, and N_4 are estimated similarly. All together, we have justified the bound $\|\vec{w}\|_X \leq C\|\vec{f}\|_Y$ for a $\vec{f}$-independent positive constant C. □

Lemma 11. *Under the condition (3.30), assume that $\lambda = e^{i\gamma/2}$ is the eigenvalue of the spectral problem (4.10) for the eigenvector $\vec{\psi} \in H^1(\mathbb{R};\mathbb{C}^2)$ determined in*

Lemma 8. Then, the eigenvector can be written in the form (3.28) with

$$\vec{\phi}(x) = \begin{bmatrix} e^{\frac{1}{2}x\sin\gamma}(1+r_{11}(x)) + e^{-\frac{1}{2}x\sin\gamma}r_{12}(x) \\ e^{\frac{1}{2}x\sin\gamma}r_{21}(x) + e^{-\frac{1}{2}x\sin\gamma}(1+r_{22}(x)) \end{bmatrix} \left|\operatorname{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|, \tag{3.46}$$

where components r_{ij} for $1 \le i,j \le 2$ satisfy the bound

$$\|r_{11}\|_{L^\infty} + \|r_{12}\|_{L^2\cap L^\infty} + \|r_{21}\|_{L^2\cap L^\infty} + \|r_{22}\|_{L^\infty} \lesssim \|u_0 - u_\gamma\|_{L^2} + \|v_0 - v_\gamma\|_{L^2}. \tag{3.47}$$

Proof. Recall the projection operators $P_\gamma : L^2(\mathbb{R};\mathbb{C}^2) \to L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\sigma_3\vec{\eta}_\gamma\}^\perp$ and $\hat{P}_\gamma : L^2(\mathbb{R};\mathbb{C}^2) \to L^2(\mathbb{R};\mathbb{C}^2) \cap \operatorname{span}\{\vec{\eta}_\gamma\}^\perp$ introduced in the proof of Lemma 8. The existence of the eigenvector $\vec{\phi} \in H^1(\mathbb{R};\mathbb{C}^2)$ of the spectral problem (3.29) for the eigenvalue $\lambda = e^{i\gamma/2}$ has been established in Lemma 8. Therefore, we are using operators P_γ and $\hat{P}_\gamma$ to prove additional properties of the eigenvector $\vec{\phi}$.

Using the projection operator P_γ, we decompose $\vec{\phi} = \vec{\phi}_\gamma + \vec{\phi}_s$ and rewrite the spectral problem (3.29) in the form

$$(\partial_x - M_\gamma)\,\vec{\phi}_s = \Delta\widetilde{M}(\vec{\phi}_\gamma + \vec{\phi}_s), \tag{3.48}$$

where $\Delta\widetilde{M}$ is the anti-diagonal matrix that contains the perturbation terms $u_0 - u_\gamma$ and $v_0 - v_\gamma$ only. Because $\vec{\phi}_s \in H^1(\mathbb{R};\mathbb{C}^2)$ exists by Lemma 8, we realize that $\Delta\widetilde{M}(\vec{\phi}_\gamma + \vec{\phi}_s) = \hat{P}\Delta\widetilde{M}(\vec{\phi}_\gamma + \vec{\phi}_s)$, which yields equivalently the constraint

$$\langle \vec{\eta}_\gamma, \Delta\widetilde{M}(\vec{\phi}_\gamma + \vec{\phi}_s)\rangle_{L^2} = 0. \tag{3.49}$$

Therefore, we write the perturbed equation (3.48) in the form

$$\vec{\phi}_s = P_\gamma\,(\partial_x - M_\gamma)^{-1}\,\hat{P}_\gamma\Delta\widetilde{M}(\vec{\phi}_\gamma + \vec{\phi}_s). \tag{3.50}$$

Note that the operator $\hat{P}_\gamma$ applies to the sum of the two terms in the right-hand-side of (3.50) thanks to (3.49) and cannot be applied to each term separately.

Since $\Delta\widetilde{M}$ is anti-diagonal, for any $\vec{\zeta} = (\zeta_1, \zeta_2)^t \in X$, we have

$$\|\Delta\widetilde{M}\vec{\zeta}\|_Y = \|(\Delta\widetilde{M})_{1,2}\zeta_2\|_{Y_1} + \|(\Delta\widetilde{M})_{2,1}\zeta_1\|_{Y_2},$$

which is bounded as follows:

$$\|(\Delta\widetilde{M})_{1,2}\zeta_2\|_{Y_1} \lesssim (\|u_0 - u_\gamma\|_{L^2} + \|v_0 - v_\gamma\|_{L^2})\|\zeta_2\|_{X_2}, \tag{3.51}$$

$$\|(\Delta\widetilde{M})_{2,1}\zeta_1\|_{Y_2} \lesssim (\|u_0 - u_\gamma\|_{L^2} + \|v_0 - v_\gamma\|_{L^2})\|\zeta_1\|_{X_1}. \tag{3.52}$$

Bound (3.51) follows simply from

$$\begin{aligned}
\|(\Delta\widetilde{M})_{1,2}\zeta_2\|_{Y_1} &\leq \inf_{\zeta_2=\xi_2+\eta_2}\Big(\|(\Delta\widetilde{M})_{1,2}\xi_2 e^{\frac{x}{2}\sin\gamma}R(x)\|_{L^2_x} \\
&\qquad + \|(\Delta\widetilde{M})_{1,2}\eta_2 e^{-\frac{x}{2}\sin\gamma}R(x)\|_{L^2_x\cap L^1_x}\Big) \\
&\lesssim \|(\Delta\widetilde{M})_{1,2}\|_{L^2}\inf_{\zeta_2=\xi_2+\eta_2}\big(\|\xi_2 e^{\frac{x}{2}\sin\gamma}R(x)\|_{L^\infty_x} \\
&\qquad + \|\eta_2 e^{-\frac{x}{2}\sin\gamma}R(x)\|_{L^\infty_x\cap L^2_x}\big) \\
&= \|(\Delta\widetilde{M})_{1,2}\|_{L^2}\|\zeta_2\|_{X_2},
\end{aligned}$$

where $R(x) = \left|\cosh\left(x\sin\gamma - i\frac{\gamma}{2}\right)\right|$. Bound (3.52) is obtained similarly. Because $\vec{\phi}_\gamma \in X$, the bound $\|\vec{w}\|_X \lesssim \|\vec{f}\|_Y$ in Lemma 10 and bounds (3.51) and (3.52) imply

$$\begin{aligned}
\|P_\gamma(\partial_x - M_\gamma)^{-1}\hat{P}_\gamma\Delta\widetilde{M}(\vec{\phi}_\gamma+\vec{\phi}_s)\|_X &\lesssim \|\Delta\widetilde{M}(\vec{\phi}_\gamma+\vec{\phi}_s)\|_Y \\
&\lesssim (\|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2})(1+\|\vec{\phi}_s\|_X).
\end{aligned}$$

Since $\|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}$ is sufficiently small, the component $\vec{\phi}_s$ in (3.50) satisfies the bound

$$\|\vec{\phi}_s\|_X \lesssim \|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}. \tag{3.53}$$

This completes the proof of the bound (3.47) in the representation (3.46), because the bound (3.53) on $\vec{\phi}_s$ in Banach space X yields the bounds on the components r_{ij} in the corresponding spaces. □

Corollary 3. *In addition to the assumptions of Lemma 11, assume that* $(u_0,v_0)\in H^m(\mathbb{R})$ *for an integer* $m\geq 0$. *Then,* r_{ij} *for* $1\leq i,j\leq 2$ *defined by* (3.46) *are* C^m*-functions of* x.

Proof. The statement is proved for $m=0$ in Lemma 11, because r_{ij} are bounded functions according to the bound (3.47) and they are continuous functions since $\vec{\phi}\in H^1(\mathbb{R};\mathbb{C}^2)$.

For $m=1$, we differentiate the equation (3.48) with respect to x to get

$$(\partial_x - M_\gamma)\,\partial_x\vec{\phi}_s = \vec{r} + R\vec{\phi}_s + \Delta\widetilde{M}\partial_x\vec{\phi}_s, \tag{3.54}$$

where $\vec{r} := \partial_x(\Delta\widetilde{M}\vec{\phi}_\gamma)$ and $R := \partial_x(M_\gamma)+\partial_x(\Delta\widetilde{M})$. Recall that $\vec{\phi}_s\in X$ by Lemma 11. If $(u_0,v_0)\in H^1(\mathbb{R})$, then $\vec{r}+R\vec{\phi}_s\in Y$ according to the bounds

$$\begin{aligned}
\|\vec{r}\|_Y &\lesssim \|u_0-u_\gamma\|_{H^1}+\|v_0-v_\gamma\|_{H^1}, \\
\|R\vec{\phi}_s\|_Y &\lesssim (1+\|u_0-u_\gamma\|_{H^1}+\|v_0-v_\gamma\|_{H^1})\|\vec{\phi}_s\|_X.
\end{aligned}$$

From bootstrapping of solution of the linear equation (3.50), we have $\partial_x\vec{\phi}_s\in$

$H^1(\mathbb{R})$. Then, since $\vec{r} + R\vec{\phi}_s \in Y$, we have

$$\vec{r} + R\vec{\phi}_s + \Delta\widetilde{M}\partial_x\vec{\phi}_s = \hat{P}_\gamma(\vec{r} + R\vec{\phi}_s + \Delta\widetilde{M}\partial_x\vec{\phi}_s)$$

Therefore, we can write the derivative equation (3.54) in the form

$$\partial_x\vec{\phi}_s = P_\gamma \left(\partial_x - M_\gamma\right)^{-1} \hat{P}_\gamma(\vec{r} + R\vec{\phi}_s + \Delta\widetilde{M}\partial_x\vec{\phi}_s). \tag{3.55}$$

Using bounds (3.51) and (3.52) and the smallness of $\|u_0 - u_\gamma\|_{L^2} + \|v_0 - v_\gamma\|_{L^2}$, we obtain

$$\|\partial_x\vec{\phi}_s\|_X \lesssim \|\vec{r} + R\vec{\phi}_s\|_Y < \infty, \tag{3.56}$$

from which it follows that $\partial_x\vec{\phi}_s \in H^1(\mathbb{R}) \cap X$, hence $\partial_x r_{ij} \in C(\mathbb{R})$ for $1 \leq i, j \leq 2$. Note that the bound (3.47) does not hold for $\partial_x r_{ij}$ because $\|u_0 - u_\gamma\|_{H^1} + \|v_0 - v_\gamma\|_{H^1}$ may not be small.

For $m \geq 2$, we differentiate (3.48) m times and obtain the expression

$$\left(\partial_x - M_\gamma\right)\partial_x^m\vec{\phi}_s = \vec{r}_m + \Delta\widetilde{M}\partial_x^m\vec{\phi}_s, \tag{3.57}$$

where $\vec{r}_m := \partial_x^m(\Delta\widetilde{M}\vec{\phi}_\gamma) + [\partial_x^m, M_\gamma + \Delta\widetilde{M}]\vec{\phi}_s$ and we denote $[\partial_x, f]g = \partial_x(fg) - f\partial_x(g)$. We note that the term $[\partial_x^m, M_\gamma + \Delta\widetilde{M}]\vec{\phi}_s$ does not contain the m-th derivative of $\vec{\phi}_s$. By an induction similar to the case $m = 1$, we find that $\vec{r}_m \in Y$ according to the bound

$$\|\vec{r}_m\|_Y \lesssim \|u_0 - u_\gamma\|_{H^m} + \|v_0 - v_\gamma\|_{H^m}.$$

Hence if $(u_0, v_0) \in H^m(\mathbb{R})$, then $\partial_x^m\vec{\phi}_s \in H^1(\mathbb{R}) \cap X$, hence $\partial_x^m r_{ij} \in C(\mathbb{R})$ for $1 \leq i, j \leq 2$. □

Lemma 12. *Under the condition (3.30), assume that $\lambda = e^{i\gamma/2}$ is the eigenvalue of the spectral problem (4.10) for the eigenvector $\vec{\psi} \in H^1(\mathbb{R};\mathbb{C}^2)$ determined in Lemma 8 and define*

$$p_0 := -u_0\frac{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2} + \frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2}, \tag{3.58}$$

$$q_0 := -v_0\frac{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2} - \frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}. \tag{3.59}$$

Then, $(p_0, q_0) \in L^2(\mathbb{R})$ satisfy the bound

$$\|p_0\|_{L^2} + \|q_0\|_{L^2} \lesssim \|u_0 - u_\gamma\|_{L^2} + \|v_0 - v_\gamma\|_{L^2}. \tag{3.60}$$

If, in addition, $(u_0, v_0) \in H^m(\mathbb{R})$ for an integer $m \geq 1$, then $(p_0, q_0) \in H^m(\mathbb{R})$.

Proof. Let us rewrite equation (3.58) as

$$p_0 S = -u_0 + \frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}, \tag{3.61}$$

where S is a module-one factor given by

$$S := \frac{e^{i\gamma/2}|\psi_1|^2 + e^{-i\gamma/2}|\psi_2|^2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2}.$$

We use the representation (3.28) and (3.46) for the eigenvector $\vec{\psi}$. Substituting $\vec{\psi}$ into the second term of (3.61), we obtain

$$\begin{aligned} \frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2} &= \frac{2i\overline{f}^2\sin\gamma\left[1+\epsilon_1+\epsilon_2 e^{x\sin\gamma}+\epsilon_3 e^{-x\sin\gamma}\right]}{e^{x\sin\gamma - i\gamma/2}(1+\epsilon_4)+e^{-x\sin\gamma+i\gamma/2}(1+\epsilon_5)+\epsilon_6} \\ &= i\overline{f}^2\sin\gamma\,\mathrm{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)\left[1+\mathcal{O}(|\epsilon_1|+|\epsilon_4|+|\epsilon_5|+|\epsilon_6|)\right]+\mathcal{O}(|\epsilon_2|+|\epsilon_3|), \end{aligned}$$

where $f(x) = e^{\frac{i}{4}\int_0^x(|u_0|^2-|v_0|^2)dx}$ and we have defined

$$\begin{aligned} \epsilon_1 &:= \overline{r}_{11}+r_{22}+\overline{r}_{11}r_{22}+\overline{r}_{12}r_{21}, \\ \epsilon_2 &:= r_{21}+\overline{r}_{11}r_{21}, \\ \epsilon_3 &:= \overline{r}_{12}+\overline{r}_{12}r_{22}, \\ \epsilon_4 &:= r_{11}+\overline{r}_{11}+|r_{11}|^2+e^{i\gamma}|r_{21}|^2, \\ \epsilon_5 &:= r_{22}+\overline{r}_{22}+|r_{22}|^2+e^{-i\gamma}|r_{12}|^2, \\ \epsilon_6 &:= 2e^{-i\gamma/2}\mathrm{Re}(r_{12}+\overline{r}_{11}r_{12})+2e^{i\gamma/2}\mathrm{Re}(r_{21}+r_{21}\overline{r}_{22}). \end{aligned}$$

Bound (3.47) in Lemma 11 implies that

$$\|\epsilon_1\|_{L^\infty}+\|\epsilon_2\|_{L^\infty\cap L^2}+\|\epsilon_3\|_{L^\infty\cap L^2}+\|\epsilon_4\|_{L^\infty}+\|\epsilon_5\|_{L^\infty}+\|\epsilon_6\|_{L^\infty\cap L^2} \lesssim \|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}.$$

Since $u_\gamma(x) = i\sin\gamma\,\mathrm{sech}\left(x\sin\gamma - i\frac{\gamma}{2}\right)$ and $|f(x)| = 1$ for all $x\in\mathbb{R}$, we obtain

$$\left\|\frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2} - \overline{f}^2 u_\gamma\right\|_{L^2} \lesssim \|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}.$$

Applying the triangle inequality to the representation (3.61), we obtain

$$\begin{aligned} \|p_0\|_{L^2} = \|p_0 S\|_{L^2} &\leq \|u_0-\overline{f}^2u_\gamma\|_{L^2}+\left\|\frac{2i\sin\gamma\overline{\psi}_1\psi_2}{e^{-i\gamma/2}|\psi_1|^2 + e^{i\gamma/2}|\psi_2|^2} - \overline{f}^2 u_\gamma\right\|_{L^2} \\ &\lesssim \|u_0-\overline{f}^2u_\gamma\|_{L^2}+\|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}. \end{aligned}$$

By the Taylor series expansion of f and the triangle inequality, we obtain

$$\begin{aligned} \|u_0-\overline{f}^2u_\gamma\|_{L^2} &\leq \|u_0-u_\gamma\|_{L^2}+\|u_\gamma\|_{L^2}\|1-\overline{f}^2\|_{L^\infty} \\ &\lesssim \|u_0-u_\gamma\|_{L^2}+\|v_0-v_\gamma\|_{L^2}, \end{aligned}$$

which finally yields the bound (3.60) for $\|p_0\|_{L^2}$. The bound (3.60) for $\|q_0\|_{L^2}$ is obtained in exactly the same way.

Now if $(u_0,v_0)\in H^m(\mathbb{R})$ for an integer $m\geq 1$, we can differentiate equation (3.61) in x m times and use Corollary 3 to conclude that $(p_0,q_0)\in H^m(\mathbb{R})$. □

3.4 From a small solution to a perturbed one-soliton solution

Here we use the auto-Bäcklund transformation given by Proposition 2 to transform a sufficiently smooth solution of the MTM system (2.1) in a L^2-neighborhood of the zero solution to the one in a L^2-neighborhood of the one-soliton solution.

Let $(p_0, q_0) \in H^2(\mathbb{R})$ be the initial data for the MTM system (2.1), which is sufficiently small in L^2 norm. Let $\vec{\phi}$ be a solution of the linear equation

$$\partial_x \vec{\phi} = L(p_0, q_0, \lambda)\vec{\phi} \tag{3.62}$$

with $\lambda = e^{i\gamma/2}$. Two linearly independent solutions of the linear equation (3.62) are constructed in Lemma 13 below.

Now, let $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ be the unique global solution to the MTM system (2.1) such that $(p, q)|_{t=0} = (p_0, q_0)$. This solution exists in $H^2(\mathbb{R})$ by the global well-posedness theory for Dirac equations [32, 84, 98]. The time evolution of the vector function $\vec{\phi}$ in t for every $x \in \mathbb{R}$ is defined by the linear equation

$$\partial_t \vec{\phi} = A(p, q, \lambda)\vec{\phi} \tag{3.63}$$

for the same $\lambda = e^{i\gamma/2}$. Lemma 14 characterizes two linearly independent solutions of the linear equation (3.63) for every $t \in \mathbb{R}$.

Lastly, Lemma 15 constructs a new solution $(u, v) \in C(\mathbb{R}; H^2(\mathbb{R}))$ to the MTM system (2.1) in a L^2-neighborhood of the one-soliton solution from the auto–Bäcklund transformation involving (p, q) and $\vec{\phi}$ for every $t \in \mathbb{R}$. Let us introduce the following unitary matrices

$$M_1 = \mathrm{diag}(m_1, \overline{m}_1) \qquad \text{and} \qquad M_2 = \mathrm{diag}(\overline{m}_2, m_2), \tag{3.64}$$

where $m_1(x) := e^{\frac{i}{4}\int_{-\infty}^{x}(|p_0|^2-|q_0|^2)ds}$ and $m_2(x) := e^{\frac{i}{4}\int_{x}^{\infty}(|p_0|^2-|q_0|^2)ds}$. We make substitution

$$\vec{\phi}_1(x) = e^{-\frac{\sin\gamma}{2}x} M_1(x) \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \end{bmatrix} \quad \text{and} \quad \vec{\phi}_2(x) = e^{\frac{\sin\gamma}{2}x} M_2(x) \begin{bmatrix} \chi_1(x) \\ \chi_2(x) \end{bmatrix}, \tag{3.65}$$

into the linear equation (3.62) with $\lambda = e^{i\gamma/2}$ and obtain two boundary value problems:

$$\begin{cases} \varphi_1' = \frac{i}{2}(e^{-i\gamma/2}\overline{p}_0 - e^{i\gamma/2}\overline{q}_0)\overline{m}_1^2\varphi_2, \\ \varphi_2' = \frac{i}{2}(e^{-i\gamma/2}p_0 - e^{i\gamma/2}q_0)m_1^2\varphi_1 + \sin\gamma\varphi_2, \end{cases} \tag{3.66}$$

and

$$\begin{cases} \chi_1' = -\sin\gamma\chi_1 + \frac{i}{2}(e^{-i\gamma/2}\overline{p}_0 - e^{i\gamma/2}\overline{q}_0)m_2^2\chi_2, \\ \chi_2' = \frac{i}{2}(e^{-i\gamma/2}p_0 - e^{i\gamma/2}q_0)\overline{m}_2^2\chi_1, \end{cases} \tag{3.67}$$

subject to the boundary conditions

$$\begin{cases} \lim_{x\to-\infty} \varphi_1(x) = 1, \\ \lim_{x\to\infty} e^{-x\sin\gamma}\varphi_2(x) = 0, \end{cases} \quad \text{and} \quad \begin{cases} \lim_{x\to-\infty} e^{x\sin\gamma}\chi_1(x) = 0, \\ \lim_{x\to\infty} \chi_2(x) = 1. \end{cases} \tag{3.68}$$

The following lemma characterizes solutions of the boundary value problems (3.66), (3.67), and (3.68) if (p_0, q_0) is small in the L^2-norm.

Lemma 13. *There exists a real positive δ such that if $\|p_0\|_{L^2} + \|q_0\|_{L^2} \leq \delta$, then the boundary value problems* (3.66)*, (3.67), and (3.68) have unique solutions in the class*

$$(\varphi_1, \varphi_2) \in L^\infty(\mathbb{R}) \times (L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})), \quad \textit{and} \quad (\chi_1, \chi_2) \in (L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})) \times L^\infty(\mathbb{R}),$$

satisfying bounds

$$\|\varphi_1 - 1\|_{L^\infty} + \|\varphi_2\|_{L^2\cap L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \tag{3.69}$$

and

$$\|\chi_1\|_{L^2\cap L^\infty} + \|\chi_2 - 1\|_{L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2}. \tag{3.70}$$

Proof. The boundary value problem (3.66) and (3.68) can be written in the integral form

$$\begin{cases} \varphi_1(x) = T_1(\varphi_1, \varphi_2)(x) := 1 + \frac{i}{2}\int_{-\infty}^{x} \left[e^{-i\gamma/2}\overline{p}_0(y) - e^{i\gamma/2}\overline{q}_0(y)\right] \overline{m}_1^2(y)\varphi_2(y)dy, \\ \varphi_2(x) = T_2(\varphi_1, \varphi_2)(x) := -\frac{i}{2}\int_{x}^{\infty} e^{(x-y)\sin\gamma}\left[e^{-i\gamma/2}p_0(y) - e^{i\gamma/2}q_0(y)\right] m_1^2(y)\varphi_1(y)dy. \end{cases} \tag{3.71}$$

We introduce a Banach space $Z := L^\infty(\mathbb{R}) \times (L^2(\mathbb{R}) \cap L^\infty(\mathbb{R}))$ equipped with the norm

$$\|\vec{u}\|_Z := \|u_1\|_{L^\infty} + \|u_2\|_{L^\infty\cap L^2}$$

and show that $\vec{T} = (T_1, T_2)^t : Z \to Z$ is a contraction mapping. Using the Schwartz inequality, the Younge inequality, and the triangle inequality, we obtain for any $\vec{\varphi}, \widetilde{\vec{\varphi}} \in Z$,

$$\begin{aligned} &\|T_1(\varphi_1, \varphi_2) - T_1(\widetilde{\varphi}_1, \widetilde{\varphi}_2)\|_{L^\infty} \\ =&\sup_{x\in\mathbb{R}} \left| \frac{i}{2}\int_{-\infty}^{x} \left[e^{-i\gamma/2}\overline{p}_0(y) - e^{i\gamma/2}\overline{q}_0(y)\right] \overline{m}_1^2(y)(\varphi_2(y) - \widetilde{\varphi}_2(y))dy \right| \\ \leq&\frac{1}{2}\left(\|p_0\|_{L^2} + \|q_0\|_{L^2}\right)\|\varphi_2 - \widetilde{\varphi}_2\|_{L^2} \end{aligned}$$

and

$$\begin{aligned} &\|T_2(\varphi_1, \varphi_2) - T_2(\widetilde{\varphi}_1, \widetilde{\varphi}_2)\|_{L^\infty\cap L^2} \\ \leq&\frac{1}{2}\|e^{x\sin\gamma}\|_{L^1_x(\mathbb{R}_-)\cap L^2_x(\mathbb{R}_-)}\|e^{-i\gamma/2}p_0 - e^{i\gamma/2}q_0\|_{L^2}\|\varphi_1 - \widetilde{\varphi}_1\|_{L^\infty} \\ \leq&\frac{1}{\sin\gamma}\left(\|p_0\|_{L^2} + \|q_0\|_{L^2}\right)\|\varphi_1 - \widetilde{\varphi}_1\|_{L^\infty}. \end{aligned}$$

If $\|p_0\|_{L^2}+\|q_0\|_{L^2} \leq \delta$ is sufficiently small such that $\delta < \sin\gamma$ for a fixed $\gamma \in (0,\pi)$, then $\vec{T} = (T_1,T_2)^t$ is a contraction mapping on Z. To prove the inequality (3.69), we have

$$\begin{aligned} \|\varphi_1 - 1\|_{L^\infty} + \|\varphi_2\|_{L^2\cap L^\infty} &\leq \|\vec{T}(\varphi_1,\varphi_2) - \vec{T}(0,0)\|_Z \\ &\leq \frac{\|p_0\|_{L^2}+\|q_0\|_{L^2}}{\sin\gamma}(1+\|\varphi_1-1\|_{L^\infty}+\|\varphi_2\|_{L^\infty\cap L^2}). \end{aligned}$$

Since $\|p_0\|_{L^2}+\|q_0\|_{L^2} \leq \delta < \sin\gamma$, the above estimates yields the inequality (3.69). Repeating similar estimates for the boundary-value problem (3.67) and (3.68), we can prove that $(\chi_1,\chi_2) \in (L^2(\mathbb{R})\cap L^\infty(\mathbb{R}))\times L^\infty(\mathbb{R})$ and the inequality (3.70) holds. □

Remark 4. *Bounds* (3.69) *and* (3.70) *imply that a parameter* λ *is not an* L^2 *eigenvalue in the spectrum problem* (3.62).

Let us now define the time evolution of the vector functions $\vec{\phi}_1$ and $\vec{\phi}_2$ in t for every $x \in \mathbb{R}$, according to the linear equation (3.63), where $\lambda = e^{i\gamma/2}$ and $(p,q) \in C(\mathbb{R};H^2(\mathbb{R}))$ is the unique solution of the MTM system (2.1) such that $(p,q)|_{t=0} = (p_0,q_0)$. We also consider the initial data for $\vec{\phi}_1$ and $\vec{\phi}_2$ at $t=0$ given by the two linearly independent solutions (3.65) of the linear equation (3.62). The linear equation (3.63) for $\vec{\phi}_{1,2}$ with $\lambda = e^{i\gamma/2}$ take the form

$$\partial_t\vec{\phi}_{1,2} = \begin{bmatrix} -\frac{i}{4}(|p|^2+|q|^2)+\frac{i}{2}\cos\gamma & -\frac{i}{2}(e^{-i\gamma/2}\overline{p}+e^{i\gamma/2}\overline{q}) \\ -\frac{i}{2}(e^{-i\gamma/2}p+e^{i\gamma/2}q) & \frac{i}{4}(|p|^2+|q|^2)-\frac{i}{2}\cos\gamma \end{bmatrix}\vec{\phi}_{1,2}. \tag{3.72}$$

We set

$$\vec{\phi}_1(x,t) := e^{-\frac{x}{2}\sin\gamma}M_1(x,t)\vec{\varphi}(x,t), \quad \vec{\phi}_2(x,t) := e^{\frac{x}{2}\sin\gamma}M_2(x,t)\vec{\chi}(x,t), \tag{3.73}$$

where $M_1(x,t)$ and $M_2(x,t)$ are given by (3.64) with

$$m_1(x,t) := e^{\frac{i}{4}\int_{-\infty}^{x}(|p(s,t)|^2-|q(s,t)|^2)ds}, \quad m_2(x,t) := e^{\frac{i}{4}\int_{x}^{\infty}(|p(s,t)|^2-|q(s,t)|^2)ds}. \tag{3.74}$$

The following lemma characterizes vector functions $\vec{\varphi}$ and $\vec{\chi}$.

Lemma 14. *Let* $(p_0,q_0) \in H^2(\mathbb{R})$ *and assume that there exists a sufficiently small* δ *such that* $\|p_0\|_{L^2}+\|q_0\|_{L^2} \leq \delta$. *Let* $(p,q) \in C(\mathbb{R};H^2(\mathbb{R}))$ *be the unique solution of the MTM system (2.1) such that* $(p,q)|_{t=0} = (p_0,q_0)$. *Let* $\vec{\phi}_1$ *and* $\vec{\phi}_2$ *be solutions of the linear equation* (3.72) *starting with the initial data given by (3.65). Then, for every* $t \in \mathbb{R}$, $\vec{\phi}_1$ *and* $\vec{\phi}_2$ *are given by (3.73), where*

$(\varphi_1,\varphi_2)(\cdot,t) \in L^\infty(\mathbb{R})\times(L^2(\mathbb{R})\cap L^\infty(\mathbb{R}))$ and $(\chi_1,\chi_2)(\cdot,t) \in (L^2(\mathbb{R})\cap L^\infty(\mathbb{R}))\times L^\infty(\mathbb{R})$

satisfy the differential equations

$$\partial_x\vec{\varphi} = \begin{bmatrix} 0 & \frac{i}{2}(e^{-i\gamma/2}\overline{p}-e^{i\gamma/2}\overline{q})\bar{m}_1^2 \\ \frac{i}{2}(e^{-i\gamma/2}p-e^{i\gamma/2}q)m_1^2 & \sin\gamma \end{bmatrix}\vec{\varphi} \tag{3.75}$$

and

$$\partial_x \vec{\chi} = \begin{bmatrix} -\sin\gamma & \frac{i}{2}(e^{-i\gamma/2}\overline{p} - e^{i\gamma/2}\overline{q})m_2^2 \\ \frac{i}{2}(e^{-i\gamma/2}p - e^{i\gamma/2}q)\bar{m}_2^2 & 0 \end{bmatrix} \vec{\chi}, \tag{3.76}$$

subject to the boundary values

$$\begin{cases} \lim_{x\to-\infty} \varphi_1(x,t) = e^{\frac{i}{2}t\cos\gamma}, \\ \lim_{x\to\infty} e^{-x\sin\gamma}\varphi_2(x,t) = 0, \end{cases} \quad \text{and} \quad \begin{cases} \lim_{x\to-\infty} e^{x\sin\gamma}\chi_1(x,t) = 0, \\ \lim_{x\to\infty} \chi_2(x,t) = e^{-\frac{i}{2}t\cos\gamma}. \end{cases} \tag{3.77}$$

Furthermore, for every $t \in \mathbb{R}$, these functions satisfy the bounds

$$\|\varphi_1(\cdot,t) - e^{\frac{i}{2}t\cos\gamma}\|_{L^\infty} + \|\varphi_2(\cdot,t)\|_{L^2\cap L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \tag{3.78}$$

and

$$\|\chi_1(\cdot,t)\|_{L^2\cap L^\infty} + \|\chi_2(\cdot,t) - e^{-\frac{i}{2}t\cos\gamma}\|_{L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2}. \tag{3.79}$$

Proof. By Sobolev embedding of $H^2(\mathbb{R})$ into $C^1(\mathbb{R})$, the x-derivatives of solutions $(p,q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ are continuous and bounded functions of x for every $t \in \mathbb{R}$. Moreover, bootstrapping arguments for the MTM system (2.1) show that the same solution (p,q) exists in $C^1(\mathbb{R}; H^1(\mathbb{R}))$. Therefore, the t-derivatives of solutions (p,q) are also continuous and bounded functions of x for every $t \in \mathbb{R}$. Thus, the technical assumption $(p_0,q_0) \in H^2(\mathbb{R})$ simplifies working with the system of Lax equations (3.62) and (3.63). In particular, we shall prove that $\vec{\varphi}$ satisfies the differential equation (3.75) for every $t \in \mathbb{R}$ if $\vec{\phi}_1$ satisfies the differential equation (3.72) for every $x \in \mathbb{R}$ and the representation (3.73) is used.

By Lemma 13, $\vec{\varphi}$ is a bounded function of x for $t = 0$ and by bootstrapping arguments, $\vec{\varphi} \in C(\mathbb{R})$ for $t = 0$. We now claim that the differential equation (3.72) preserves this property for every $t \in \mathbb{R}$. From the differential equation (3.72) and the representation (3.73), we obtain

$$\begin{aligned} \partial_t(|\varphi_1|^2 + |\varphi_2|^2) &= \sin\left(\frac{\gamma}{2}\right)\left[(\bar{q}-\bar{p})\bar{m}_1^2\bar{\varphi}_1\varphi_2 + (q-p)m_1^2\varphi_1\bar{\varphi}_2\right] \\ &\leq (|p|+|q|)(|\varphi_1|^2 + |\varphi_2|^2). \end{aligned}$$

By Gronwall's inequality, for any $T > 0$, we obtain

$$|\varphi_1(x,t)|^2 + |\varphi_2(x,t)|^2 \leq e^{\alpha_T T}(|\varphi_1(x,0)|^2 + |\varphi_2(x,0)|^2) \quad x \in \mathbb{R}, \quad t \in [-T,T], \tag{3.80}$$

where

$$\alpha_T := \sup_{t\in[-T,T]} \sup_{x\in\mathbb{R}} \left(|p(x,t)| + |q(x,t)|\right).$$

Since the exponential factor remains bounded for any finite time $T > 0$, then it follows that $\vec{\varphi}(\cdot,t) \in L^\infty(\mathbb{R})$ for every $t \in \mathbb{R}$. Bootstrapping then yields $\vec{\varphi}(\cdot,t) \in C(\mathbb{R})$ for every $t \in \mathbb{R}$.

Since coefficients of the linear system (3.72) are continuous functions of (x,t), we have $\partial_t\vec{\varphi}(\cdot,t) \in C(\mathbb{R})$ for every $t \in \mathbb{R}$. Now, if (p,q) are C^1 functions of x and

t, then a similar method shows that $\partial_x\vec{\varphi}, \partial_t\partial_x\vec{\varphi}, \partial_t^2\vec{\varphi} \in C(\mathbb{R})$ for every $t \in \mathbb{R}$.

We shall now establish the validity of the differential equation (3.75). For $\vec{\phi}_1$ in (3.73), we write this equation in the abstract form $\partial_x\vec{\phi}_1 = L\vec{\phi}_1$. We also write the differential equation (3.72) for $\vec{\phi}_1$ in the abstract form $\partial_t\vec{\phi}_1 = A\vec{\phi}_1$. To establish (3.75) for every $t \in \mathbb{R}$, we construct the residual function $\vec{F} := \partial_x\vec{\phi}_1 - L\vec{\phi}_1$. This function is zero for every $x \in \mathbb{R}$ and $t = 0$. We shall prove that $\vec{F}$ is zero for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$.

The compatibility condition $\partial_x A - \partial_t L + [A, L] = 0$ is satisfied for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$, if (p, q) is a C^1 solution of the MTM system (2.1). After differentiating $\vec{F}$ with respect to t, we obtain

$$\begin{aligned}
\partial_t\vec{F} &= \partial_t\partial_x\vec{\phi}_1 - (\partial_t L)\vec{\phi}_1 - L\partial_t\vec{\phi}_1 \\
&= \partial_x(A\vec{\phi}_1) - (\partial_t L)\vec{\phi}_1 - LA\vec{\phi}_1 \\
&= (\partial_x A - \partial_t L + [A, L])\vec{\phi}_1 + A\vec{F} \\
&= A\vec{F}.
\end{aligned}$$

Let $\vec{F} = (F_1, F_2)^t$. From the linear evolution $\partial_t\vec{F} = AF$, we again obtain

$$\begin{aligned}
\partial_t(|F_1|^2 + |F_2|^2) &= \sin\left(\frac{\gamma}{2}\right)\left[(\bar{q} - \bar{p})\bar{F}_1 F_2 + (q - p)F_1\bar{F}_2\right] \\
&\leq (|p| + |q|)(|F_1|^2 + |F_2|^2),
\end{aligned}$$

which yields with Gronwall's inequality for any $T > 0$

$$|F_1(x,t)|^2 + |F_2(x,t)|^2 \leq e^{\alpha_T T}(|F_1(x,0)|^2 + |F_2(x,0)|^2), \quad x \in \mathbb{R}, \quad t \in [-T, T],$$

with the same definition of α_T. Since $\vec{F}(x, 0) = \vec{0}$, then the above inequality yields $\vec{F}(x,t) = \vec{0}$ for every $x \in \mathbb{R}$ and $t \in [-T, T]$. Hence, $\vec{\varphi}$ satisfies the differential equation (3.75).

We have shown that $\vec{\varphi}(\cdot, t) \in L^\infty(\mathbb{R})$ for every $t \in \mathbb{R}$. We now show that $\varphi_2(\cdot, t) \in L^2(\mathbb{R})$ for every $t \in \mathbb{R}$. It follows from the differential equation (3.72) and the representation (3.73) that

$$\begin{aligned}
\partial_t(|\varphi_2|^2) &\leq (|p| + |q|)|\bar{\varphi}_1\varphi_2| \\
&\lesssim |\varphi_2|^2 + (|p|^2 + |q|^2)|\varphi_1|^2.
\end{aligned}$$

Using Gronwall's inequality and the previous bound (3.80), we have for any $T > 0$

$$\begin{aligned}
&|\varphi_2(x,t)|^2 \\
&\leq e^T\left[|\varphi_2(x,0)|^2 + \int_{-T}^{T}(|p(x,s)|^2 + |q(x,s)|^2)|\varphi_1(x,s)|^2 ds\right] \\
&\leq e^T|\varphi_2(x,0)|^2 + e^{(1+\alpha_T)T}\int_{-T}^{T}(|p(x,s)|^2 + |q(x,s)|^2)\left(|\varphi_1(x,0)|^2 + |\varphi_2(x,0)|^2\right) ds,
\end{aligned}$$

where $x \in \mathbb{R}$ and $t \in [-T, T]$. Therefore, we have

$$\|\varphi_2(\cdot,t)\|_{L^2}^2 \leq e^T \|\varphi_2(\cdot,0)\|_{L^2}^2$$
$$+ e^{(1+\alpha_T)T} \left(\|\varphi_1(\cdot,0)\|_{L^\infty}^2 + \|\varphi_2(\cdot,0)\|_{L^\infty}^2 \right) \int_{-T}^{T} (\|p(\cdot,s)\|_{L^2}^2 + \|q(\cdot,s)\|_{L^2}^2) ds.$$

Since the right-hand side of this inequality remains bounded for any finite time $T > 0$, then it follows that $\varphi_2(\cdot,t) \in L^2(\mathbb{R})$ for every $t \in \mathbb{R}$.

It remains to prove the boundary values for $\vec{\varphi}_1(x,t)$ as $x \to \pm\infty$ in (3.77). The second boundary condition

$$\lim_{x\to\infty} e^{-x\sin\gamma} \varphi_2(x,t) = 0$$

follows from the fact that $\varphi_2(\cdot,t) \in L^\infty(\mathbb{R})$ for every $t \in \mathbb{R}$. To prove the first boundary condition, we use Duhamel's formula to write the differential equation (3.72) in the integral form:

$$\vec{\phi}_1(x,t) = e^{\frac{i}{2}t\sigma_3\cos\gamma} \vec{\phi}_1(x,0) + \int_0^t e^{\frac{i}{2}(t-s)\sigma_3\cos\gamma} A_1(x,s) \vec{\phi}_1(x,s) ds,$$

where

$$A_1(x,t) := \begin{bmatrix} -\frac{i}{4}(|p|^2 + |q|^2) & -\frac{i}{2}(e^{-i\gamma/2}\overline{p} + e^{i\gamma/2}\overline{q}) \\ -\frac{i}{2}(e^{-i\gamma/2}p + e^{i\gamma/2}q) & \frac{i}{4}(|p|^2 + |q|^2) \end{bmatrix}.$$

Using the representation (3.73), we have for $t \in \mathbb{R}$

$$|M_1\vec{\varphi}(x,t) - e^{\frac{i}{2}t\sigma_3\cos\gamma} M_1\vec{\varphi}(x,0)| \leq \int_0^{|t|} |A_1(x,s) M_1 \vec{\varphi}(x,s)| ds,$$

where $|\vec{f}|$ denotes the vector norm of the 2-vector $\vec{f}$. Since $\vec{\varphi}(\cdot,t) \in L^\infty(\mathbb{R}) \times (L^\infty(\mathbb{R}) \cap L^2(\mathbb{R}))$ for every $t \in \mathbb{R}$ and $p(\cdot,t), q(\cdot,t) \in H^2(\mathbb{R})$, we claim that

- $|A_1(x,s) M_1\vec{\varphi}(x,s)|$ is bounded by some s-independent constant for every $x \in \mathbb{R}$ and $|s| \leq |t|$
- $\lim_{|x|\to\infty} A_1(x,s) M_1 \vec{\varphi}(x,s) = \vec{0}$ pointwise for every $|s| \leq |t|$.

Then, the dominated convergence theorem gives

$$\lim_{x\to-\infty} |M_1(x,t)\vec{\varphi}(x,t) - e^{\frac{i}{2}t\sigma_3\cos\gamma} M_1(x,0)\vec{\varphi}(x,0)| = 0, \quad t \in \mathbb{R}.$$

Since $\vec{\varphi}(x,0) \to (1,0)^t$ as $x \to -\infty$ and $M_1(x,t) \to I$ as $x \to -\infty$ for every $t \in \mathbb{R}$, the above limit recovers the first boundary condition

$$\lim_{x\to-\infty} \varphi_1(x,t) = e^{\frac{i}{2}t\cos\gamma}.$$

The proof of the differential equation (3.76) and the boundary condition for $\vec{\chi}$ in (3.77) is analogous. Finally, since the L^2 norm of solutions of the MTM system

(2.1) is constant in time t, according to (3.2), the proof of bounds (3.78) and (3.79) is analogous to the proof in Lemma 13. $\square$

Lemma 15. *Let $(p_0, q_0) \in H^2(\mathbb{R})$ and assume that there exists a sufficiently small δ such that $\|p_0\|_{L^2} + \|q_0\|_{L^2} \leq \delta$. Let $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ be the unique solution to the MTM system (2.1) such that $(p, q)|_{t=0} = (p_0, q_0)$. Using solutions $\vec{\varphi}$ and $\vec{\chi}$ in Lemma 14, let us define*

$$\begin{bmatrix} \phi_1(x,t) \\ \phi_2(x,t) \end{bmatrix} := c_1(t) e^{-\frac{x}{2}\sin\gamma} M_1(x,t)\vec{\varphi}(x,t) + c_2(t) e^{\frac{x}{2}\sin\gamma} M_2(x,t)\vec{\chi}(x,t), \quad (3.81)$$

where $c_1(t) := e^{(a+i\theta)/2}$, $c_2(t) := e^{-(a+i\theta)/2}$ are given in terms of the real coefficients a, θ, which may depend on t. Then, the auto–Bäcklund transformation

$$u := -p\frac{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2} + \frac{2i\sin\gamma\overline{\phi}_1\phi_2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2} \quad (3.82)$$

and

$$v := -q\frac{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2}{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2} - \frac{2i\sin\gamma\overline{\phi}_1\phi_2}{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2} \quad (3.83)$$

generates a new solution $(u, v) \in C(\mathbb{R}; H^2(\mathbb{R}))$ to the MTM system (2.1) satisfying the bound

$$\left\| u(x,t) - ie^{-i\theta - it\cos\gamma}\sin\gamma\,\mathrm{sech}\left(x\sin\gamma - i\frac{\gamma}{2} - a\right)\right\|_{L^2_x} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \quad (3.84)$$

and

$$\left\| v(x,t) + ie^{-i\theta - it\cos\gamma}\sin\gamma\,\mathrm{sech}\left(x\sin\gamma + i\frac{\gamma}{2} - a\right)\right\|_{L^2_x} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \quad (3.85)$$

for every $t \in \mathbb{R}$.

Proof. Let us introduce $\vec{\psi} = (\psi_1, \psi_2)^t$ by

$$\psi_1 := \frac{\overline{\phi}_2}{|e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2|}, \quad \psi_2 := \frac{\overline{\phi}_1}{|e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2|}. \quad (3.86)$$

The inequalities (3.78) and (3.79) imply that (u, v) and $\vec{\psi}$ are bounded for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$. If (p, q) are C^1 functions of (x, t) and $\vec{\phi}$ is a C^2 function of (x, t), then (u, v) are C^1 functions of (x, t) and $\vec{\psi}$ is a C^2 function of (x, t). Proposition 2 states that $\vec{\psi}$ given by (3.86) satisfies the evolution equations

$$\partial_x\vec{\psi} = L(u, v, \lambda)\vec{\psi}, \quad \partial_t\vec{\psi} = A(u, v, \lambda)\vec{\psi},$$

for $\lambda = e^{i\gamma/2}$. As a result, the compatibility condition $\partial_x\partial_t\vec{\psi} = \partial_t\partial_x\vec{\psi}$ for every $x \in \mathbb{R}$ and $t \in \mathbb{R}$ yields the MTM system (2.1) for the functions (u, v).

We shall now prove inequality (3.84). The proof of inequality (3.85) is analogous. First, we write (3.82) in the form of

$$R := \frac{2i\sin\gamma\overline{\phi}_1\phi_2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2} = u + p\frac{e^{-i\gamma/2}|\phi_1|^2 + e^{i\gamma/2}|\phi_2|^2}{e^{i\gamma/2}|\phi_1|^2 + e^{-i\gamma/2}|\phi_2|^2}. \tag{3.87}$$

Explicit substitutions of (3.81) into (3.87) yield

$$R := \frac{2i\sin\gamma\left(\overline{m}_1 m_2 e^{-i\theta}\overline{\varphi}_1\chi_2 + R_1\right)}{e^{i\gamma/2+a-x\sin\gamma}|\varphi_1|^2 + e^{-i\gamma/2-a+x\sin\gamma}|\chi_2|^2 + R_2},$$

where

$$R_1 := \overline{m}_1^2 e^{a-x\sin\gamma}\overline{\varphi}_1\varphi_2 + \overline{m}_1 m_2 e^{i\theta}\varphi_2\overline{\chi}_1 + m_2^2 e^{-a+x\sin\gamma}\overline{\chi}_1\chi_2$$

and

$$\begin{aligned} R_2 := {} & e^{i\gamma/2-a+x\sin\gamma}|\chi_1|^2 + e^{-i\gamma/2+a-x\sin\gamma}|\varphi_2|^2 + 2e^{i\gamma/2}\mathrm{Re}[m_1 m_2 e^{i\theta}\varphi_1\overline{\chi}_1] \\ & + 2e^{-i\gamma/2}\mathrm{Re}[\overline{m}_1\overline{m}_2 e^{i\theta}\varphi_2\overline{\chi}_2]. \end{aligned}$$

By bounds (3.78) and (3.79) in Lemma 14, we have $|\varphi_1|, |\chi_2| \sim 1$ and $|\varphi_2|, |\chi_1| \sim 0$, so that for $a - x\sin\gamma \leq 0$,

$$R = \frac{2i\sin\gamma\overline{m}_1 m_2 e^{-i\theta+a-x\sin\gamma}\overline{\varphi}_1\chi_2}{e^{i\gamma/2+2(a-x\sin\gamma)}|\varphi_1|^2 + e^{-i\gamma/2}|\chi_2|^2} + \mathcal{O}(|\varphi_2| + |\chi_1|) \tag{3.88}$$

and for $a - x\sin\gamma \geq 0$,

$$R = \frac{2i\sin\gamma\overline{m}_1 m_2 e^{-i\theta-a+x\sin\gamma}\overline{\varphi}_1\chi_2}{e^{i\gamma/2}|\varphi_1|^2 + e^{-i\gamma/2-2(a-x\sin\gamma)}|\chi_2|^2} + \mathcal{O}(|\varphi_2| + |\chi_1|). \tag{3.89}$$

Combining (3.88) and (3.89), we get

$$\begin{aligned} & \left| R - \frac{2i\sin\gamma e^{-i\theta-it\cos\gamma}}{e^{i\gamma/2+a-x\sin\gamma} + e^{-i\gamma/2} - a + x\sin\gamma} \right| \\ \lesssim {} & e^{-|a-x\sin\gamma|}(|\varphi_1 - e^{it\frac{\cos\gamma}{2}}| + |\chi_2 - e^{-it\frac{\cos\gamma}{2}}| + |m_1 - 1| + |m_2 - 1|) + |\varphi_2| + |\chi_1| \end{aligned}$$

Since $m_1 = e^{\frac{i}{4}\int_{-\infty}^{x}(|p|^2-|q|^2)ds}$ and $m_2 = e^{\frac{i}{4}\int_{x}^{\infty}(|p|^2-|q|^2)ds}$, we obtain the bounds

$$\|m_{1,2}(\cdot,t) - 1\|_{L^\infty} \lesssim \|p\|_{L^2}^2 + \|q\|_{L^2}^2,$$

provided that $\|p\|_{L^2}$ and $\|q\|_{L^2}$ are sufficiently small. Then, by Lemma 14 and the L^2 conservation law (3.2), the previous estimate yields

$$\left\| R(x,t) - ie^{-i\theta-it\cos\gamma}\sin\gamma\,\mathrm{sech}\left(x\sin\gamma - i\frac{\gamma}{2} - a\right)\right\|_{L_x^2} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2}. \tag{3.90}$$

Using the definition (3.87), the bound (3.90), and the triangle inequality, we obtain inequality (3.84).

Lastly, if $(p,q) \in C(\mathbb{R}; H^2(\mathbb{R}))$, we can differentiate equations (3.88) and (3.89) in x twice to show from (3.82) and (3.83) that $(u,v) \in C(\mathbb{R}; H^2(\mathbb{R}))$. □

3.5 Proof of the main result

Proof of Theorem 3. Thanks to the Lorentz transformation given by Proposition 3, we may choose $\lambda_0 = e^{i\gamma_0/2}$, $\gamma_0 \in (0,\pi)$ in Theorem 3. For a given initial data (u_0, v_0) satisfying the inequality (3.4) for sufficiently small ϵ, we map a L^2-neighborhood of one-soliton solution to that of the zero solution. To do so, we use Lemma 8 and obtain an eigenvector $\vec{\psi}$ of the spectral problem (4.10) for an eigenvalue $\lambda \in \mathbb{C}$ satisfying

$$|\lambda - e^{i\gamma_0/2}| \lesssim \|u_0 - u_{\gamma_0}\|_{L^2} + \|v_0 - v_{\gamma_0}\|_{L^2} \le \epsilon. \tag{3.91}$$

We should note that the same Lorentz transformation cannot be used twice to consider the cases of $\lambda_0 = e^{i\gamma_0/2}$ and $\lambda = e^{i\gamma/2}$ simultaneously; the assumption $\lambda_0 = e^{i\gamma_0/2}$ implies that λ is not generally on the unit circle, and vice versa. Hence, if $\lambda_0 = e^{i\gamma_0/2}$ is set, all formulas in Section 3 below Remark 3 must in fact be generalized for a general λ. However, this generalization is straightforward thanks again to the existence of the Lorentz transformation given by Proposition 3. In what follows, we then use the general MTM solitons (u_λ, v_λ) given by (3.3).

By Lemma 12, the auto–Bäcklund transformation (3.24) and (3.25) with $\vec{\psi}$ in Lemma 8 yields an initial data $(p_0, q_0) \in L^2(\mathbb{R})$ of the MTM system (2.1) satisfying the estimate

$$\begin{aligned}
\|p_0\|_{L^2} + \|q_0\|_{L^2} &\lesssim \|u_0 - u_\lambda(\cdot,0)\|_{L^2} + \|v_0 - v_\lambda(\cdot,0)\|_{L^2} \\
&\lesssim \|u_0 - u_{\gamma_0}\|_{L^2} + \|v_0 - v_{\gamma_0}\|_{L^2} + \|u_\lambda(\cdot,0) - u_{\gamma_0}\|_{L^2} \\
&\qquad + \|v_\lambda(\cdot,0) - v_{\gamma_0}\|_{L^2} \\
&\lesssim \|u_0 - u_{\gamma_0}\|_{L^2} + \|v_0 - v_{\gamma_0}\|_{L^2} =: \epsilon,
\end{aligned} \tag{3.92}$$

where we have used the triangle inequality and the bound (3.91).

Since the time evolution in Section 4 is well-defined if $(p_0, q_0) \in H^2(\mathbb{R})$, let us first assume that the initial data $(u_0, v_0) \in L^2(\mathbb{R})$ satisfying the inequality (3.4) also satisfy $(u_0, v_0) \in H^2(\mathbb{R})$. Then, $(p_0, q_0) \in H^2(\mathbb{R})$ by Lemma 12. Let $(p, q) \in C(\mathbb{R}; H^2(\mathbb{R}))$ be the unique solution of the MTM system (2.1) such that $(p, q)|_{t=0} = (p_0, q_0)$. Next we will map a L^2-neighborhood of the zero solution to that of one-soliton solution for all $t \in \mathbb{R}$.

By Lemma 14, we construct a solution of the Lax equations

$$\partial_x \vec{\phi} = L(p, q, \lambda)\vec{\phi} \quad \text{and} \quad \partial_t \vec{\phi} = A(p, q, \lambda)\vec{\phi} \tag{3.93}$$

for the same eigenvalue λ as in (3.91). Let

$$k_1(\lambda) := \frac{i}{4}\left(\lambda^2 - \frac{1}{\lambda^2}\right), \quad k_2(\lambda) := \frac{1}{4}\left(\lambda^2 + \frac{1}{\lambda^2}\right).$$

The solution of the Lax system (3.93) is constructed in the form

$$\vec{\phi}(x,t) = c_1(t) M_1(x,t) e^{x k_1(\lambda)} \vec{\varphi}(x,t) + c_2(t) M_2(x,t) e^{-x k_1(\lambda)} \vec{\chi}(x,t), \tag{3.94}$$

where unitary matrices M_1 and M_2 are given in (3.64) with m_1 and m_2 given by (3.74), whereas the vectors $\vec{\varphi}$ and $\vec{\chi}$ satisfy the estimates

$$\|\varphi_1(\cdot,t) - e^{itk_2(\lambda)}\|_{L^\infty} + \|\varphi_2(\cdot,t)\|_{L^2\cap L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \leq \epsilon \tag{3.95}$$

and

$$\|\chi_1(\cdot,t)\|_{L^2\cap L^\infty} + \|\chi_2(\cdot,t) - e^{-itk_2(\lambda)}\|_{L^\infty} \lesssim \|p_0\|_{L^2} + \|q_0\|_{L^2} \leq \epsilon. \tag{3.96}$$

The coefficients c_1 and c_2 of the linear superposition (3.94) can be parameterized by parameters a and θ as follows:

$$c_1 = e^{(a+i\theta)/2}, \quad c_2 = e^{-(a+i\theta)/2},$$

where parameters a and θ may depend on the time variable t but not on the space variable x. These parameters determine the spatial and gauge translations of the MTM solitons according to the transformation (**??**).

By Lemma 15, the auto–Bäcklund transformation generates a new solution (u,v) of the MTM system (2.1) satisfying the bound for every $t \in \mathbb{R}$,

$$\inf_{a,\theta\in\mathbb{R}} (\|u(\cdot+a,t)-e^{-i\theta}u_\lambda(\cdot,t)\|_{L^2}+\|v(\cdot+a,t)-e^{-i\theta}v_\lambda(\cdot,t)\|_{L^2}) \lesssim \|p_0\|_{L^2}+\|q_0\|_{L^2} \leq \epsilon. \tag{3.97}$$

Theorem 3 is proved if $(u_0,v_0) \in H^2(\mathbb{R})$. To obtain the same result for $(u_0,v_0) \in L^2(\mathbb{R})$ but $(u_0,v_0) \notin H^2(\mathbb{R})$, we construct an approximating sequence $(u_{0,n},v_{0,n}) \in H^2(\mathbb{R})$ $(n \in \mathbb{N})$ that converges as $n \to \infty$ to $(u_0,v_0) \in L^2(\mathbb{R})$ in the L^2-norm. For a sufficiently small $\epsilon > 0$, we let

$$\|u_{0,n} - u_{\gamma_0}\|_{L^2} + \|v_{0,n} - v_{\gamma_0}\|_{L^2} \leq \epsilon, \quad \text{for every } n \in \mathbb{N}.$$

Under this condition, for each $(u_{0,n},v_{0,n}) \in H^2(\mathbb{R})$, we obtain inequalities (3.91), (3.92), and (3.97) independently of n. Therefore, there is a subsequence of solutions $(u_n,v_n) \in C(\mathbb{R};H^2(\mathbb{R}))$ $(n \in \mathbb{N})$ of the MTM system (2.1) such that it converges as $n \to \infty$ to a solution $(u,v) \in C(\mathbb{R};L^2(\mathbb{R}))$ of the MTM system (2.1) satisfying inequalities (3.5) and (3.6). The proof of Theorem 3 is now complete. □

Chapter 4

Global well-posedness of the derivative NLS by the inverse scattering transform

4.1 Main result

We consider the Cauchy problem for the derivative nonlinear Schrödinger (DNLS) equation

$$\begin{cases} iu_t + u_{xx} + i(|u|^2 u)_x = 0, & t > 0, \\ u|_{t=0} = u_0, \end{cases} \tag{4.1}$$

where the subscripts denote partial derivatives and u_0 is defined in a suitable function space, e.g., in Sobolev space $H^m(\mathbb{R})$ of distributions with square integrable derivatives up to the order m.

Local existence of solutions for $u_0 \in H^s(\mathbb{R})$ with $s > \frac{3}{2}$ was established by Tsutsumi & Fukuda [106] by using a parabolic regularization. Later, the same authors [107] used the first five conserved quantities of the DNLS equation and established the global existence of solutions for $u_0 \in H^2(\mathbb{R})$ provided the initial data is small in the $H^1(\mathbb{R})$ norm.

Using a gauge transformation of the DNLS equation to a system of two semi-linear NLS equations, for which a contraction argument can be used in the space $L^2(\mathbb{R})$ with the help of the Strichartz estimates, Hayashi [43] proved local and global existence of solutions to the DNLS equation for $u_0 \in H^1(\mathbb{R})$ provided that the initial data is small in the $L^2(\mathbb{R})$ norm. More specifically, the initial data u_0 is required to satisfy the precise inequality:

$$\|u_0\|_{L^2} < \sqrt{2\pi}. \tag{4.2}$$

The space $H^1(\mathbb{R})$ is referred to as the energy space for the DNLS equation because its first three conserved quantities having the meaning of the mass, momentum, and energy are well-defined in the space $H^1(\mathbb{R})$:

$$I_0 = \int_{\mathbb{R}} |u|^2 dx, \tag{4.3}$$

$$I_1 = i \int_{\mathbb{R}} (\bar{u}u_x - u\bar{u}_x) dx - \int_{\mathbb{R}} |u|^4 dx, \tag{4.4}$$

$$I_2 = \int_{\mathbb{R}} |u_x|^2 dx + \frac{3i}{4} \int_{\mathbb{R}} |u|^2 (u\bar{u}_x - u_x\bar{u}) dx + \frac{1}{2} \int_{\mathbb{R}} |u|^6 dx. \tag{4.5}$$

Using the gauge transformation $u = ve^{-\frac{3i}{4}\int_{-\infty}^{x} |v(y)|^2 dy}$ and the Gagliardo–Nirenberg inequality [109]

$$\|u\|_{L^6}^6 \leq \frac{4}{\pi^2} \|u\|_{L^2}^4 \|u_x\|_{L^2}^2, \tag{4.6}$$

one can obtain

$$I_2 = \|v_x\|_{L^2}^2 - \frac{1}{16}\|v\|_{L^6}^6 \geq \left(1 - \frac{1}{4\pi^2}\|v\|_{L^2}^4\right) \|v_x\|_{L^2}^2.$$

Under the small-norm assumption (4.2), the $H^1(\mathbb{R})$ norm of the function v (and hence, the $H^1(\mathbb{R})$ norm of the solution u to the DNLS equation) is controlled by the conserved quantities I_0 and I_2, once the local existence of solutions in $H^1(\mathbb{R})$ is established.

Developing the approach based on the gauge transformation and a priori energy estimates, Hayashi & Ozawa [44, 45, 81] considered global solutions to the DNLS equation in weighted Sobolev spaces under the same small-norm assumption (4.2), e.g., for $u_0 \in H^m(\mathbb{R}) \cap L^{2,m}(\mathbb{R})$, where $m \in \mathbb{N}$. Here and in what follows, $L^{2,m}(\mathbb{R})$ denotes the weighted $L^2(\mathbb{R})$ space with the norm

$$\|u\|_{L^{2,m}} := \left(\int_{\mathbb{R}} (1+x^2)^m |u|^2 dx\right)^{1/2} = \left(\int_{\mathbb{R}} \langle x\rangle^{2m} |u|^2 dx\right)^{1/2},$$

where $\langle x\rangle := (1+x^2)^{1/2}$.

More recently, local well-posedness of solutions to the DNLS equation was established in spaces of lower regularity, e.g., for $u_0 \in H^s(\mathbb{R})$ with $s \geq \frac{1}{2}$ by Takaoka [101] who used the Fourier transform restriction method. This result was shown to be sharp in the sense that the flow map fails to be uniformly continuous for $s < \frac{1}{2}$ [10]. Global existence under the constraint (4.2) was established in $H^s(\mathbb{R})$ with subsequent generations of the Fourier transform restriction method and the so-called I-method, e.g., for $s > \frac{32}{33}$ by Takaoka [102], for $s > \frac{2}{3}$ and $s > \frac{1}{2}$ by Colliander *et al.* [19] and [20] respectively, and finally for $s = \frac{1}{2}$ by Miao, Wu and Xu [75].

The key question, which goes back to the paper of Hayashi & Ozawa [44], is to find out *if the bound (4.2) is optimal for existence of global solutions to the DNLS equation*. By analogy with the quintic nonlinear Schrödinger (NLS) and Korteweg–de Vries (KdV) equations, one can ask if solutions with the $L^2(\mathbb{R})$ norm exceeding the threshold value in the inequality (4.2) can blow up in a finite time.

The threshold value $\sqrt{2\pi}$ for the $L^2(\mathbb{R})$ norm corresponds to the constant value of the $L^2(\mathbb{R})$ norm of the stationary solitary wave solutions to the DNLS equation. These solutions can be written in the explicit form:

$$u(x,t) = \phi_\omega(x) e^{i\omega^2 t - \frac{3i}{4}\int_{-\infty}^{x} |\phi_\omega(y)|^2 dy}, \quad \phi_\omega(x) = \sqrt{4\omega \operatorname{sech}(2\omega x)}, \quad \omega \in \mathbb{R}^+, \quad (4.7)$$

from which we have $\|\phi_\omega\|_{L^2} = \sqrt{2\pi}$ for every $\omega \in \mathbb{R}^+$. Although the solitary wave solutions are unstable in the quintic NLS and KdV equations, it was proved by Colin & Ohta [21] that the solitary wave of the DNLS equation is orbitally stable with respect to perturbations in $H^1(\mathbb{R})$. This result indicates that there exist global solutions to the DNLS equation (4.1) in $H^1(\mathbb{R})$ with the $L^2(\mathbb{R})$ norm exceeding the threshold value in (4.2).

Moreover, Colin & Ohta [21] proved that the moving solitary wave solutions of the DNLS equation are also orbitally stable in $H^1(\mathbb{R})$. Since the $L^2(\mathbb{R})$ norm of the moving solitary wave solutions is bounded from above by $2\sqrt{\pi}$, the orbital stability result indicates that there exist global solutions to the DNLS equation (4.1) if the initial data u_0 satisfies the inequality

$$\|u_0\|_{L^2} < 2\sqrt{\pi}. \quad (4.8)$$

These orbital stability results suggest that the inequality (4.2) is not sharp for the global existence in the DNLS equation (4.1). Furthermore, recent numerical explorations of the DNLS equation (4.1) indicate no blow-up phenomenon for initial data with any large $L^2(\mathbb{R})$ norm [70, 71]. The same conclusion is indicated by the asymptotic analysis in the recent work [16].

Towards the same direction, Wu [110] proved that the solution to the DNLS equations with $u_0 \in H^1(\mathbb{R})$ does not blow up in a finite time if the $L^2(\mathbb{R})$ norm of the initial data u_0 slightly exceed the threshold value in (4.2). The technique used in [110] is a combination of a variational argument together with the mass, momentum and energy conservation in (4.3)–(4.5). On the other hand, the solution to the DNLS equation restricted on the half line $\mathbb{R}^+$ blows up in a finite time if the initial data $u_0 \in H^2(\mathbb{R}^+) \cap L^{2,1}(\mathbb{R}^+)$ yields the negative energy $I_2 < 0$ given by (4.5) [110]. Proceeding further with sharper Gagliardo–Nirenberg-type inequalities, Wu [111] proved very recently that the global solutions to the DNLS equation exists in $H^1(\mathbb{R})$ if the initial data $u_0 \in H^1(\mathbb{R})$ satisfies the inequality (4.8), which is larger than the inequality (4.2).

Our approach to address the same question concerning global existence in the Cauchy problem for the DNLS equation (4.1) without the small $L^2(\mathbb{R})$-norm assumption relies on a different technique involving the inverse scattering transform theory [6, 7]. As was shown by Kaup & Newell [57], the DNLS equation appears to be a compatibility condition for suitable solutions to the linear system given by

$$\partial_x \psi = \left[-i\lambda^2 \sigma_3 + \lambda Q(u) \right] \psi \quad (4.9)$$

and

$$\partial_t \psi = \left[-2i\lambda^4 \sigma_3 + 2\lambda^3 Q(u) + i\lambda^2 |u|^2 \sigma_3 - \lambda |u|^2 Q(u) + i\lambda \sigma_3 Q(u_x) \right] \psi, \qquad (4.10)$$

where $\psi \in \mathbb{C}^2$ is assumed to be a C^2 function of x and t, $\lambda \in \mathbb{C}$ is the (x,t)-independent spectral parameter, and $Q(u)$ is the (x,t)-dependent matrix potential given by

$$Q(u) = \begin{bmatrix} 0 & u \\ -\bar{u} & 0 \end{bmatrix}. \qquad (4.11)$$

The Pauli matrices that include σ_3 are given by

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \qquad (4.12)$$

A long but standard computation shows that the compatibility condition $\partial_t \partial_x \psi = \partial_x \partial_t \psi$ for eigenfunctions $\psi \in C^2(\mathbb{R} \times \mathbb{R})$ is equivalent to the DNLS equation $iu_t + u_{xx} + i(|u|^2 u)_x = 0$ for classical solutions u. The linear equation (4.9) is usually referred to as the Kaup–Newell spectral problem.

In a similar context of the cubic NLS equation, it is well known that the inverse scattering transform technique applied to the linear system (associated with the so-called Zakharov–Shabat spectral problem) provides a rigorous framework to solve the Cauchy problem in weighted L^2 spaces, e.g., for $u_0 \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ [29, 31, 121] or for $u_0 \in H^1(\mathbb{R}) \cap L^{2,s}(\mathbb{R})$ with $s > \frac{1}{2}$ [?]. In comparison with the spectral problem (4.9), the Zakharov–Shabat spectral problem has no multiplication of matrix potential $Q(u)$ by λ. As a result, Neumann series solutions for the Jost functions of the Zakharov–Shabat spectral problem converge if u belongs to the space $L^1(\mathbb{R})$, see, e.g., Chapter 2 in [1]. As was shown originally by Deift & Zhou [31, 121], the inverse scattering problem based on the Riemann–Hilbert problem of complex analysis with a jump along the real line can be solved uniquely if u is defined in a subspace of $L^{2,1}(\mathbb{R})$, which is continuously embedded into the space $L^1(\mathbb{R})$. The time evolution of the scattering data is well defined if u is posed in space $H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ [29, 31].

For the Kaup–Newell spectral problem (4.9), the key feature is the presence of the spectral parameter λ that multiplies the matrix potential $Q(u)$. As a result, Neumann series solutions for the Jost functions do not converge uniformly if u is only defined in the space $L^1(\mathbb{R})$. Although the Lax system (4.9)–(4.10) appeared long ago and was used many times for formal methods, such as construction of soliton solutions [57], temporal asymptotics [59, 108], and long-time asymptotic expansions [112, 113], no rigorous results on the function spaces for the matrix $Q(u)$ have been obtained so far to ensure bijectivity of the direct and inverse scattering transforms for the Kaup–Newell spectral problem (4.9).

In this connection, we mention the works of Lee [66, 67] on the local solvability of a generalized Lax system with λ^n dependence for an integer $n \geq 2$ and generic small initial data u_0 in Schwarz class. In the follow-up paper [68], Lee also claimed existence of a global solution to the Cauchy problem (4.1) for large u_0 in Schwarz class, but the analysis of [68] relies on a "Basic Lemma", where the Jost functions are claimed to be defined for u_0 in $L^2(\mathbb{R})$. However, equation (4.9) shows that the condition $u_0 \in L^2(\mathbb{R})$ is insufficient for construction of the Jost functions uniformly in λ.

We address the bijectivity of the direct and inverse scattering transform for the Lax system (4.9)–(4.10) in this work. We show that the direct scattering transform for the Jost functions of the Lax system (4.9)–(4.18) can be developed under the requirement $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $\partial_x u_0 \in L^1(\mathbb{R})$. This requirement is satisfied if u_0 is defined in the weighted Sobolev space $H^{1,1}(\mathbb{R})$ defined by

$$H^{1,1}(\mathbb{R}) = \left\{ u \in L^{2,1}(\mathbb{R}), \quad \partial_x u \in L^{2,1}(\mathbb{R}) \right\}. \tag{4.13}$$

Note that it is quite common to use notation $H^{1,1}(\mathbb{R})$ to denote $H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ [31, 121], which is not what is used here in (4.13). Moreover, we show that asymptotic expansions of the Jost functions are well defined if $u_0 \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, which also provide a rigorous framework to study the inverse scattering transform based on the Riemann–Hilbert problem of complex analysis. Finally, the time evolution of the scattering data is well defined if $u_0 \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$.

We shall now define eigenvalues and resonances for the spectral problem (4.9) and present the global existence result for the DNLS equation (4.1).

Definition 3. *We say that $\lambda \in \mathbb{C}$ is an eigenvalue of the spectral problem (4.9) if the linear equation (4.9) with this λ admits a solution in $L^2(\mathbb{R})$.*

Definition 4. *We say that $\lambda \in \mathbb{R} \cup i\mathbb{R}$ is a resonance of the spectral problem (4.9) if the linear equation (4.9) with this λ admits a solution in $L^\infty(\mathbb{R})$ with the asymptotic behavior*

$$\psi(x) \sim \begin{cases} a_+ e^{-i\lambda^2 x} e_1, & x \to -\infty, \\ a_- e^{+i\lambda^2 x} e_2, & x \to +\infty, \end{cases}$$

where a_+ and a_- are nonzero constant coefficients, whereas $e_1 = [1,0]^t$ and $e_2 = [0,1]^t$.

Theorem 4. *For every $u_0 \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ such that the linear equation (4.9) admits no eigenvalues or resonances in the sense of Definitions 3 and 4, there exists a unique global solution $u(t,\cdot) \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ of the Cauchy problem* (4.1) *for every $t \in \mathbb{R}$. Furthermore, the map*

$$H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}) \ni u_0 \mapsto u \in C(\mathbb{R}; H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}))$$

is Lipschitz continuous.

Remark 5. *A sufficient condition that the spectral problem (4.9) admits no eigenvalues was found in [84]. This condition is satisfied under the small-norm assumption on the $H^{1,1}(\mathbb{R})$ norm of the initial data u_0. See Remark 9 below. Although we believe that there exist initial data u_0 with large $H^{1,1}(\mathbb{R})$ norm that yield no eigenvalues in the spectral problem (4.9), we have no constructive examples of such initial data. Nevertheless, a finite number of eigenvalues $\lambda \in \mathbb{C}$ in the spectral* *problem (4.9) can be included by using algebraic methods such as the Backlünd, Darboux, or dressing transformations [22, 24].*

Remark 6. *The condition that the spectral problem (4.9) admits no resonance is used to identify the so-called generic initial data u_0. The non-generic initial data u_0 violating this condition are at the threshold case in the sense that a small perturbation to u_0 may change the number of eigenvalues λ in the linear equation (4.9).*

Remark 7. *Compared to the results of Hayashi & Ozawa [43, 44, 45, 81], where global well-posedness of the Cauchy problem for the DNLS equation (4.1) was established in $H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ under the small $L^2(\mathbb{R})$ norm assumption (4.2), the inverse scattering transform theory is developed without the smallness assumption on the initial data u_0.*

Remark 8. *An alternative proof of Theorem 4 is developed in [69] by using a different version of the inverse scattering transform for the Lax system (4.9)–(4.10). The results of [69] are formulated in space $H^2(\mathbb{R}) \cap L^{2,2}(\mathbb{R})$, which is embedded into space $H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$.*

The paper is organized as follows. Section 2 reports the solvability results on the direct scattering transform for the spectral problem (4.9). Section 3 gives equivalent formulations of the Riemann–Hilbert problem associated with the spectral problem (4.9). Section 4 is devoted to the solvability results on the inverse scattering transform for the spectral problem (4.9). Section 5 incorporates the time evolution of the linear equation (4.10) and contains the proof of Theorem 4.

4.2 Direct scattering transform

The direct scattering transform is developed for the Kaup–Newell spectral problem (4.9), which we rewrite here for convenience:

$$\partial_x \psi = \left[-i\lambda^2 \sigma_3 + \lambda Q(u) \right] \psi, \tag{4.14}$$

where $\psi \in \mathbb{C}^2$, $\lambda \in \mathbb{C}$, and the matrices $Q(u)$ and σ_3 are given by (4.11) and (4.12).

The formal construction of the Jost functions is based on the construction of the fundamental solution matrices $\Psi_\pm(x;\lambda)$ of the linear equation (4.14), which satisfy the same asymptotic behavior at infinity as the linear equation (4.14) with $Q(u) \equiv 0$:

$$\Psi_\pm(x;\lambda) \to e^{-i\lambda^2 x \sigma_3} \quad \text{as } x \to \pm\infty, \tag{4.15}$$

where parameter λ is fixed in an unbounded subset of $\mathbb{C}$. However, the standard fixed point argument for Volterra's integral equations associated with the linear equation (4.14) is not uniform in λ as $|\lambda| \to \infty$ if $Q(u) \in L^1(\mathbb{R})$. Integrating by parts, it was suggested in [84] that uniform estimates on the Jost functions of the linear equation (4.14) can be obtained under the condition

$$\|u\|_{L^1}(\|u\|_{L^\infty} + \|\partial_x u\|_{L^1}) < \infty.$$

Here we explore this idea further and introduce a transformation of the linear equation (4.14) to a spectral problem of the Zakharov–Shabat type. This will allow us to adopt the direct and inverse scattering transforms, which were previously used for the cubic NLS equation [31, 121] (see also [24, 29] for review). Note that the pioneer idea of a transformation of the linear equation (4.14) to a spectral problem of the Zakharov–Shabat type can be found already in the formal work of Kaup & Newell [57].

Let us define the transformation matrices for any $u \in L^\infty(\mathbb{R})$ and $\lambda \in \mathbb{C}$,

$$T_1(x;\lambda) = \begin{bmatrix} 1 & 0 \\ -\overline{u}(x) & 2i\lambda \end{bmatrix} \quad \text{and} \quad T_2(x;\lambda) = \begin{bmatrix} 2i\lambda & -u(x) \\ 0 & 1 \end{bmatrix}, \tag{4.16}$$

If the vector $\psi \in \mathbb{C}^2$ is transformed by $\psi_{1,2} = T_{1,2}\psi$, then straightforward computations show that $\psi_{1,2}$ satisfy the linear equations

$$\partial_x \psi_1 = \left[-i\lambda^2\sigma_3 + Q_1(u)\right]\psi_1, \quad Q_1(u) = \frac{1}{2i}\begin{bmatrix} |u|^2 & u \\ -2i\overline{u}_x - \overline{u}|u|^2 & -|u|^2 \end{bmatrix} \tag{4.17}$$

and

$$\partial_x \psi_2 = \left[-i\lambda^2\sigma_3 + Q_2(u)\right]\psi_2, \quad Q_2(u) = \frac{1}{2i}\begin{bmatrix} |u|^2 & -2iu_x + u|u|^2 \\ -\overline{u} & -|u|^2 \end{bmatrix}. \tag{4.18}$$

Note that $Q_{1,2}(u) \in L^1(\mathbb{R})$ if $u \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$. The linear equations (4.17) and (4.18) are of the Zakharov–Shabat-type, after we introduce the complex variable $z = \lambda^2$. In what follows, we study the Jost functions and the scattering coefficients for the linear equations (4.17) and (4.18).

4.2.1 Jost functions

Let us introduce the normalized Jost functions from solutions $\psi_{1,2}$ of the linear equations (4.17) and (4.18) with $z = \lambda^2$ in the form

$$m_\pm(x;z) = \psi_1(x;z)e^{ixz}, \quad n_\pm(x;z) = \psi_2(x;z)e^{-ixz}, \tag{4.19}$$

according to the asymptotic behavior

$$\left.\begin{array}{l} m_\pm(x;z) \to e_1, \\ n_\pm(x;z) \to e_2, \end{array}\right\} \quad \text{as} \quad x \to \pm\infty, \tag{4.20}$$

where $e_1 = [1,0]^t$ and $e_2 = [0,1]^t$. The normalized Jost functions satisfy the following Volterra's integral equations

$$m_\pm(x;z) = e_1 + \int_{\pm\infty}^x \begin{bmatrix} 1 & 0 \\ 0 & e^{2iz(x-y)} \end{bmatrix} Q_1(u(y)) m_\pm(y;z)dy \tag{4.21}$$

and

$$n_\pm(x;z) = e_2 + \int_{\pm\infty}^x \begin{bmatrix} e^{-2iz(x-y)} & 0 \\ 0 & 1 \end{bmatrix} Q_2(u(y)) n_\pm(y;z)dy. \tag{4.22}$$

The next two lemmas describe properties of the Jost functions, which are analogues to similar properties of the Jost functions in the Zakharov–Shabat spectral problem (see, e.g., Lemma 2.1 in [1]).

Lemma 16. *Let $u \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$. For every $z \in \mathbb{R}$, there exist unique solutions $m_\pm(\cdot; z) \in L^\infty(\mathbb{R})$ and $n_\pm(\cdot; z) \in L^\infty(\mathbb{R})$ satisfying the integral equations (4.21) and (4.22). Moreover, for every $x \in \mathbb{R}$, $m_-(x; \cdot)$ and $n_+(x; \cdot)$ are continued analytically in $\mathbb{C}^+$, whereas $m_+(x; \cdot)$ and $n_-(x; \cdot)$ are continued analytically in $\mathbb{C}^-$. Finally, there exists a positive z-independent constant C such that*

$$\|m_\mp(\cdot; z)\|_{L^\infty} + \|n_\pm(\cdot; z)\|_{L^\infty} \le C, \quad z \in \mathbb{C}^\pm. \tag{4.23}$$

Proof. It suffices to prove the statement for one Jost function, e.g., for m_-. The proof for other Jost functions is analogous. Let us define the integral operator K by

$$(Kf)(x; z) := \frac{1}{2i} \int_{-\infty}^{x} \begin{bmatrix} 1 & 0 \\ 0 & e^{2iz(x-y)} \end{bmatrix} \begin{bmatrix} |u(y)|^2 & u(y) \\ -2i\partial_y \overline{u}(y) - \overline{u(y)}|u(y)|^2 & -|u(y)|^2 \end{bmatrix} f(y)dy. \tag{4.24}$$

For every $z \in \mathbb{C}^+$ and every $x_0 \in \mathbb{R}$, we have

$$\begin{aligned} &\|(Kf)(\cdot; z)\|_{L^\infty(-\infty,x_0)} \\ &\le \frac{1}{2} \begin{bmatrix} \|u\|^2_{L^2(-\infty,x_0)} & \|u\|_{L^1(-\infty,x_0)} \\ 2\|\partial_x u\|_{L^1(-\infty,x_0)} + \|u\|^3_{L^3(-\infty,x_0)} & \|u\|^2_{L^2(-\infty,x_0)} \end{bmatrix} \|f(\cdot; z)\|_{L^\infty(-\infty,x_0)}. \end{aligned}$$

The operator K is a contraction from $L^\infty(-\infty, x_0)$ to $L^\infty(-\infty, x_0)$ if the two eigenvalues of the matrix

$$A = \frac{1}{2} \begin{bmatrix} \|u\|^2_{L^2(-\infty,x_0)} & \|u\|_{L^1(-\infty,x_0)} \\ 2\|\partial_x u\|_{L^1(-\infty,x_0)} + \|u\|^3_{L^3(-\infty,x_0)} & \|u\|^2_{L^2(-\infty,x_0)} \end{bmatrix}$$

are located inside the unit circle. The two eigenvalues are given by

$$\lambda_\pm = \frac{1}{2}\|u\|^2_{L^2(-\infty,x_0)} \pm \frac{1}{2}\sqrt{\|u\|_{L^1(-\infty,x_0)}(2\|\partial_x u\|_{L^1(-\infty,x_0)} + \|u\|^3_{L^3(-\infty,x_0)})},$$

so that $|\lambda_-| < |\lambda_+|$. Hence, the operator K is a contraction if $x_0 \in \mathbb{R}$ is chosen so that

$$\frac{1}{2}\|u\|^2_{L^2(-\infty,x_0)} + \frac{1}{2}\sqrt{\|u\|_{L^1(-\infty,x_0)}(2\|\partial_x u\|_{L^1(-\infty,x_0)} + \|u\|^3_{L^3(-\infty,x_0)})} < 1. \tag{4.25}$$

By the Banach Fixed Point Theorem, for this x_0 and every $z \in \mathbb{C}^+$, there exists a unique solution $m_-(\cdot; z) \in L^\infty(-\infty, x_0)$ of the integral equation (4.21). To extend this result to $L^\infty(\mathbb{R})$, we can split $\mathbb{R}$ into a finite number of subintervals such that the estimate (4.25) is satisfied in each subinterval. Unique solutions in each subinterval can be glued together to obtain the unique solution $m_-(\cdot; z) \in L^\infty(\mathbb{R})$ for every $z \in \mathbb{C}^+$.

Analyticity of $m_-(x;\cdot)$ in $\mathbb{C}^+$ for every $x \in \mathbb{R}$ follows from the absolute and uniform convergence of the Neumann series of analytic functions in z. Indeed, let us denote the L^1 matrix norm of the 2-by-2 matrix function Q as

$$\|Q\|_{L^1} := \sum_{i,j=1}^{2} \|Q_{i,j}\|_{L^1}.$$

If $u \in H^{1,1}(\mathbb{R})$, then $u \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$ so that $Q_1(u) \in L^1(\mathbb{R})$, where the matrix $Q_1(u)$ appears in the integral kernel K given by (4.24). For every $f(x;z) \in L^\infty(\mathbb{R} \times \mathbb{C}^+)$, we have

$$\|(K^n f)\|_{L^\infty} \leq \frac{1}{n!}\|Q_1(u)\|_{L^1}^n \|f\|_{L^\infty}. \tag{4.26}$$

As a result, the Neumann series for Volterra's integral equation (4.21) for m_- converges absolutely and uniformly for every $x \in \mathbb{R}$ and $z \in \mathbb{C}^+$ and contains analytic functions of z for $z \in \mathbb{C}^+$. Therefore, $m_-(x;\cdot)$ is analytic in $\mathbb{C}^+$ for every $x \in \mathbb{R}$ and it satisfies the bound (4.23). □

Remark 9. *If u is sufficiently small so that the estimate*

$$\frac{1}{2}\|u\|_{L^2}^2 + \frac{1}{2}\sqrt{\|u\|_{L^1}(2\|\partial_x u\|_{L^1} + \|u\|_{L^3}^3)} < \frac{1}{2} \tag{4.27}$$

holds on $\mathbb{R}$, then Banach Fixed Point Theorem yields the existence of the unique solution $m_-(\cdot;z) \in L^\infty(\mathbb{R})$ of the integral equation (4.21) such that $\|m_-(\cdot;z) - e_1\|_{L^\infty} < 1$. This is in turn equivalent to the conditions that the linear equation (4.14) has no $L^2(\mathbb{R})$ solutions for every $\lambda \in \mathbb{C}$ and the linear equation (4.14) has no resonances for every $\lambda \in \mathbb{R} \cup i\mathbb{R}$ in the sense of Definitions 3 and 4. Therefore, the small-norm constraint (4.27) is a sufficient condition that the assumptions of Theorem 4 are satisfied.

Lemma 17. *Under the conditions of Lemma 16, for every $x \in \mathbb{R}$, the Jost functions $m_\pm(x;z)$ and $n_\pm(x;z)$ satisfy the following limits as $|\mathrm{Im}(z)| \to \infty$ along a contour in the domains of their analyticity:*

$$\lim_{|z|\to\infty} m_\pm(x;z) = m_\pm^\infty(x) e_1, \quad m_\pm^\infty(x) := e^{\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2 dy} \tag{4.28}$$

and

$$\lim_{|z|\to\infty} n_\pm(x;z) = n_\pm^\infty(x) e_2, \quad n_\pm^\infty(x) := e^{-\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2 dy}. \tag{4.29}$$

If in addition, $u \in C^1(\mathbb{R})$, then for every $x \in \mathbb{R}$, the Jost functions $m_\pm(x;z)$ and $n_\pm(x;z)$ satisfy the following limits as $|\mathrm{Im}(z)| \to \infty$ along a contour in the domains of their analyticity:

$$\lim_{|z|\to\infty} z\left[m_\pm(x;z) - m_\pm^\infty(x) e_1\right] = q_\pm^{(1)}(x) e_1 + q_\pm^{(2)}(x) e_2 \tag{4.30}$$

and

$$\lim_{|z|\to\infty} z\left[n_\pm(x;z) - n_\pm^\infty(x)e_2\right] = s_\pm^{(1)}(x)e_1 + s_\pm^{(2)}(x)e_2, \tag{4.31}$$

where

$$\begin{aligned} q_\pm^{(1)}(x) &:= -\frac{1}{4}e^{\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2dy}\int_{\pm\infty}^x \left[u(y)\partial_y\bar{u}(y) + \frac{1}{2i}|u(y)|^4\right]dy,\\ q_\pm^{(2)}(x) &:= \frac{1}{2i}\partial_x\left(\bar{u}(x)e^{\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2dy}\right),\\ s_\pm^{(1)}(x) &:= -\frac{1}{2i}\partial_x\left(u(x)e^{-\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2dy}\right),\\ s_\pm^{(2)}(x) &:= \frac{1}{4}e^{-\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2dy}\int_{\pm\infty}^x \left[\bar{u}(y)\partial_y u(y) - \frac{1}{2i}|u(y)|^4\right]dy. \end{aligned}$$

Proof. Again, we prove the statement for the Jost function m_- only. The proof for other Jost functions is analogous. Let $m_- = [m_-^{(1)}, m_-^{(2)}]^t$ and rewrite the integral equation (4.21) in the component form:

$$m_-^{(1)}(x;z) = 1 + \frac{1}{2i}\int_{-\infty}^x u(y)\left[\bar{u}(y)m_-^{(1)}(y;z) + m_-^{(2)}(y;z)\right]dy, \tag{4.32}$$

and

$$\begin{aligned} m_-^{(2)}(x;z) = -\frac{1}{2i}\int_{-\infty}^x e^{2iz(x-y)}&\left[(2i\partial_y\bar{u}(y)\right.\\ &\left.+|u(y)|^2\bar{u}(y))m_-^{(1)}(y;z) + |u(y)|^2 m_-^{(2)}(y;z)\right]dy. \end{aligned} \tag{4.33}$$

Recall that for every $x \in \mathbb{R}$, $m_-(x;\cdot)$ is analytic in $\mathbb{C}^+$. By bounds (4.23) in Lemma 16, for every $u \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$, the integrand of the second equation (4.33) is bounded for every $z \in \mathbb{C}^+$ by an absolutely integrable z-independent function. Also, the integrand converges to zero for every $y \in (-\infty, x)$ as $|z| \to \infty$ in $\mathbb{C}^+$. By Lebesgue's Dominated Convergence Theorem, we obtain $\lim_{|z|\to\infty} m_-^{(2)}(x;z) = 0$, hence $m_-^\infty(x) := \lim_{|z|\to\infty} m_-^{(1)}(x;z)$ satisfies the inhomogeneous integral equation

$$m_-^\infty(x) = 1 + \frac{1}{2i}\int_{-\infty}^x |u(y)|^2 m_-^\infty(y)dy, \tag{4.34}$$

with the unique solution $m_-^\infty(x) = e^{\frac{1}{2i}\int_{-\infty}^x |u(y)|^2dy}$. This proves the limit (4.28) for m_-.

We now add the condition $u \in C^1(\mathbb{R})$ and use the technique behind Watson's Lemma related to the Laplace method of asymptotic analysis [76]. For every $x \in \mathbb{R}$ and every small $\delta > 0$, we split integration in the second equation (4.33) for $(-\infty, x-\delta)$ and $(x-\delta, x)$, rewriting it in the equivalent form:

$$\begin{aligned} m_-^{(2)}(x;z) &= \int_{-\infty}^{x-\delta} e^{2iz(x-y)}\phi(y;z)dy + \phi(x;z)\int_{x-\delta}^{x} e^{2iz(x-y)}dy \\ &+ \int_{x-\delta}^{x} e^{2iz(x-y)}\left[\phi(y;z)-\phi(x;z)\right]dy \equiv I + II + III, \end{aligned} \quad (4.35)$$

where

$$\phi(x;z) := -\frac{1}{2i}\left[(2i\partial_x\bar{u}(x) + |u(x)|^2\bar{u}(x))m_-^{(1)}(x;z) + |u(x)|^2 m_-^{(2)}(x;z)\right].$$

Since $\phi(\cdot;z) \in L^1(\mathbb{R})$, we have

$$|I| \leq e^{-2\delta \mathrm{Im}(z)}\|\phi(\cdot;z)\|_{L^1}.$$

Since $\phi(\cdot;z) \in C^0(\mathbb{R})$, we have

$$|III| \leq \frac{1}{2\mathrm{Im}(z)}\|\phi(x-\cdot;z)-\phi(x;z)\|_{L^\infty(x-\delta,x)}.$$

On the other hand, we have the exact value

$$II = -\frac{1}{2iz}\left[1-e^{2iz\delta}\right]\phi(x;z).$$

Let us choose $\delta := [\mathrm{Im}(z)]^{-1/2}$ such that $\delta \to 0$ as $\mathrm{Im}(z) \to \infty$. Then, by taking the limit along the contour in $\mathbb{C}^+$ such that $\mathrm{Im}(z) \to \infty$, we obtain

$$\lim_{|z|\to\infty} z m_-^{(2)}(x;z) = -\frac{1}{2i}\lim_{|z|\to\infty}\phi(x;z) = -\frac{1}{4}(2i\partial_x\bar{u}(x) + |u(x)|^2\bar{u}(x))m_-^\infty(x), \quad (4.36)$$

which yields the limit (4.30) for $m_-^{(2)}$. On the other hand, the first equation (4.32) can be rewritten as the differential equation

$$\partial_x m_-^{(1)}(x;z) = \frac{1}{2i}|u(x)|^2 m_-^{(1)}(x;z) + \frac{1}{2i}u(x)m_-^{(2)}(x;z).$$

Using $\bar{m}_-^\infty$ as the integrating factor,

$$\partial_x(\overline{m}_-^\infty(x)m_-^{(1)}(x;z)) = \frac{1}{2i}u(x)\overline{m}_-^\infty(x)m_-^{(2)}(x;z),$$

we obtain another integral equation for $m_-^{(1)}$:

$$m_-^{(1)}(x;z) = m_-^\infty(x) + \frac{1}{2i}m_-^\infty(x)\int_{-\infty}^{x} u(y)\overline{m}_-^\infty(y)m_-^{(2)}(y;z)dy, \quad (4.37)$$

Multiplying this equation by z and taking the limit $|z| \to \infty$, we obtain

$$\lim_{|z|\to\infty} z \left[m_-^{(1)}(x;z) - m_-^\infty(x) \right] = -\frac{1}{4} m_-^\infty(x) \int_{-\infty}^x \left[u(y)\partial_y \bar{u}(y) + \frac{1}{2i}|u(y)|^4 \right] dy, \tag{4.38}$$

which yields the limit (4.30) for $m_-^{(1)}$. $\square$

We shall now study properties of the Jost functions on the real axis of z. First, we note that following elementary result from the Fourier theory. For notational convenience, we use sometimes $\|f(z)\|_{L^2_z}$ instead of $\|f(\cdot)\|_{L^2}$.

Proposition 4. *If $w \in H^1(\mathbb{R})$, then*

$$\sup_{x\in\mathbb{R}} \left\| \int_{-\infty}^x e^{2iz(x-y)} w(y) dy \right\|_{L^2_z(\mathbb{R})} \leq \sqrt{\pi} \|w\|_{L^2}. \tag{4.39}$$

and

$$\sup_{x\in\mathbb{R}} \left\| 2iz \int_{-\infty}^x e^{2iz(x-y)} w(y) dy + w(x) \right\|_{L^2_z(\mathbb{R})} \leq \sqrt{\pi} \|\partial_x w\|_{L^2}. \tag{4.40}$$

Moreover, if $w \in L^{2,1}(\mathbb{R})$, then for every $x_0 \in \mathbb{R}^-$, we have

$$\sup_{x\in(-\infty,x_0)} \left\| \langle x \rangle \int_{-\infty}^x e^{2iz(x-y)} w(y) dy \right\|_{L^2_z(\mathbb{R})} \leq \sqrt{\pi} \|w\|_{L^{2,1}(-\infty,x_0)}, \tag{4.41}$$

where $\langle x \rangle := (1+x^2)^{1/2}$.

Proof. Here we give a quick proof based on Plancherel's theorem of Fourier analysis. For every $x \in \mathbb{R}$ and every $z \in \mathbb{R}$, we write

$$f(x;z) := \int_{-\infty}^x e^{2iz(x-y)} w(y) dy = \int_{-\infty}^0 e^{-2izy} w(y+x) dy,$$

so that

$$\begin{aligned} \|f(x;\cdot)\|_{L^2}^2 &= \int_{-\infty}^\infty \int_{-\infty}^0 \int_{-\infty}^0 \bar{w}(y_1+x) w(y_2+x) e^{2i(y_1-y_2)z} dy_1 dy_2 dz \\ &= \pi \int_{-\infty}^0 |w(y+x)|^2 dy = \pi \int_{-\infty}^x |w(y)|^2 dy. \end{aligned} \tag{4.42}$$

Bound (4.39) holds if $w \in L^2(\mathbb{R})$.

If $y \leq x \leq 0$, we have $1+y^2 \geq 1+x^2$, so that equation (4.42) implies

$$\|f(x;\cdot)\|_{L^2}^2 \leq \frac{\pi}{1+x^2} \int_{-\infty}^x (1+y^2)|w(y)|^2 dy \leq \frac{\pi}{1+x^2} \|w\|_{L^{2,1}(-\infty,x)}^2,$$

which yields the bound (4.41) for any fixed $x_0 \in \mathbb{R}^-$.

To get the bound (4.40), we note that if $w \in H^1(\mathbb{R})$, then $w \in L^\infty(\mathbb{R})$ and

$w(x) \to 0$ as $|x| \to \infty$. As a result, we have

$$2izf(x;z) + w(x) = \int_{-\infty}^{x} e^{2iz(x-y)} \partial_y w(y) dy.$$

The bound (4.40) follows from the computation similar to (4.42). □

Subtracting the asymptotic limits (4.28) and (4.29) in Lemma 17 from the Jost functions $m_\pm$ and $n_\pm$ in Lemma 16, we prove that for every fixed $x \in \mathbb{R}^\pm$, the remainder terms belongs to $H^1(\mathbb{R})$ with respect to the variable z if u belongs to the space $H^{1,1}(\mathbb{R})$ defined in (4.13). Moreover, subtracting also the $\mathcal{O}(z^{-1})$ terms as defined by (4.30) and (4.31) and multiplying the result by z, we prove that the remainder term belongs to $L^2(\mathbb{R})$ if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$. Note that if $u \in H^{1,1}(\mathbb{R})$, then the conditions of Lemma 16 are satisfied, so that $u \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$. Also if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, then the additional condition $u \in C^1(\mathbb{R})$ of Lemma 17 is also satisfied.

Lemma 18. *If $u \in H^{1,1}(\mathbb{R})$, then for every $x \in \mathbb{R}^\pm$, we have*

$$m_\pm(x;\cdot) - m_\pm^\infty(x)e_1 \in H^1(\mathbb{R}), \quad n_\pm(x;\cdot) - n_\pm^\infty(x)e_2 \in H^1(\mathbb{R}). \tag{4.43}$$

Moreover, if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, then for every $x \in \mathbb{R}$, we have

$$z\left[m_\pm(x;z) - m_\pm^\infty(x)e_1\right] - (q_\pm^{(1)}(x)e_1 + q_\pm^{(2)}(x)e_2) \in L_z^2(\mathbb{R}) \tag{4.44}$$

and

$$z\left[n_\pm(x;z) - n_\pm^\infty(x)e_2\right] - (s_\pm^{(1)}(x)e_1 + s_\pm^{(2)}(x)e_2) \in L_z^2(\mathbb{R}). \tag{4.45}$$

Proof. Again, we prove the statement for the Jost function m_-. The proof for other Jost functions is analogous. We write the integral equation (4.21) for m_- in the abstract form

$$m_- = e_1 + Km_-, \tag{4.46}$$

where the operator K is given by (4.24). Although equation (4.46) is convenient for verifying the boundary condition $m_-(x;z) \to e_1$ as $x \to -\infty$, we note that the asymptotic limit as $|z| \to \infty$ is different by the complex exponential factor. Indeed, for every $x \in \mathbb{R}$, the asymptotic limit (4.28) is written as

$$m_-(x;z) \to m_-^\infty(x)e_1 \quad \text{as} \quad |z| \to \infty, \quad \text{where} \quad m_-^\infty(x) := e^{\frac{1}{2i}\int_{-\infty}^{x}|u(y)|^2 dy}.$$

Therefore, we rewrite equation (4.46) in the equivalent form

$$(I - K)(m_- - m_-^\infty e_1) = he_2, \tag{4.47}$$

where we have used the integral equation (4.34) that yields $e_1 - (I-K)m_-^\infty e_1 = he_2$ with

$$h(x;z) = \int_{-\infty}^{x} e^{2iz(x-y)} w(y) dy, \quad w(x) := -\partial_x\left(\overline{u}(x) e^{\frac{1}{2i}\int_{-\infty}^{x}|u(y)|^2 dy}\right). \tag{4.48}$$

If $u \in H^{1,1}(\mathbb{R})$, then $w \in L^2(\mathbb{R})$. By the bounds (4.39) and (4.41) in Proposition 4, we have $h(x;z) \in L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$ and for every $x_0 \in \mathbb{R}^-$, the following bound is satisfied:

$$\begin{aligned} \sup_{x\in(-\infty,x_0)} \|\langle x\rangle\, h(x;z)\|_{L^2_z(\mathbb{R})} &\leq \sqrt{\pi}\left(\|\partial_x u\|_{L^{2,1}} + \frac{1}{2}\|u^3\|_{L^{2,1}}\right) \\ &\leq C(\|u\|_{H^{1,1}} + \|u\|^3_{H^{1,1}}), \end{aligned} \tag{4.49}$$

where C is a positive u-independent constant and the Sobolev inequality $\|u\|_{L^\infty} \leq \frac{1}{\sqrt{2}}\|u\|_{H^1}$ is used.

By using estimates similar to those in the derivation of the bound (4.26) in Lemma 16, we find that for every $f(x;z) \in L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$, we have

$$\|(K^n f)(x;z)\|_{L^\infty_x L^2_z} \leq \frac{1}{n!}\|Q_1(u)\|^n_{L^1}\|f(x;z)\|_{L^\infty_x L^2_z}. \tag{4.50}$$

Therefore, the operator $I-K$ is invertible on the space $L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$ and a bound on the inverse operator is given by

$$\|(I-K)^{-1}\|_{L^\infty_x L^2_z \to L^\infty_x L^2_z} \leq \sum_{n=0}^{\infty}\frac{1}{n!}\|Q_1(u)\|^n_{L^1} = e^{\|Q_1(u)\|_{L^1}}. \tag{4.51}$$

Moreover, the same estimate (4.51) can be obtained in the norm $L^\infty_x((-\infty,x_0); L^2_z(\mathbb{R}))$ for every $x_0 \in \mathbb{R}$. By using (4.47), (4.49), and (4.51), we obtain the following estimate for every $x_0 \in \mathbb{R}^-$:

$$\sup_{x\in(-\infty,x_0)} \left\|\langle x\rangle\left(m_-(x;z) - m^\infty_-(x)e_1\right)\right\|_{L^2_z(\mathbb{R})} \leq Ce^{\|Q_1(u)\|_{L^1}}\left(\|u\|_{H^{1,1}} + \|u\|^3_{H^{1,1}}\right). \tag{4.52}$$

Next, we want to show $\partial_z m_-(x;z) \in L^\infty_x((-\infty,x_0); L^2_z(\mathbb{R}))$ for every $x_0 \in \mathbb{R}^-$. We differentiate the integral equation (4.46) in z and introduce the vector $v = [v^{(1)}, v^{(2)}]^t$ with the components

$$v^{(1)}(x;z) := \partial_z m^{(1)}_-(x;z) \quad \text{and} \quad v^{(2)}(x;z) := \partial_z m^{(2)}_-(x;z) - 2ixm^{(2)}_-(x;z).$$

Thus, we obtain from (4.46):

$$(I-K)v = h_1 e_1 + h_2 e_2 + h_3 e_2, \tag{4.53}$$

where

$$\begin{aligned} h_1(x;z) &= \int_{-\infty}^{x} yu(y)m^{(2)}_-(y;z)dy, \\ h_2(x;z) &= \int_{-\infty}^{x} ye^{2iz(x-y)}(2i\overline{u}_y(y) + |u(y)|^2\overline{u}(y))(m^{(1)}_-(y;z) - m^\infty_-(y))dy, \\ h_3(x;z) &= \int_{-\infty}^{x} ye^{2iz(x-y)}(2i\overline{u}_y(y) + |u(y)|^2\overline{u}(y))m^\infty_-(y)dy. \end{aligned}$$

For every $x_0 \in \mathbb{R}^-$, each inhomogeneous term of the integral equation (4.53) can be estimated by using Hölder's inequality and the bound (4.39) of Proposition 4:

$$\begin{aligned}
\sup_{x\in(-\infty,x_0)} \|h_1(x;z)\|_{L^2_z(\mathbb{R})} &\leq \|u\|_{L^1} \sup_{x\in(-\infty,x_0)} \|\langle x\rangle\, m_-^{(2)}(x;z)\|_{L^2_z(\mathbb{R})}, \\
\sup_{x\in(-\infty,x_0)} \|h_2(x;z)\|_{L^2_z(\mathbb{R})} & \\
&\leq (2\|\partial_x u\|_{L^1} + \|u^3\|_{L^1}) \sup_{x\in(-\infty,x_0)} \left\|\langle x\rangle \left(m_-^{(1)}(x;z) - m_-^{\infty}(x)\right)\right\|_{L^2_z(\mathbb{R})}, \\
\sup_{x\in(-\infty,x_0)} \|h_3(x;z)\|_{L^2_z(\mathbb{R})} &\leq \sqrt{\pi}\left(2\|\partial_x u\|_{L^{2,1}} + \|u^3\|_{L^{2,1}}\right).
\end{aligned}$$

The upper bounds in the first two inequalities are finite due to estimate (4.52) and the embedding of $L^{2,1}(\mathbb{R})$ into $L^1(\mathbb{R})$. Using the bounds (4.51), (4.52), and the integral equation (4.53), we conclude that $v(x;z) \in L^\infty_x((-\infty,x_0);L^2_z(\mathbb{R}))$ for every $x_0 \in \mathbb{R}^-$. Since $xm_-^{(2)}(x;z)$ is bounded in $L^\infty_x((-\infty,x_0);L^2_z(\mathbb{R}))$ by the same estimate (4.52), we finally obtain $\partial_z m_-(x;z) \in L^\infty_x((-\infty,x_0);L^2_z(\mathbb{R}))$ for every $x_0 \in \mathbb{R}^-$. This completes the proof of (4.43) for m_-.

To prove (4.44) for m_-, we subtract the $\mathcal{O}(z^{-1})$ term as defined by (4.30) from the integral equation (4.47) and multiply the result by z. Thus, we obtain

$$(I-K)\left[z\left(m_- - m_-^\infty e_1\right) - (q_-^{(1)}e_1 + q_-^{(2)}e_2)\right] = zhe_2 - (I-K)(q_-^{(1)}e_1 + q_-^{(2)}e_2), \tag{4.54}$$

where the limiting values $q_-^{(1)}$ and $q_-^{(2)}$ are defined in Lemma 17. Using the integral equation (4.37), we obtain cancelation of the first component of the source term, so that

$$zhe_2 - (I-K)(q_-^{(1)}e_1 + q_-^{(2)}e_2) = \tilde{h}e_2$$

with

$$\begin{aligned}
\tilde{h}(x;z) &= z\int_{-\infty}^x e^{2iz(x-y)}w(y)dy + \frac{1}{2i}w(x) \\
&\quad -\frac{1}{2i}\int_{-\infty}^x e^{2iz(x-y)}\left[(2i\partial_y\bar{u}(y) + \bar{u}(y)|u(y)|^2)q_-^{(1)}(y) + |u(y)|^2 q_-^{(2)}(y)\right]dy,
\end{aligned}$$

where w is the same as in (4.48). By using bounds (4.39) and (4.40) in Proposition 4, we have $\tilde{h}(x;z) \in L^\infty_x(\mathbb{R};L^2_z(\mathbb{Z}))$ if $w \in H^1(\mathbb{R})$ in addition to $u \in H^{1,1}(\mathbb{R})$, that is, if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$. Inverting $(I-K)$ on $L^\infty_x(\mathbb{R};L^2_z(\mathbb{Z}))$, we finally obtain (4.44) for m_-. □

The following result is deduced from Lemma 18 to show that the mapping

$$H^{1,1}(\mathbb{R}) \ni u \to [m_\pm(x;z) - m_\pm^\infty(x)e_1, n_\pm(x;z) - n_\pm^\infty(x)] \in L^\infty_x(\mathbb{R}^\pm;H^1_z(\mathbb{R})) \tag{4.55}$$

is Lipschitz continuous. Moreover, by restricting the potential to $H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, subtracting $\mathcal{O}(z^{-1})$ terms from the Jost functions, and multiplying them by z, we also have Lipschitz continuity of remainders of the Jost functions in function space $L^\infty_x(\mathbb{R};L^2_z(\mathbb{R}))$.

Corollary 4. *Let $u, \tilde{u} \in H^{1,1}(\mathbb{R})$ satisfy $\|u\|_{H^{1,1}}, \|\tilde{u}\|_{H^{1,1}} \leq U$ for some $U > 0$. Denote the corresponding Jost functions by $[m_\pm, n_\pm]$ and $[\tilde{m}_\pm, \tilde{n}_\pm]$ respectively. Then, there is a positive U-dependent constant $C(U)$ such that for every $x \in \mathbb{R}^\pm$, we have*

$$\|m_\pm(x;\cdot) - m_\pm^\infty(x)e_1 - \tilde{m}_\pm(x;\cdot) + \tilde{m}_\pm^\infty(x)e_1\|_{H^1} \leq C(U)\|u - \tilde{u}\|_{H^{1,1}} \quad (4.56)$$

and

$$\|n_\pm(x;\cdot) - n_\pm^\infty(x)e_2 - \tilde{n}_\pm(x;\cdot) + \tilde{n}_\pm^\infty(x)e_2\|_{H^1} \leq C(U)\|u - \tilde{u}\|_{H^{1,1}}. \quad (4.57)$$

Moreover, if $u, \tilde{u} \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ satisfy $\|u\|_{H^2\cap H^{1,1}}, \|\tilde{u}\|_{H^2\cap H^{1,1}} \leq U$, then for every $x \in \mathbb{R}$, there is a positive U-dependent constant $C(U)$ such that

$$\|\hat{m}_\pm(x;\cdot) - \hat{\tilde{m}}_\pm(x;\cdot)\|_{L^2} + \|\hat{n}_\pm(x;\cdot) - \hat{\tilde{n}}_\pm(x;\cdot)\|_{L^2} \leq C(U)\|u - \tilde{u}\|_{H^2\cap H^{1,1}}. \quad (4.58)$$

where

$$\begin{aligned} \hat{m}_\pm(x;z) &:= z\left[m_\pm(x;z) - m_\pm^\infty(x)e_1\right] - (q_\pm^{(1)}(x)e_1 + q_\pm^{(2)}(x)e_2), \\ \hat{n}_\pm(x;z) &:= z\left[n_\pm(x;z) - n_\pm^\infty(x)e_2\right] - (s_\pm^{(1)}(x)e_1 + s_\pm^{(2)}(x)e_2). \end{aligned}$$

Proof. Again, we prove the statement for the Jost function m_-. The proof for other Jost functions is analogous. First, let us consider the limiting values of m_- and $\tilde{m}_-$ given by

$$m_-^\infty(x) := e^{\frac{1}{2i}\int_{-\infty}^x |u(y)|^2 dy}, \quad \tilde{m}_-^\infty(x) := e^{\frac{1}{2i}\int_{-\infty}^x |\tilde{u}(y)|^2 dy}$$

Then, for every $x \in \mathbb{R}$, we have

$$\begin{aligned} |m_-^\infty(x) - \tilde{m}_-^\infty(x)| &= \left|e^{\frac{1}{2i}\int_{-\infty}^x (|u(y)|^2 - |\tilde{u}(y)|^2)dy} - 1\right| \\ &\leq C_1(U)\int_{-\infty}^x (|u(y)|^2 - |\tilde{u}(y)|^2)dy \\ &\leq 2UC_1(U)\|u - \tilde{u}\|_{L^2}, \end{aligned} \quad (4.59)$$

where $C_1(U)$ is a U-dependent positive constant. Using the integral equation (4.47), we obtain

$$\begin{aligned} &(m_- - m_-^\infty e_1) - (\tilde{m}_- - \tilde{m}_-^\infty e_1) \\ &= (I - K)^{-1}he_2 - (I - \tilde{K})^{-1}\tilde{h}e_2 \\ &= (I - K)^{-1}(h - \tilde{h})e_2 + [(I - K)^{-1} - (I - \tilde{K})^{-1}]\tilde{h}e_2 \\ &= (I - K)^{-1}(h - \tilde{h})e_2 + (I - K)^{-1}(K - \tilde{K})(I - \tilde{K})^{-1}\tilde{h}e_2, \end{aligned} \quad (4.60)$$

where $\tilde{K}$ and $\tilde{h}$ denote the same as K and h but with u being replaced by $\tilde{u}$. To estimate the first term, we write

$$h(x;z) - \tilde{h}(x;z) = \int_{-\infty}^{x} e^{2iz(x-y)} \left[w(y) - \tilde{w}(y)\right] dy, \tag{4.61}$$

where

$$w - \tilde{w} \;\; = \;\; \left(\partial_x \bar{\tilde{u}} + \frac{1}{2i}|\tilde{u}|^2\bar{\tilde{u}}\right) \tilde{m}_-^{\infty} - \left(\partial_x \bar{u} + \frac{1}{2i}|u|^2\bar{u}\right) m_-^{\infty}.$$

By using (4.59), we obtain $\|w-\tilde{w}\|_{L^{2,1}} \leq C_2(U)\|u-\tilde{u}\|_{H^{1,1}}$, where $C_2(U)$ is another U-dependent positive constant. By using (4.61) and Proposition 4, we obtain for every $x_0 \in \mathbb{R}^-$:

$$\sup_{x\in(-\infty,x_0)} \left\|\langle x\rangle \left(h(x;z) - \tilde{h}(x;z)\right)\right\|_{L^2_z(\mathbb{R})} \;\; \leq \;\; \sqrt{\pi}C_2(U)\|u-\tilde{u}\|_{H^{1,1}}. \tag{4.62}$$

This gives the estimate for the first term in (4.60). To estimate the second term, we use (4.24) and observe that K is a Lipschitz continuous operator from $L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$ to $L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$ in the sense that for every $f \in L^\infty_x(\mathbb{R}; L^2_z(\mathbb{R}))$, we have

$$\|(K - \tilde{K})f\|_{L^\infty_x L^2_z} \leq C_3(U)\|u-\tilde{u}\|_{H^{1,1}}\|f\|_{L^\infty_x L^2_z}, \tag{4.63}$$

where $C_3(U)$ is another U-dependent positive constant that is independent of f. By using (4.49), (4.51), (4.60), (4.62), and (4.63), we obtain for every $x_0 \in \mathbb{R}^-$:

$$\sup_{x\in(-\infty,x_0)} \left\|\langle x\rangle \left(m_-(x;\cdot) - m_-^{\infty}(x)e_1 - \tilde{m}_-(x;\cdot) + \tilde{m}_-^{\infty}(x)e_1\right)\right\|_{L^2_z(\mathbb{R})} \leq C(U)\|u-\tilde{u}\|_{H^{1,1}}.$$

This yields the first part of the bound (4.56) for m_- and $\tilde{m}_-$. The other part of the bound (4.56) and the bound (4.58) for m_- and $\tilde{m}_-$ follow by repeating the same analysis to the integral equations (4.53) and (4.54). □

4.2.2 Scattering coefficients

Let us define the Jost functions of the original Kaup–Newell spectral problem (4.14). These Jost functions are related to the Jost functions of the Zakharov–Shabat spectral problems (4.17) and (4.18) by using the matrix transformations (4.16). To be precise, we define

$$\varphi_\pm(x;\lambda) = T_1^{-1}(x;\lambda)m_\pm(x;z), \quad \phi_\pm(x;\lambda) = T_2^{-1}(x;\lambda)n_\pm(x;z), \tag{4.64}$$

where the inverse matrices are given by

$$T_1^{-1}(x;\lambda) = \frac{1}{2i\lambda}\begin{bmatrix} 2i\lambda & 0 \\ \bar{u}(x) & 1 \end{bmatrix} \quad \text{and} \quad T_2^{-1}(x;\lambda) = \frac{1}{2i\lambda}\begin{bmatrix} 1 & u(x) \\ 0 & 2i\lambda \end{bmatrix}. \tag{4.65}$$

It follows from the integral equations (4.21)–(4.22) and the transformation (4.64) that the original Jost functions $\varphi_\pm$ and $\phi_\pm$ satisfy the following Volterra's integral equations

$$\varphi_{\pm}(x;\lambda) = e_1 + \lambda \int_{\pm\infty}^{x} \begin{bmatrix} 1 & 0 \\ 0 & e^{2i\lambda^2(x-y)} \end{bmatrix} Q(u(y))\varphi_{\pm}(y;\lambda)dy, \tag{4.66}$$

and

$$\phi_{\pm}(x;\lambda) = e_2 + \lambda \int_{\pm\infty}^{x} \begin{bmatrix} e^{-2i\lambda^2(x-y)} & 0 \\ 0 & 1 \end{bmatrix} Q(u(y))\phi_{\pm}(y;\lambda)dy. \tag{4.67}$$

The following corollary is obtained from Lemma 16 and the representations (4.64)–(4.65).

Corollary 5. *Let $u \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $\partial_x u \in L^1(\mathbb{R})$. For every $\lambda^2 \in \mathbb{R}\backslash\{0\}$, there exist unique functions $\varphi_{\pm}(\cdot;\lambda) \in L^\infty(\mathbb{R})$ and $\phi_{\pm}(\cdot;\lambda) \in L^\infty(\mathbb{R})$ such that*

$$\left.\begin{array}{l} \varphi_{\pm}(x;\lambda) \to e_1, \\ \phi_{\pm}(x;\lambda) \to e_2, \end{array}\right\} \quad \text{as} \quad x \to \pm\infty. \tag{4.68}$$

Moreover, $\varphi_{\pm}^{(1)}(x;\lambda)$ and $\phi_{\pm}^{(2)}(x;\lambda)$ are even in λ, whereas $\varphi_{\pm}^{(2)}(x;\lambda)$ and $\phi_{\pm}^{(1)}(x;\lambda)$ are odd in λ.

Proof. To the conditions of Lemma 16, we added the condition $u \in L^\infty(\mathbb{R})$, which ensures that $T_{1,2}^{-1}(x;\lambda)$ are bounded for every $x \in \mathbb{R}$ and for every $\lambda \in \mathbb{C}\backslash\{0\}$. Then, the existence and uniqueness of the functions $\varphi_{\pm}(\cdot;\lambda) \in L^\infty(\mathbb{R})$ and $\phi_{\pm}(\cdot;\lambda) \in L^\infty(\mathbb{R})$, as well as the limits (4.68) follow by the representation (4.64)–(4.65) and by the first assertion of Lemma 16. The parity argument for components of $\varphi_{\pm}(x;\lambda)$ and $\psi_{\pm}(x;\lambda)$ in λ follow from the representation (4.64)–(4.65) and the fact that $m_{\pm}(x;z)$ and $n_{\pm}(x;z)$ are even in λ since $z = \lambda^2$. □

Remark 10. *There is no singularity in the definition of Jost functions at the value $\lambda = 0$. The integral equations (4.66) and (4.67) with $\lambda = 0$ admit unique Jost functions $\varphi_{\pm}(x;0) = e_1$ and $\phi_{\pm}(x;0) = e_2$, which yield unique definitions for $m_{\pm}(x;0)$ and $n_{\pm}(x;0)$:*

$$m_{\pm}(x;0) = \begin{bmatrix} 1 \\ -\bar{u}(x) \end{bmatrix}, \quad n_{\pm}(x;0) = \begin{bmatrix} -u(x) \\ 1 \end{bmatrix},$$

which follow from the unique solutions to the integral equations (4.21) and (4.22) at $z = 0$.

Remark 11. *The only purpose in the definition of the original Jost functions (4.64) is to introduce the standard form of the scattering relations, similar to the one used in the literature [57]. After introducing the scattering data for $\lambda \in \mathbb{R} \cup i\mathbb{R}$, we analyze their behavior in the complex z-plane, instead of the complex λ-plane, where $z = \lambda^2$.*

Analytic properties of the Jost functions $\varphi_{\pm}(x;\cdot)$ and $\psi_{\pm}(x;\cdot)$ for every $x \in \mathbb{R}$ are summarized in the following result. The result is a corollary of Lemmas 16 and 18.

Corollary 6. *Under the same assumption as Corollary 5, for every $x \in \mathbb{R}$, the Jost functions $\varphi_-(x;\cdot)$ and $\phi_+(x;\cdot)$ are analytic in the first and third quadrant of the λ plane (where $\mathrm{Im}(\lambda^2) > 0$), whereas the Jost functions $\varphi_+(x;\cdot)$ and $\phi_-(x;\cdot)$ are analytic in the second and fourth quadrant of the λ plane (where $\mathrm{Im}(\lambda^2) < 0$). Moreover, if $u \in H^{1,1}(\mathbb{R})$, then for every $x \in \mathbb{R}^\pm$, we have*

$$\varphi_\pm^{(1)}(x;\lambda) - m_\pm^\infty(x), \;\; 2i\lambda\varphi_\pm^{(2)}(x;\lambda) - \bar{u}(x)m_\pm^\infty(x), \;\; \lambda^{-1}\varphi_\pm^{(2)}(x;\lambda) \in H_z^1(\mathbb{R}) \quad (4.69)$$

and

$$\lambda^{-1}\phi_\pm^{(1)}(x;\lambda), \;\; 2i\lambda\phi_\pm^{(1)}(x;\lambda) - u(x)n_\pm^\infty(x), \;\; \phi_\pm^{(2)}(x;\lambda) - n_\pm^\infty(x) \in H_z^1(\mathbb{R}), \quad (4.70)$$

where $m_\pm^\infty$ and $n_\pm^\infty$ are the same as in Lemma 17.

Proof. By chain rule, we obtain

$$\frac{\partial}{\partial\bar{\lambda}} = 2\bar{\lambda}\frac{\partial}{\partial\bar{z}}.$$

As a result, the analyticity result for the Jost functions $\varphi_\pm$ and $\phi_\pm$ follows from the corresponding result of Lemma 16. With the transformation (4.64)–(4.65) and the result of Lemma 18, we obtain (4.69) and (4.70) for $\varphi_\pm^{(1)}$, $\lambda\varphi_\pm^{(2)}$, $\lambda\phi_\pm^{(1)}$, and $\phi_\pm^{(2)}$.

It remains to consider $\lambda^{-1}\varphi_\pm^{(2)}$ and $\lambda^{-1}\phi_\pm^{(1)}$. Although the result also follows from Remark 10, we will give a direct proof. We write explicitly from the integral equation (4.66):

$$\begin{aligned}\lambda^{-1}\varphi_\pm^{(2)}(x;\lambda) = &-\int_{\pm\infty}^{x} e^{2iz(x-y)}\overline{u}(y)m_\pm^\infty(y)dy \\ &-\int_{\pm\infty}^{x} e^{2iz(x-y)}\overline{u}(y)\left(m_\pm^{(1)}(y;z) - m_\pm^\infty(y)\right)dy, \end{aligned} \quad (4.71)$$

where $m_\pm^\infty = e^{\frac{1}{2i}\int_{\pm\infty}^{x}|u(y)|^2dy}$ and $z = \lambda^2$ as the same as in Lemma 18. By using Proposition 4 in the same way as it was used in the proof of Lemma 18, we obtain $\lambda^{-1}\varphi_\pm^{(2)}(x;\lambda) \in H_z^1(\mathbb{R})$ for every $x \in \mathbb{R}^\pm$. The proof of $\lambda^{-1}\phi_\pm^{(1)}(x;\lambda) \in H_z^1(\mathbb{R})$ is similar. □

We note that $\psi(x) := \varphi_\pm(x;\lambda)e^{-i\lambda^2 x}$ and $\psi(x) := \phi_\pm(x;\lambda)e^{i\lambda^2 x}$ satisfies the Kaup–Newell spectral problem (4.14), see asymptotic limits (4.15) and (4.68). By the ODE theory for the second-order differential systems, only two solutions are linearly independent. Therefore, for every $x \in \mathbb{R}$ and every $\lambda^2 \in \mathbb{R}\backslash\{0\}$, we define the scattering data according to the following transfer matrix

$$\begin{bmatrix}\varphi_-(x;\lambda)\\ \phi_-(x;\lambda)\end{bmatrix} = \begin{bmatrix} a(\lambda) & b(\lambda)e^{2i\lambda^2 x}\\ c(\lambda)e^{-2i\lambda^2 x} & d(\lambda)\end{bmatrix}\begin{bmatrix}\varphi_+(x;\lambda)\\ \phi_+(x;\lambda)\end{bmatrix}. \quad (4.72)$$

By Remark 10, the transfer matrix is extended to $\lambda = 0$ with $a(0) = d(0) = 1$ and $b(0) = c(0) = 0$.

Since the coefficient matrix in the Kaup–Newell spectral problem (4.14) has zero trace, the Wronskian determinant, denoted by W, of two solutions to the differential system (4.14) for any $\lambda \in \mathbb{C}$ is independent of x. As a result, we verify that the scattering coefficients a, b, c, and d are independent of x:

$$a(\lambda) = W(\varphi_-(x;\lambda)e^{-i\lambda^2 x}, \phi_+(x;\lambda)e^{+i\lambda^2 x}) = W(\varphi_-(0;\lambda), \phi_+(0;\lambda)), \quad (4.73)$$

$$b(\lambda) = W(\varphi_+(x;\lambda)e^{-i\lambda^2 x}, \varphi_-(x;\lambda)e^{-i\lambda^2 x}) = W(\varphi_+(0;\lambda), \varphi_-(0;\lambda)), \quad (4.74)$$

where we have used the Wronskian relation $W(\varphi_+, \phi_+) = 1$, which follows from the boundary conditions (4.68) as $x \to +\infty$.

Now we note the symmetry on solutions to the linear equation (4.14). If ψ is a solution for any $\lambda \in \mathbb{C}$, then $\sigma_1\sigma_3\overline{\psi}$ is also a solution for $\bar{\lambda} \in \mathbb{C}$, where σ_1 and σ_3 are Pauli matrices in (4.12). As a result, using the boundary conditions for the normalized Jost functions, we obtain the following relations:

$$\phi_\pm(x;\lambda) = \sigma_1\sigma_3\overline{\varphi_\pm(x;\overline{\lambda})},$$

where $\overline{\varphi_\pm(x;\overline{\lambda})}$ means that we take complex conjugation of $\varphi_\pm$ constructed from the system of integral equations (4.66) for $\bar{\lambda}$. By applying complex conjugation to the first equation in system (4.72) for $\bar{\lambda}$, multiplying it by $\sigma_1\sigma_3$, and using the relations $\sigma_1\sigma_3 = -\sigma_3\sigma_1$ and $\sigma_1^2 = \sigma_3^2 = 1$, we obtain the second equation in system (4.72) with the correspondence

$$c(\lambda) = -\overline{b(\overline{\lambda})}, \quad d(\lambda) = \overline{a(\overline{\lambda})}, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}. \quad (4.75)$$

From the Wronskian relation $W(\varphi_-, \phi_-) = 1$, which can be established from the boundary conditions (4.68) as $x \to -\infty$, we verify that the transfer matrix in system (4.72) has the determinant equals to unity. In view of the correspondence (4.75), this yields the result

$$a(\lambda)\overline{a(\overline{\lambda})} + b(\lambda)\overline{b(\overline{\lambda})} = 1, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}. \quad (4.76)$$

We now study properties of the scattering coefficients a and b in suitable function spaces. We prove that

$$a(\lambda) \to a_\infty := e^{\frac{1}{2i}\int_{\mathbb{R}} |u|^2 dx} \quad \text{as} \quad |\lambda| \to \infty, \quad (4.77)$$

whereas $a(\lambda) - a_\infty$, $\lambda b(\lambda)$, and $\lambda^{-1}b(\lambda)$ are $H^1_z(\mathbb{R})$ functions with respect to z if u belongs to $H^{1,1}(\mathbb{R})$ defined in (4.13). Moreover, we show that $\lambda b(\lambda)$ is also in $L^{2,1}_z(\mathbb{R})$ if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$.

Lemma 19. *If $u \in H^{1,1}(\mathbb{R})$, then the functions $a(\lambda)$ and $\overline{a(\overline{\lambda})}$ are continued analytically in $\mathbb{C}^+$ and $\mathbb{C}^-$ with respect to z, and, in addition,*

$$a(\lambda) - a_\infty, \ \lambda b(\lambda), \ \lambda^{-1}b(\lambda) \in H^1_z(\mathbb{R}), \quad (4.78)$$

where $a_\infty := e^{\frac{1}{2i}\int_{\mathbb{R}}|u|^2dx}$. *Moreover, if* $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, *then*

$$\lambda b(\lambda),\ \lambda^{-1}b(\lambda) \in L^{2,1}_z(\mathbb{R}). \tag{4.79}$$

Proof. We consider the integral equations (4.66) and (4.67). By taking the limit $x \to +\infty$, which is justified due to Corollary 5 and Remark 10 for every $\lambda \in \mathbb{R}\cup i\mathbb{R}$, and using the scattering relation (4.72) and the transformation (4.64)–(4.65), we obtain

$$a(\lambda) = 1 + \lambda \int_{\mathbb{R}} u(x)\varphi_-^{(2)}(x;\lambda)dx \tag{4.80}$$

and

$$\overline{a(\overline{\lambda})} = 1 - \lambda \int_{\mathbb{R}} \overline{u}(x)\phi_-^{(1)}(x;\lambda)dx. \tag{4.81}$$

It follows from the representations (4.80) and (4.81), as well as Corollary 6, that $a(\lambda)$ is continued analytically in $\mathbb{C}^+$ with respect to z, whereas $\overline{a(\overline{\lambda})}$ is continued analytically in $\mathbb{C}^-$ with respect to z. Using limits (4.28) in Lemma 17 and transformation (4.65), we obtain the following limit for the scattering coefficient $a(\lambda)$ as $|\mathrm{Im}(z)| \to \infty$ along a contour in $\mathbb{C}^+$:

$$\lim_{|z|\to\infty} a(\lambda) = 1 + \frac{1}{2i}\int_{\mathbb{R}} |u(x)|^2 e^{\frac{1}{2i}\int_{-\infty}^x |u(y)|^2dy}dx = e^{\frac{1}{2i}\int_{\mathbb{R}}|u(x)|^2dx} =: a_\infty.$$

In order to prove that $a(\lambda) - a_\infty$ is a $H^1_z(\mathbb{R})$ function, we use the Wronskian representation (4.73). Recall from the transformation (4.64)–(4.65) that

$$\varphi_\pm^{(1)}(x;\lambda) = m_\pm^{(1)}(x;z) \quad \text{and} \quad \phi_\pm^{(2)}(x;\lambda) = n_\pm^{(2)}(x;z).$$

Subtracting the limiting values for a and the normalized Jost functions $m_\pm$ and $n_\pm$, we rewrite the Wronskian representation (4.73) explicitly

$$\begin{aligned} a(\lambda) - a_\infty &= (m_-^{(1)}(0;z) - m_-^\infty(0))(n_+^{(2)}(0;z) - n_+^\infty(0)) \\ &\quad + m_-^\infty(0)(n_+^{(2)}(0;z) - n_+^\infty(0)) + n_+^\infty(0)(m_-^{(1)}(0;z) - m_-^\infty(0)) \\ &\quad - \varphi_-^{(2)}(0;\lambda)\phi_+^{(1)}(0;\lambda). \end{aligned} \tag{4.82}$$

By (4.43) in Lemma 18, all but the last term in (4.82) belong to $H^1_z(\mathbb{R})$. Furthermore, $\lambda^{-1}\varphi_\pm^{(2)}(0;\lambda)$ and $2i\lambda\phi_\pm^{(1)}(0;\lambda) - u(0)n_\pm^\infty(0)$ also belong to $H^1_z(\mathbb{R})$ by Corollary 6. Using the representation (4.82) and the Banach algebra property of $H^1(\mathbb{R})$, we conclude that $a(\lambda) - a_\infty \in H^1_z(\mathbb{R})$.

Next, we analyze the scattering coefficient b. By using the representation (4.64)–(4.65) and the Wronskian representation (4.74), we write

$$2i\lambda b(\lambda) = m_+^{(1)}(0;z)m_-^{(2)}(0;z) - m_+^{(2)}(0;z)m_-^{(1)}(0;z). \tag{4.83}$$

By (4.43) in Lemma 18 (after the corresponding limiting values are subtracted from $m_\pm^{(1)}(0;z)$), we establish that $\lambda b(\lambda) \in H_z^1(\mathbb{R})$. On the other hand, the same Wronskian representation (4.74) can also be written in the form

$$\lambda^{-1}b(\lambda) = m_+^{(1)}(0;z)\lambda^{-1}\varphi_-^{(2)}(0;\lambda) - m_-^{(1)}(0;z)\lambda^{-1}\varphi_+^{(2)}(0;\lambda). \tag{4.84}$$

Recalling that $\lambda^{-1}\varphi_\pm^{(2)}(0;\lambda)$ belongs to $H_z^1(\mathbb{R})$ by Corollary 6, we obtain $\lambda^{-1}b(\lambda) \in H_z^1(\mathbb{R})$. The first assertion (4.78) of the lemma is proved.

To prove the second assertion (4.79) of the lemma, we note that $\lambda^{-1}b(\lambda) \in L_z^{2,1}(\mathbb{R})$ because $z\lambda^{-1}b(\lambda) = \lambda b(\lambda) \in H_z^1(\mathbb{R})$. On the other hand, to show that $\lambda b(\lambda) \in L_z^{2,1}(\mathbb{R})$, we multiply equation (4.83) by z and write the resulting equation in the form

$$\begin{aligned} 2i\lambda z b(\lambda) &= m_+^{(1)}(0;z)\left(zm_-^{(2)}(0;z) - q_-^{(2)}(0)\right) - m_-^{(1)}(0;z)\left(zm_+^{(2)}(0;z) - q_+^{(2)}(0)\right) \\ &\quad + q_-^{(2)}(0)\left(m_+^{(1)}(0;z) - m_+^{\infty}(0)\right) - q_+^{(2)}(0)\left(m_-^{(1)}(0;z) - m_-^{\infty}(0)\right), \end{aligned} \tag{4.85}$$

where we have used the identity $q_-^{(2)}(0)m_+^{\infty}(0) - q_+^{(2)}(0)m_-^{\infty}(0) = 0$ that follows from limits (4.28) and (4.30). By (4.43) and (4.44) in Lemma 18, all the terms in the representation (4.85) are in $L_z^2(\mathbb{R})$, hence $\lambda b(\lambda) \in L_z^{2,1}(\mathbb{R})$. The second assertion (4.79) of the lemma is proved. □

We show that the mapping

$$H^{1,1}(\mathbb{R}) \ni u \to a(\lambda) - a_\infty, \lambda b(\lambda), \lambda^{-1}b(\lambda) \in H_z^1(\mathbb{R}) \tag{4.86}$$

is Lipschitz continuous. Moreover, we also have Lipschitz continuity of the mapping

$$H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}) \ni u \to \lambda b(\lambda), \lambda^{-1}b(\lambda) \in L_z^{2,1}(\mathbb{R}). \tag{4.87}$$

The corresponding result is deduced from Lemma 19 and Corollary 4.

Corollary 7. *Let $u, \tilde{u} \in H^{1,1}(\mathbb{R})$ satisfy $\|u\|_{H^{1,1}}, \|\tilde{u}\|_{H^{1,1}} \le U$ for some $U > 0$. Denote the corresponding scattering coefficients by (a,b) and $(\tilde{a},\tilde{b})$ respectively. Then, there is a positive U-dependent constant $C(U)$ such that*

$$\begin{aligned} &\|a(\lambda) - a_\infty - \tilde{a}(\lambda) + \tilde{a}_\infty\|_{H_z^1} + \|\lambda b(\lambda) - \lambda\tilde{b}(\lambda)\|_{H_z^1} \\ &\qquad + \|\lambda^{-1}b(\lambda) - \lambda^{-1}\tilde{b}(\lambda)\|_{H_z^1} \le C(U)\|u - \tilde{u}\|_{H^{1,1}}. \end{aligned} \tag{4.88}$$

Moreover, if $u, \tilde{u} \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ satisfy $\|u\|_{H^2\cap H^{1,1}}, \|\tilde{u}\|_{H^2\cap H^{1,1}} \le U$, then there is a positive U-dependent constant $C(U)$ such that

$$\|\lambda b(\lambda) - \lambda\tilde{b}(\lambda)\|_{L_z^{2,1}} + \|\lambda^{-1}b(\lambda) - \lambda^{-1}\tilde{b}(\lambda)\|_{L_z^{2,1}} \le C(U)\|u - \tilde{u}\|_{H^2\cap H^{1,1}}. \tag{4.89}$$

Proof. The assertion follows from the representations (4.82), (4.83), (4.84), and (4.85), as well as the Lipschitz continuity of the Jost functions $m_\pm$ and $n_\pm$ established in Corollary 4. □

Remark 12. *Since Corollary 7 yields Lipschitz continuity of the mappings (4.86) and (4.87) for every u, $\tilde{u}$ in a ball of a fixed (but possibly large) radius U, the mappings (4.86) and (4.87) are one-to-one for every u in the ball.*

Another result, which follows from Lemma 19, is the parity property of the scattering coefficients a and b with respect to λ. The corresponding result is given in the following corollary.

Corollary 8. *The scattering coefficients a and b are even and odd functions in λ for $\lambda \in \mathbb{R} \cup i\mathbb{R}$. Moreover, they satisfy the following scattering relation*

$$\begin{cases} |a(\lambda)|^2 + |b(\lambda)|^2 = 1, & \lambda \in \mathbb{R}, \\ |a(\lambda)|^2 - |b(\lambda)|^2 = 1, & \lambda \in i\mathbb{R}. \end{cases} \tag{4.90}$$

Proof. Because $a(\lambda)$ and $\lambda^{-1}b(\lambda)$ are functions of $z = \lambda^2$, as follows from Lemma 19, we have $a(-\lambda) = a(\lambda)$ and $b(-\lambda) = -b(\lambda)$ for all $\lambda \in \mathbb{R} \cup i\mathbb{R}$. For $\lambda \in \mathbb{R}$, the scattering relation (4.76) yields the first line of (4.90). For $\lambda = i\gamma$ with $\gamma \in \mathbb{R}$, the parity properties of a and b imply

$$\overline{a(\bar{\lambda})} = \overline{a(-i\gamma)} = \overline{a(i\gamma)} = \overline{a(\lambda)} \quad \text{and} \quad \overline{b(\bar{\lambda})} = \overline{b(-i\gamma)} = -\overline{b(i\gamma)} = -\overline{b(\lambda)}.$$

Substituting these relations to the scattering relation (4.76), we obtain the second line of (4.90) □

4.3 Formulations of the Riemann–Hilbert problem

We deduce the Riemann–Hilbert problem of complex analysis from the jump condition for normalized Jost functions on $\mathbb{R} \cup i\mathbb{R}$ in the λ plane, which corresponds to $\mathbb{R}$ in the z plane, where $z = \lambda^2$. The jump condition yields boundary conditions for the Jost functions extended to sectionally analytic functions in different domains of the corresponding complex plane. In the beginning, we derive the jump condition in the λ plane by using the Jost functions of the original Kaup–Newell spectral problem (4.14).

Let us define the reflection coefficient by

$$r(\lambda) := \frac{b(\lambda)}{a(\lambda)}, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}. \tag{4.91}$$

Each zero of a on $\mathbb{R} \cup i\mathbb{R}$ corresponds to the resonance, according to Definition 4. By the assumptions of Theorem 4, the spectral problem (4.14) admits no resonances, therefore, there exists a positive number A such that

$$|a(\lambda)| \geq A > 0, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}. \tag{4.92}$$

Thus, $r(\lambda)$ is well-defined for every $\lambda \in \mathbb{R} \cup i\mathbb{R}$.

Under the condition (4.92), the scattering relations (4.72) with (4.75) can be rewritten in the equivalent form:

$$\frac{\varphi_-(x;\lambda)}{a(\lambda)} - \varphi_+(x;\lambda) = r(\lambda)e^{2i\lambda^2 x}\phi_+(x;\lambda) \tag{4.93}$$

and

$$\frac{\phi_-(x;\lambda)}{\overline{a(\bar{\lambda})}} - \phi_+(x;\lambda) = -\overline{r(\bar{\lambda})}e^{-2i\lambda^2 x}\varphi_+(x;\lambda), \tag{4.94}$$

where $\lambda \in \mathbb{R} \cup i\mathbb{R}$.

By Lemma 19, $a(\lambda)$ is continued analytically in the first and third quadrants of the λ plane, where $\text{Im}(\lambda^2) > 0$. Also $a(\lambda)$ approaches to a finite limit $a_\infty \neq 0$ as $|\lambda| \to \infty$. By a theorem of complex analysis on zeros of analytic functions, a has at most finite number of zeros in each quadrant of the λ plane. Each zero of a corresponds to an eigenvalue of the spectral problem (4.14) with the $L^2(\mathbb{R})$ solution $\psi(x)$ decaying to zero exponentially fast as $|x| \to \infty$. Indeed, this follows from the Wronskian relation (4.73) between the Jost functions φ_- and ψ_+ extended to the first and third quadrant of the λ plane by Corollary 6. By the assumptions of Theorem 4, the spectral problem (4.14) admits no eigenvalues, hence the bound (4.92) is extended to the first and third quadrants of the λ plane. Therefore, the functions $\frac{\varphi_-(x;\lambda)}{a(\lambda)}$ and $\frac{\phi_-(x;\lambda)}{\overline{a(\bar{\lambda})}}$ are analytic in the corresponding domains of the λ plane.

From the scattering relations (4.93) and (4.94), we can define the complex functions

$$\Phi_+(x;\lambda) := \left[\frac{\varphi_-(x;\lambda)}{a(\lambda)}, \phi_+(x;\lambda)\right], \quad \Phi_-(x;\lambda) := \left[\varphi_+(x;\lambda), \frac{\phi_-(x;\lambda)}{\overline{a(\bar{\lambda})}}\right]. \tag{4.95}$$

By Corollary 6, Lemma 19, and the condition (4.92) on a, for every $x \in \mathbb{R}$, the function $\Phi_+(x;\cdot)$ is analytic in the first and third quadrants of the λ plane, whereas the function $\Phi_-(x;\cdot)$ is analytic in the second and fourth quadrants of the λ plane. For every $x \in \mathbb{R}$ and $\lambda \in \mathbb{R} \cup i\mathbb{R}$, the two functions are related by the jump condition

$$\Phi_+(x;\lambda) - \Phi_-(x;\lambda) = \Phi_-(x;\lambda)S(x;\lambda), \tag{4.96}$$

where

$$S(x;\lambda) := \begin{bmatrix} |r(\lambda)|^2 & \overline{r(\lambda)}e^{-2i\lambda^2 x} \\ r(\lambda)e^{2i\lambda^2 x} & 0 \end{bmatrix}, \quad \lambda \in \mathbb{R} \tag{4.97}$$

and

$$S(x;\lambda) := \begin{bmatrix} -|r(\lambda)|^2 & -\overline{r(\lambda)}e^{-2i\lambda^2 x} \\ r(\lambda)e^{2i\lambda^2 x} & 0 \end{bmatrix}, \quad \lambda \in i\mathbb{R}. \tag{4.98}$$

Note that $r(-\lambda) = -r(\lambda)$ by Corollary 8, so that $r(0) = 0$. By Corollary 6, the functions $\Phi_\pm(x;\lambda)$ satisfy the limiting behavior as $|\lambda| \to \infty$ along a contour in the corresponding domains of their analyticity in the λ plane:

$$\Phi_\pm(x;\lambda) \to \Phi_\infty(x) := \left[e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} e_1, \quad e^{-\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} e_2 \right] \quad \text{as} \quad |\lambda| \to \infty. \tag{4.99}$$

The jump conditions (4.96) and the boundary conditions (4.99) set up a Riemann–Hilbert problem to find sectionally analytic functions $\Phi(x;\cdot)$ for every $x \in \mathbb{R}$. It is quite remarkable that the matrix S is Hermitian for $\lambda \in \mathbb{R}$. In this case, we can use the theory of Zhou [122] to obtain a unique solution to the Riemann–Hilbert problem (4.96), (4.97), and (4.99). However, the matrix S is not Hermitian for $\lambda \in i\mathbb{R}$. Nevertheless, the second scattering relation (4.90) yields a useful constraint:

$$1 - |r(\lambda)|^2 = \frac{1}{|a(\lambda)|^2} \geq c_0^2 > 0, \quad \lambda \in i\mathbb{R}, \tag{4.100}$$

where $c_0 := \sup_{\lambda \in i\mathbb{R}} |a(\lambda)|$. The constraint (4.100) will be used to obtain a unique solution to the Riemann–Hilbert problem (4.96), (4.98), and (4.99).

We note that only the latter case (4.98), which is relevant to the imaginary values of λ, was considered in the context of the Kaup–Newell spectral problem by Kitaev & Vartanian [59], who studied the long time asymptotic solution of the derivative NLS equation (4.1, also in the case of no solitons. The smallness condition (4.100) does not need to be assumed a priori, as it is done in Lemma 2.2 in [59], but appears naturally from the second scattering relation (4.90). The Hermitian case of real values of λ was missed in [59].

We also note that the scattering matrix $S(x;\lambda)$ is analogous to the one known for the focusing NLS equation if $\lambda \in \mathbb{R}$ and the one known for the defocusing NLS equation if $\lambda \in i\mathbb{R}$. As a result, the inverse scattering transform for the derivative NLS equation combines elements of the inverse scattering transforms developed for the focusing and defocusing cubic NLS equations [29, 31, 121].

In the rest of this section, we reformulate the jump condition in the z plane and introduce two scattering coefficients $r_\pm$, which are defined on the real line in the function space $H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$. The scattering coefficients $r_\pm$ allow us to recover a potential u in the function space $H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ (in Section 4).

4.3.1 Reformulation of the Riemann–Hilbert problem

Using transformation matrices in (4.64)–(4.65), we can rewrite the scattering relations (4.93) and (4.94) in terms of the z-dependent Jost functions $m_\pm$ and $n_\pm$:

$$\frac{m_-(x;z)}{a(\lambda)} - m_+(x;z) = \frac{2i\lambda b(\lambda)}{a(\lambda)} e^{2izx} p_+(x;z) \tag{4.101}$$

and

$$\frac{p_-(x;z)}{\overline{a(\bar{\lambda})}} - p_+(x;z) = -\frac{\overline{b(\bar{\lambda})}}{2i\bar{\lambda}\overline{a(\bar{\lambda})}} e^{-2izx} m_+(x;z), \tag{4.102}$$

where $z \in \mathbb{R}$, $m_\pm$ are defined by Lemma 16, and $p_\pm$ are given explicitly by

$$p_\pm(x;z) = \frac{1}{2i\lambda} T_1(x;\lambda) T_2^{-1}(x;\lambda) n_\pm(x;z) = -\frac{1}{4z} \begin{bmatrix} 1 & u(x) \\ -\bar{u}(x) & -|u(x)|^2 - 4z \end{bmatrix} n_\pm(x;z). \tag{4.103}$$

Properties of the new functions $p_\pm$ are summarized in the following result.

Lemma 20. *Under the conditions of Lemma 16, for every $x \in \mathbb{R}$, the functions $p_\pm(x;z)$ are continued analytically in $\mathbb{C}^\pm$ and satisfy the following limits as $|\mathrm{Im}(z)| \to \infty$ along a contour in the domains of their analyticity:*

$$\lim_{|z|\to\infty} p_\pm(x;z) = n_\pm^\infty(x) e_2, \tag{4.104}$$

where $n_\pm^\infty$ are the same as in the limits (4.29).

Proof. The asymptotic limits (4.104) follow from the representation (4.103) and the asymptotic limits (4.29) for $n_\pm(x;z)$ as $|z| \to \infty$ in Lemma 17. Using the transformation (4.64)–(4.65), functions $p_\pm$ can be written in the equivalent form

$$p_\pm(x;z) = n_\pm^{(2)}(x;z) e_2 + \frac{1}{2i\lambda} \begin{bmatrix} 1 \\ -\bar{u}(x) \end{bmatrix} \phi_\pm^{(1)}(x;\lambda), \tag{4.105}$$

where both $n_\pm^{(2)}(x;z)$ and $\lambda^{-1}\phi_\pm^{(1)}(x;\lambda)$ are continued analytically in $\mathbb{C}^\pm$ with respect to z by Lemma 16 and Corollary 6. From the Volterra integral equation (4.67), we also obtain

$$\lambda^{-1}\phi_\pm^{(1)}(x;\lambda) = \int_{\pm\infty}^{x} e^{-2iz(x-y)} u(y) n_\pm^{(2)}(y;z) dy, \tag{4.106}$$

therefore, $p_\pm(x;0)$ exists for every $x \in \mathbb{R}$. Thus, for every $x \in \mathbb{R}$, the analyticity properties of $p_\pm(x;\cdot)$ are the same as those of $n_\pm(x;\cdot)$. □

Let us now introduce the new scattering data:

$$r_+(z) := -\frac{b(\lambda)}{2i\lambda a(\lambda)}, \quad r_-(z) := \frac{2i\lambda b(\lambda)}{a(\lambda)}, \quad z \in \mathbb{R}. \tag{4.107}$$

which satisfy the relation

$$r_-(z) = 4z r_+(z), \quad z \in \mathbb{R}. \tag{4.108}$$

It is worthwhile noting that

$$\begin{cases} \overline{r}_+(z) r_-(z) = |r(\lambda)|^2, & z \in \mathbb{R}^+, \quad \lambda \in \mathbb{R}, \\ \overline{r}_+(z) r_-(z) = -|r(\lambda)|^2, & z \in \mathbb{R}^-, \quad \lambda \in i\mathbb{R}. \end{cases} \tag{4.109}$$

The scattering data $r_\pm$ satisfy the following properties, which are derived from the previous results.

Lemma 21. *Assume the condition (4.92) on a. If $u \in H^{1,1}(\mathbb{R})$, then $r_\pm \in H^1(\mathbb{R})$, whereas if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, then $r_\pm \in L^{2,1}(\mathbb{R})$. Moreover, the mapping*

$$H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}) \ni u \to (r_+, r_-) \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R}) \tag{4.110}$$

is Lipschitz continuous.

Proof. The first assertion on $r_\pm$ follows from Lemma 19. To prove Lipschitz continuity of the mapping (4.110), we use the following representation for r_- and $\tilde{r}_-$ that correspond to two potentials u and $\tilde{u}$,

$$r_- - \widetilde{r}_- = \frac{2i\lambda(b-\widetilde{b})}{a} + \frac{2i\lambda\widetilde{b}}{a\widetilde{a}}[(\widetilde{a}-\widetilde{a}_\infty)-(a-a_\infty)] + \frac{2i\lambda\widetilde{b}}{a\widetilde{a}}(\widetilde{a}_\infty - a_\infty). \tag{4.111}$$

Lipschitz continuity of the mapping (4.110) for r_- follows from the representation (4.111) and Corollary 7. Lipschitz continuity of the mapping (4.110) for r_+ is studied by using a representation similar to (4.111). □

Remark 13. *By Corollary 8, $a(-\lambda) = a(\lambda)$ for every $\lambda \in \mathbb{R} \cup i\mathbb{R}$. Therefore, when we introduce $z = \lambda^2$ and start considering functions of z, it makes sense to introduce $\mathbf{a}(z) := a(\lambda)$ for every $z \in \mathbb{R}$. In what follows, we drop the bold notations in the definition of $a(z)$.*

For every $x \in \mathbb{R}$ and $z \in \mathbb{R}$, we define two matrices $P_+(x;z)$ and $P_-(x;z)$ by

$$P_+(x;z) := \left[\frac{m_-(x;z)}{a(z)},\ p_+(x;z)\right], \quad P_-(x;z) := \left[m_+(x;z),\ \frac{p_-(x;z)}{\overline{a}(z)}\right]. \tag{4.112}$$

By Lemmas 16, 19, and 20, as well as the condition (4.92) on a, the functions $P_\pm(x;\cdot)$ for every $x \in \mathbb{R}$ are continued analytically in $\mathbb{C}^\pm$. The scattering relations (4.101) and (4.102) are now rewritten as the jump condition between functions $P_\pm(x;z)$ across the real axis in z for every $x \in \mathbb{R}$:

$$P_+(x;z) - P_-(x;z) = P_-(x;z)R(x;z), \quad R(x;z) := \begin{bmatrix} \overline{r}_+(z)r_-(z) & \overline{r}_+(z)e^{-2izx} \\ r_-(z)e^{2izx} & 0 \end{bmatrix}. \tag{4.113}$$

By Lemmas 17, 19, and 20, the functions $P_\pm(x;\cdot)$ satisfy the limiting behavior as $|z| \to \infty$ along a contour in the domain of their analyticity in the z plane:

$$P_\pm(x;z) \to \Phi_\infty(x) \quad \text{as} \quad |z| \to \infty, \tag{4.114}$$

where Φ_∞ is the same as in (4.99). The boundary conditions (4.114) depend on x, which represents an obstacle in the inverse scattering transform, where we reconstruct the potential $u(x)$ from the behavior of the analytic continuations of the Jost functions $P_\pm(x;\cdot)$ for $x \in \mathbb{R}$. Therefore, we fix the boundary conditions to the identity matrix by defining new matrices

$$M_\pm(x;z) := [\Phi_\infty(x)]^{-1} P_\pm(x;z), \quad x \in \mathbb{R}, \quad z \in \mathbb{C}^\pm. \tag{4.115}$$

As a result, we obtain the Riemann–Hilbert problem for analytic functions $M_\pm(x;\cdot)$ in $\mathbb{C}^\pm$, which is given by the jump condition equipped with the uniform boundary conditions:

$$\begin{cases} M_+(x;z) - M_-(x;z) = M_-(x;z)R(x;z), & z \in \mathbb{R}, \\ M_\pm(x;z) \to I & \text{as } |z| \to \infty. \end{cases} \tag{4.116}$$

The scattering data $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ are defined in Lemma 21.

Figure 4.1 shows the regions of analyticity of functions $\Phi_\pm$ in the λ plane (left) and those of functions $M_\pm$ in the z plane (right).

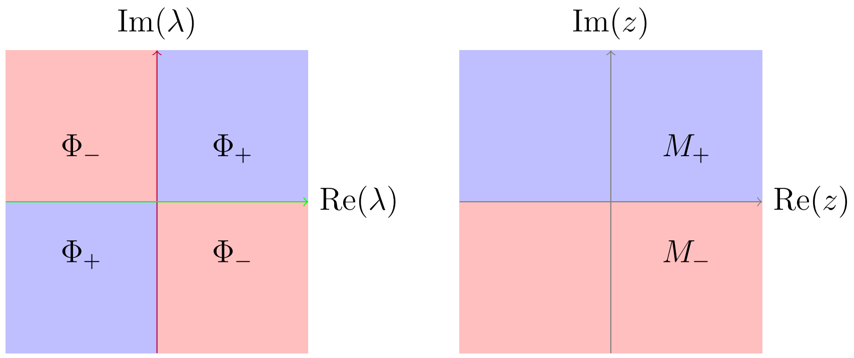


Figure 4.1: Blue and red regions mark domains of analyticity of $\Phi_\pm$ in the λ plane (left) and those of $M_\pm$ in the z plane (right).

The scattering matrix R in the Riemann–Hilbert problem (4.116) is not Hermitian. As a result, it is difficult to use the theory of Zhou [122] in order to construct a unique solution for $M_\pm$ in the Riemann–Hilbert problem (4.116) without restricting the scattering data $r_\pm$ to be small in their norms. On the other hand, the original Riemann–Hilbert problem (4.96) in the λ plane does not have these limitations. Therefore, in the following subsection, we consider two equivalent reductions of the Riemann–Hilbert problem (4.116) in the z plane to those related with the scattering matrix S instead of the scattering matrix R.

4.3.2 Two transformations of the Riemann-Hilbert problem

For every $\lambda \in \mathbb{C}\backslash\{0\}$, we denote

$$\tau_1(\lambda) := \begin{bmatrix} 1 & 0 \\ 0 & 2i\lambda \end{bmatrix}, \quad \tau_2(\lambda) := \begin{bmatrix} (2i\lambda)^{-1} & 0 \\ 0 & 1 \end{bmatrix} \tag{4.117}$$

and observe that

$$\tau_1^{-1}(\lambda)R(x;z)\tau_1(\lambda) = \tau_2^{-1}(\lambda)R(x;z)\tau_2(\lambda) = S(x;\lambda), \quad z \in \mathbb{R}, \quad \lambda \in \mathbb{R} \cup i\mathbb{R},$$

where $S(x;\lambda)$ is defined in (4.97) and (4.98), whereas $R(x;z)$ is defined in (4.113). Using these properties, we introduce two formally equivalent reformulations of the

Riemann–Hilbert problem (4.116):

$$\begin{cases} G_{+1,2}(x;\lambda) - G_{-1,2}(x;\lambda) = G_{-1,2}(x;\lambda)S(x;\lambda) + F_{1,2}(x;\lambda), & \lambda \in \mathbb{R} \cup i\mathbb{R}, \\ \lim_{|\lambda|\to\infty} G_{\pm 1,2}(x;\lambda) = 0, \end{cases} \tag{4.118}$$

where

$$G_{\pm 1,2}(x;\lambda) := M_\pm(x;z)\tau_{1,2}(\lambda) - \tau_{1,2}(\lambda), \quad F_{1,2}(x;\lambda) := \tau_{1,2}(\lambda)S(x;\lambda). \tag{4.119}$$

The functions $G_{+1,2}(x;\lambda)$ are analytic in the first and third quadrants of the λ plane, whereas the functions $G_{-1,2}(x;\lambda)$ are analytic in the second and fourth quadrants of the λ plane. Although the behavior of functions $M_\pm(x;z)\tau_{1,2}(\lambda)$ may become singular as $\lambda \to 0$, we prove in Corollary 9 below that $G_{\pm 1,2}(x;\lambda)$ are free of singularities as $\lambda \to 0$.

Figure 4.2 summarizes on the transformations of the Riemann–Hilbert problems.

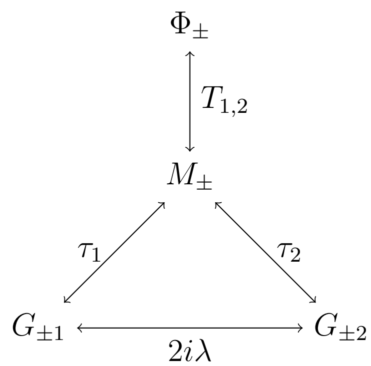


Figure 4.2: A useful diagram showing transformations of the Riemann–Hilbert problems

Solvability of the Riemann–Hilbert problem (4.118) is obtained in Section 4.1. Then, in Section 4.2, we show that the solution to the two related Riemann-Hilbert problems (4.118) can be used to obtain the solution to the Riemann-Hilbert problem (4.116). In Section 4.3, we show how this procedure defines the inverse scattering transform to recover the potential u of the Kaup–Newell spectral problem (4.14) from the scattering data $r_\pm$.

4.4 Inverse scattering transform

We are now concerned with the solvability of the Riemann-Hilbert problem (4.116) for the given scattering data $r_+, r_- \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ satisfying the constraint (4.108). We are looking for analytic matrix functions $M_\pm(x;\cdot)$ in $\mathbb{C}^\pm$ for every $x \in \mathbb{R}$. Let us introduce the following notations for the column vectors of the matrices $M_\pm$ as

$$M_\pm(x;z) = [\mu_\pm(x;z), \eta_\pm(x;z)]. \tag{4.120}$$

Before we proceed, let us inspect regularity of the reflection coefficient $r(\lambda)$ as a function of z on $\mathbb{R}$.

Proposition 5. *If $r_\pm(z) \in H^1_z(\mathbb{R}) \cap L^{2,1}_z(\mathbb{R})$, then $r(\lambda) \in L^{2,1}_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$.*

Proof. Since $r_\pm \in L^{2,1}(\mathbb{R})$ and $|r(\lambda)|^2 = \text{sign}(z)\, \overline{r}_+(z) r_-(z)$ for every $z \in \mathbb{R}$, we have $r(\lambda) \in L^{2,1}_z(\mathbb{R})$ by Cauchy–Schwarz inequality.

To show that $r(\lambda) \in L^\infty_z(\mathbb{R})$, we notice that $r(\lambda)$ can be defined equivalently from (4.107) in the following form:

$$r(\lambda) = \begin{cases} -2i\lambda r_+(z) & |\lambda| \leq 1 \\ (2i\lambda)^{-1} r_-(z) & |\lambda| \geq 1. \end{cases}$$

Since $r_\pm \in L^\infty(\mathbb{R})$ as it follows from $r_\pm \in H^1(\mathbb{R})$, then we have $r(\lambda) \in L^\infty_z(\mathbb{R})$. □

Remark 14. *We do not expect generally that $r(\lambda)$ belongs to $H^1_z(\mathbb{R})$. For instance, if*

$$h(\lambda) := \frac{\lambda}{(1+\lambda^4)^s}, \quad s > \frac{5}{4},$$

then $\lambda h(\lambda), \lambda^{-1} h(\lambda) \in H^1_z(\mathbb{R}) \cap L^{2,1}_z(\mathbb{R})$, $h(\lambda) \in L^{2,1}_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$ but $h(\lambda) \notin H^1_z(\mathbb{R})$.

We also note another useful elementary result.

Proposition 6. *If $r_-(z) \in H^1_z(\mathbb{R}) \cap L^{2,1}_z(\mathbb{R})$, then $\|\lambda r_-(z)\|_{L^\infty_z} \leq \|r_-\|_{H^1 \cap L^{2,1}}$.*

Proof. The result follows from the representation

$$z r_-(z)^2 = \int_0^z \left(r_-(z)^2 + 2z r_-(z) r_-'(z) \right) dz.$$

Using Cauchy–Schwarz inequality for $r_-(z) \in H^1_z(\mathbb{R}) \cap L^{2,1}_z(\mathbb{R})$, we obtain the desired bound. □

4.4.1 Solution to the Riemann–Hilbert problem

Let us start with the definition of the Cauchy operator, which can be found in many sources, e.g., in [31]. For any function $h \in L^p(\mathbb{R})$ with $1 \leq p < \infty$, the Cauchy operator denoted by $\mathcal{C}$ is given by

$$\mathcal{C}(h)(z) := \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{h(s)}{s-z} ds, \quad z \in \mathbb{C} \setminus \mathbb{R}. \tag{4.121}$$

The function $\mathcal{C}(h)$ is analytic off the real line such that $\mathcal{C}(h)(\cdot + iy)$ is in $L^p(\mathbb{R})$ for each $y \neq 0$. When z approaches to a point on the real line transversely from the upper and lower half planes, that is, if $y \to \pm 0$, the Cauchy operator $\mathcal{C}$ becomes the Plemelj projection operators, denoted respectively by $\mathcal{P}^\pm$. These projection operators are given explicitly by

$$\mathcal{P}^\pm(h)(z) := \lim_{\epsilon \downarrow 0} \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{h(s)}{s - (z \pm \epsilon i)} ds, \quad z \in \mathbb{R}. \tag{4.122}$$

The following proposition summarizes the basic properties of the Cauchy and projection operators.

Proposition 7. *For every $h \in L^p(\mathbb{R})$, $1 \leq p < \infty$, the Cauchy operator $\mathcal{C}(h)$ is analytic off the real line, decays to zero as $|z| \to \infty$, and approaches to $\mathcal{P}^{\pm}(h)$ almost everywhere, when a point $z \in \mathbb{C}^{\pm}$ approaches to a point on the real axis by any non-tangential contour from $\mathbb{C}^{\pm}$. If $1 < p < \infty$, then there exists a positive constant C_p (with $C_{p=2} = 1$) such that*

$$\|\mathcal{P}^{\pm}(h)\|_{L^p} \leq C_p \|h\|_{L^p}. \tag{4.123}$$

If $h \in L^1(\mathbb{R})$, then the Cauchy operator admits the following asymptotic limit in either $\mathbb{C}^+$ or $\mathbb{C}^-$:

$$\lim_{|z|\to\infty} z\mathcal{C}(h)(z) = -\frac{1}{2\pi i}\int_{\mathbb{R}} h(s)ds. \tag{4.124}$$

Proof. Analyticity, decay, and boundary values of $\mathcal{C}$ on the real axis follow from Theorem 11.2 and Corollary 2 on pp. 190–191 in [35]. By Sokhotski–Plemelj theorem, we have the relations

$$\mathcal{P}^{\pm}(h)(z) = \pm\frac{1}{2}h(z) - \frac{i}{2}\mathcal{H}(h)(z), \quad z \in \mathbb{R}, \tag{4.125}$$

where $\mathcal{H}$ is the Hilbert transform given by

$$\mathcal{H}(h)(z) := \frac{1}{\pi}\lim_{\epsilon\downarrow 0}\left(\int_{-\infty}^{z-\epsilon} + \int_{z+\epsilon}^{\infty}\right)\frac{h(s)}{s-z}ds, \quad z \in \mathbb{R}.$$

By Riesz's theorem (Theorem 3.2 in [34]), $\mathcal{H}$ is a bounded operator from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$ for every $1 < p < \infty$, so that the bound (4.123) holds with $C_2 = 1$ and $C_p \to +\infty$ as $p \to 1$ and $p \to \infty$. Finally, the asymptotic limit (4.124) is justified by Lebesgue's dominated convergence theorem if $h \in L^1(\mathbb{R})$. □

We recall the scattering matrix $S(x;\lambda)$ given explicitly by (4.97) and (4.98). The following proposition states that if $r(\lambda)$ is bounded and satisfies (4.100), then the quadratic form associated with the matrix $I + S(x;\lambda)$ is strictly positive for every $x \in \mathbb{R}$ and every $\lambda \in \mathbb{R} \cup i\mathbb{R}$, whereas the matrix $I + S(x;\lambda)$ is bounded. In what follows, $\|\cdot\|$ denotes the Euclidean norm of vectors in $\mathbb{C}^2$.

Proposition 8. *For every $r(\lambda) \in L^{\infty}_z(\mathbb{R})$ satisfying (4.100), there exist positive constants C_- and C_+ such that for every $x \in \mathbb{R}$ and every column-vector $g \in \mathbb{C}^2$, we have*

$$\mathrm{Re}\, g^t \left(I + S(x;\lambda)\right) g \geq C_- g^t g, \quad \lambda \in \mathbb{R} \cup i\mathbb{R} \tag{4.126}$$

and

$$\|\left(I + S(x;\lambda)\right) g\| \leq C_+ \|g\|, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}. \tag{4.127}$$

Proof. For $\lambda \in \mathbb{R}$, we use representation (4.97). Since $I + S(x;\lambda)$ is Hermitian for every $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, we compute the two real eigenvalues of $I + S(x;\lambda)$ given by

$$\mu_\pm(\lambda) = 1 + \frac{1}{2}|r(\lambda)|^2 \pm |r(\lambda)|\sqrt{1+\frac{1}{4}|r(\lambda)|^2} = \left(\sqrt{1+\frac{1}{4}|r(\lambda)|^2} \pm \frac{1}{2}|r(\lambda)|\right)^2 > 0.$$

Note that

$$\frac{1}{(1+|r(\lambda)|)^2} \le \mu_-(\lambda) \le \mu_+(\lambda) \le (1+|r(\lambda)|)^2, \quad \lambda \in \mathbb{R}.$$

It follows from the above inequalities that the bounds (4.126) and (4.127) for $\lambda \in \mathbb{R}$ hold with

$$C_- := \frac{1}{(1+\sup_{\lambda\in\mathbb{R}}|r(\lambda)|)^2} > 0 \quad \text{and} \quad C_+ := (1+\sup_{\lambda\in\mathbb{R}}|r(\lambda)|)^2 < \infty.$$

For $\lambda \in i\mathbb{R}$, we use representation (4.98). Since $I + S(x;\lambda)$ is no longer Hermitian, we define the Hermitian part of $S(x;\lambda)$ by

$$S_H(\lambda) := \frac{1}{2}S(x;\lambda) + \frac{1}{2}S^*(x;\lambda) = \begin{bmatrix} -|r(\lambda)|^2 & 0 \\ 0 & 0 \end{bmatrix},$$

where the asterisk denotes Hermite conjugate (matrix transposition and complex conjugate). It follows from (4.100) that $\sup_{\lambda\in i\mathbb{R}} r(\lambda) \le 1 - c_0^2 < 1$ so that the diagonal matrix $I + S_H(\lambda)$ is positive definite for every $\lambda \in i\mathbb{R}$. The bound (4.126) for $\lambda \in i\mathbb{R}$ follows from this estimate with $C_- := 1 - \sup_{\lambda\in i\mathbb{R}} |r(\lambda)|^2 \ge c_0^2 > 0$. Finally, estimating componentwise

$$\begin{aligned} \|(I+S(x;\lambda))g\|^2 &\le (1+|r(\lambda)|^2)\|g\|^2 + |r(\lambda)|^2\left(r(\lambda)g^{(1)}\overline{g^{(2)}} + \overline{r(\lambda)g^{(1)}}g^{(2)}\right) \\ &\le \left(1+|r(\lambda)|^2\right)\left(1+\frac{1}{2}|r(\lambda)|^2\right)\|g\|^2, \end{aligned}$$

we obtain the bound (4.127) for $\lambda \in i\mathbb{R}$ with $C_+ := (1+\sup_{\lambda\in i\mathbb{R}}|r(\lambda)|^2) < \infty$. □

Thanks to the result of Proposition 8, we shall prove solvability of the two related Riemann–Hilbert problems (4.118) by using the method of Zhou [122]. Dropping the subscripts, we rewrite the two related Riemann–Hilbert problems (4.118) in the following abstract form

$$\begin{cases} G_+(x;\lambda) - G_-(x;\lambda) = G_-(x;\lambda)S(x;\lambda) + F(x;\lambda), & \lambda \in \mathbb{R}\cup i\mathbb{R}, \\ G_\pm(x,\lambda) \to 0 & \text{as } |\lambda| \to \infty. \end{cases} \tag{4.128}$$

If $r_\pm \in H_z^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$, then Proposition 5 implies that $S(x;\lambda) \in L_z^1(\mathbb{R}) \cap L_z^\infty(\mathbb{R})$ and $F(x;\lambda) \in L_z^2(\mathbb{R})$ for every $x \in \mathbb{R}$. We consider the class of solutions to the Riemann–Hilbert problem (4.128) such that for every $x \in \mathbb{R}$,

- $G_\pm(x;\lambda)$ are analytic functions of $z = \lambda^2$ in $\mathbb{C}^\pm$
- $G_\pm(x;\lambda) \in L_z^2(\mathbb{R})$

- The same columns of $G_\pm(x;\lambda)$, $G_-(x;\lambda)S(x;\lambda)$, and $F(x;\lambda)$ are either even or odd in λ.

By Proposition 7 with $p = 2$, for every $x \in \mathbb{R}$, the Riemann-Hilbert problem (4.128) has a solution given by the Cauchy operator

$$G_\pm(x;\lambda) = \mathcal{C}\left(G_-(x;\lambda)S(x;\lambda) + F(x;\lambda)\right)(z), \quad z \in \mathbb{C}^\pm \tag{4.129}$$

if and only if there is a solution $G_-(x;\lambda) \in L^2_z(\mathbb{R})$ of the Fredholm integral equation:

$$G_-(x;\lambda) = \mathcal{P}^-\left(G_-(x;\lambda)S(x;\lambda) + F(x;\lambda)\right)(z), \quad z \in \mathbb{R}. \tag{4.130}$$

Once $G_-(x;\lambda) \in L^2_z(\mathbb{R})$ is found from the Fredholm integral equation (4.130), then $G_+(x;\lambda) \in L^2_z(\mathbb{R})$ is obtained from the projection formula

$$G_+(x;\lambda) = \mathcal{P}^+\left(G_-(x;\lambda)S(x;\lambda) + F(x;\lambda)\right)(z), \quad z \in \mathbb{R}. \tag{4.131}$$

Remark 15. *The complex integrals in $\mathcal{C}$ and $\mathcal{P}^\pm$ over the real line $z = \lambda^2$ can be parameterized by λ on $\mathbb{R}^+ \cup i\mathbb{R}^+$. Extensions of integral representations (4.129), (4.130), and (4.131) for $\lambda \in \mathbb{R}^- \cup i\mathbb{R}^-$ is performed with the account of parity symmetries of the corresponding columns of $G_\pm(x;\lambda)$, $G_-(x;\lambda)S(x;\lambda)$, and $F(x;\lambda)$. See Proposition 9, Corollary 10, and Remark 16 below.*

The following lemma relies on the positivity result of Proposition 8 and states solvability of the integral equation (4.130) in $L^2_z(\mathbb{R})$. For simplicity of notations, we drop dependence of S, F and $G_\pm$ from the variable x.

Lemma 22. *For every $r(\lambda) \in L^2_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$ satisfying (4.100) and every $F(\lambda) \in L^2_z(\mathbb{R})$, there is a unique solution $G(\lambda) \in L^2_z(\mathbb{R})$ of the linear inhomogeneous equation*

$$(I - \mathcal{P}^-_S)G(\lambda) = F(\lambda), \quad \lambda \in \mathbb{R} \cup i\mathbb{R}, \tag{4.132}$$

where $\mathcal{P}^-_S G := \mathcal{P}^-(GS)$.

Proof. The operator $I - \mathcal{P}^-_S$ is known to be a Fredholm operator of the index zero [6, 7, 122]. By Fredholm's alternative, a unique solution to the linear integral equation (4.132) exists for $G(\lambda) \in L^2_z(\mathbb{R})$ if and only if the zero solution to the homogeneous equation $(I - \mathcal{P}^-_S)g = 0$ is unique in $L^2_z(\mathbb{R})$.

Suppose that there exists nonzero $g \in L^2_z(\mathbb{R})$ such that $(I - \mathcal{P}^-_S)g = 0$. Since $S(\lambda) \in L^2_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$, we define two analytic functions in $\mathbb{C} \setminus \mathbb{R}$ by

$$g_1(z) := \mathcal{C}(gS)(z) \quad \text{and} \quad g_2(z) := \mathcal{C}(gS)^*(z),$$

where the asterisk denotes Hermite conjugate. We multiply the two functions by each other and integrate along the semi-circle of radius R centered at zero in $\mathbb{C}^+$. Because g_1 and g_2 are analytic functions in $\mathbb{C}^+$, the Cauchy–Goursat theorem implies that

$$0 = \oint g_1(z)g_2(z)dz.$$

Because $g(\lambda), S(\lambda) \in L^2_z(\mathbb{R})$, we have $g(\lambda)S(\lambda) \in L^1_z(\mathbb{R})$, so that the asymptotic limit (4.124) in Proposition 7 implies that $g_{1,2}(z) = \mathcal{O}(z^{-1})$ as $|z| \to \infty$. Therefore, the integral on arc goes to zero as $R \to \infty$, so that we obtain

$$\begin{aligned} 0 &= \int_{\mathbb{R}} g_1(z) g_2(z) dz \\ &= \int_{\mathbb{R}} \mathcal{P}^+(gS)\, [\mathcal{P}^-(gS)]^* dz \\ &= \int_{\mathbb{R}} \left[\mathcal{P}^-(gS) + gS\right] [\mathcal{P}^-(gS)]^* dz, \end{aligned}$$

where we have used the identity $\mathcal{P}^+ - \mathcal{P}^- = I$ following from relations (4.125). Since $\mathcal{P}^-(gS) = g$, we finally obtain

$$0 = \int_{\mathbb{R}} g(I + S) g^* dz. \tag{4.133}$$

By bound (4.126) in Proposition 8, the real part of the quadratic form associated with the matrix $I + S$ is strictly positive definite for every $z \in \mathbb{R}$. Therefore, equation (4.133) implies that $g = 0$ is the only solution to the homogeneous equation $(I - \mathcal{P}^-_S)g = 0$ in $L^2_z(\mathbb{R})$. □

As a consequence of Lemma 22, we obtain solvability of the two related Riemann–Hilbert problems (4.118).

Corollary 9. *Let $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied. There exists a unique solution to the Riemann–Hilbert problems (4.118) for every $x \in \mathbb{R}$ such that the functions*

$$G_{\pm 1,2}(x;\lambda) := M_\pm(x;z)\tau_{1,2}(\lambda) - \tau_{1,2}(\lambda)$$

are analytic functions of z in $\mathbb{C}^\pm$ and $G_{\pm 1,2}(x;\lambda) \in L^2_z(\mathbb{R})$.

Proof. For every $x \in \mathbb{R}$, the two related Riemann–Hilbert problems (4.118) are rewritten for $G_{\pm 1,2}$ and $F_{1,2}$ given by (4.119) in the form (4.128). By Proposition 5, we have $S(x;\lambda) \in L^1_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$ and $F_{1,2}(x;\lambda) \in L^2_z(\mathbb{R})$, hence $\mathcal{P}^-(F_{1,2}) \in L^2_z(\mathbb{R})$. By Lemma 22, equation (4.130) admits a unique solution for $G_{-1,2}(x;\lambda) \in L^2_z(\mathbb{R})$ for every $x \in \mathbb{R}$. Then, we define a unique solution for $G_{+1,2}(x;\lambda) \in L^2_z(\mathbb{R})$ by equation (4.131). Analytic extensions of $G_{\pm 1,2}(x;\lambda)$ as functions of z in $\mathbb{C}^\pm$ are defined by the Cauchy integrals (4.129). These functions solve the Riemann–Hilbert problem (4.128) by Proposition 7 with $p = 2$. □

For further estimates, we modify the method of Lemma 22 and prove that the operator $(I - \mathcal{P}^-_S)^{-1}$ in the integral Fredholm equation (4.132) is invertible with a bounded inverse in space $L^2_z(\mathbb{R})$.

Lemma 23. *For every $r(\lambda) \in L^2_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$ satisfying (4.100), the inverse operator $(I - \mathcal{P}^-_S)^{-1}$ is a bounded operator from $L^2_z(\mathbb{R})$ to $L^2_z(\mathbb{R})$. In particular,* *there is a positive constant C that only depends on $\|r(\lambda)\|_{L^\infty_z}$ such that for every row-vector $f \in L^2_z(\mathbb{R})$, we have*

$$\|(I - \mathcal{P}^-_S)^{-1} f\|_{L^2_z} \leq C\|f\|_{L^2_z}. \tag{4.134}$$

Proof. We consider the linear inhomogeneous equation (4.132) with $F \in L^2_z(\mathbb{R})$. Recalling that $\mathcal{P}^+ - \mathcal{P}^- = I$, we write $G = G_+ - G_-$, where G_+ and G_- satisfy the inhomogeneous equations

$$G_- - \mathcal{P}^-(G_-S) = \mathcal{P}^-(F), \quad G_+ - \mathcal{P}^-(G_+S) = \mathcal{P}^+(F). \tag{4.135}$$

By Lemma 22, since $\mathcal{P}^\pm(F) \in L^2_z(\mathbb{R})$, there are unique solutions to the inhomogeneous equations (4.132) and (4.135), so that the decomposition $G = G_+ - G_-$ is unique. Therefore, we only need to find the estimates of G_+ and G_- in $L^2_z(\mathbb{R})$.

To deal with G_-, we define two analytic functions in $\mathbb{C} \setminus \mathbb{R}$ by

$$g_1(z) := \mathcal{C}(G_-S)(z) \quad \text{and} \quad g_2(z) := \mathcal{C}(G_-S + F)^*(z),$$

similarly to the proof of Lemma 22. By Proposition 7, $g_1(z) = \mathcal{O}(z^{-1})$ and $g_2(z) \to 0$ as $|z| \to \infty$, since $F \in L^2_z(\mathbb{R})$, $G_- \in L^2_z(\mathbb{R})$, and $S(\lambda) \in L^2_z(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$. Therefore, the integral on the semi-circle of radius $R > 0$ in the upper half-plane still goes to zero as $R \to \infty$ by Lebesgue's dominated convergence theorem. Performing the same manipulations as in the proof of Lemma 22, we obtain

$$\begin{aligned} 0 &= \oint g_1(z)g_2(z)dz \\ &= \int_{\mathbb{R}} \mathcal{P}^+(G_-S)\left[\mathcal{P}^-(G_-S + F)\right]^* dz \\ &= \int_{\mathbb{R}} \left[\mathcal{P}^-(G_-S) + G_-S\right]\left[\mathcal{P}^-(G_-S + F)\right]^* dz \\ &= \int_{\mathbb{R}} \left[G_- - \mathcal{P}^-(F) + G_-S\right] G_-^* dz, \end{aligned}$$

where we have used the first inhomogeneous equation in system (4.135). By the bound (4.126) in Proposition 8, there is a positive constant C_- such that

$$C_-\|G_-\|^2_{L^2} \leq \text{Re} \int_{\mathbb{R}} G_-(I + S)G_-^* dz = \text{Re} \int_{\mathbb{R}} \mathcal{P}^-(F)G_-^* dz \leq \|F\|_{L^2}\|G_-\|_{L^2},$$

where we have used the Cauchy–Schwarz inequality and bound (4.123) with $C_{p=2} = 1$. Note that the above estimate holds independently for the corresponding row-vectors of the matrices G_- and F. Since $G_- = (I - \mathcal{P}^-_S)^{-1}\mathcal{P}^- F$, for every row-vector $f \in L^2_z(\mathbb{R})$ of the matrix $F \in L^2_z(\mathbb{R})$, the above inequality yields

$$\|(I - \mathcal{P}^-_S)^{-1}\mathcal{P}^- f\|_{L^2_z} \leq C_-^{-1}\|f\|_{L^2_z}. \tag{4.136}$$

To deal with G_+, we use $\mathcal{P}^+ - \mathcal{P}^- = I$ and rewrite the second inhomogeneous equation in system (4.135) as follows:

$$G_+(I+S) - \mathcal{P}^+(G_+S) = \mathcal{P}^+(F). \tag{4.137}$$

We now define two analytic functions in $\mathbb{C} \setminus \mathbb{R}$ by

$$g_1(z) := \mathcal{C}(G_+S)(z) \quad \text{and} \quad g_2(z) := \mathcal{C}(G_+S+F)^*(z)$$

and integrate the product of g_1 and g_2 on the semi-circle of radius $R > 0$ in the lower half-plane. Performing the same manipulations as above, we obtain

$$\begin{aligned} 0 &= \oint g_1(z)g_2(z)dz \\ &= \int_{\mathbb{R}} \mathcal{P}^-(G_+S)\left[\mathcal{P}^+(G_+S+F)\right]^* dz \\ &= \int_{\mathbb{R}} \left[G_+ - \mathcal{P}^+(F)\right]\left[G_+(I+S)\right]^* dz, \end{aligned}$$

where we have used equation (4.137).

By the bounds (4.126) and (4.127) in Proposition 8, there are positive constants C_+ and

$$\begin{aligned} C_-\|G_+\|^2_{L^2} \le \text{Re} \int_{\mathbb{R}} G_+(I+S)^* G_+^* dz = \text{Re} \int_{\mathbb{R}} \mathcal{P}^+(F)(I+S)^* G_+^* dz \\ \le C_+\|F\|_{L^2}\|G_+\|_{L^2}, \end{aligned}$$

where we have used the Cauchy–Schwarz inequality and bound (4.123) with $C_{p=2} = 1$. Again, the above estimate holds independently for the corresponding row-vectors of the matrices G_+ and F. Since $G_+ = (I - \mathcal{P}_S^-)^{-1}\mathcal{P}^+F$, for every row-vector $f \in L^2_z(\mathbb{R})$ of the matrix $F \in L^2_z(\mathbb{R})$, the above inequality yields

$$\|(I - \mathcal{P}_S^-)^{-1}\mathcal{P}^+ f\|_{L^2_z} \le C_-^{-1}C_+\|f\|_{L^2_z}. \tag{4.138}$$

The assertion of the lemma is proved with bounds (4.136), (4.138), and the triangle inequality. □

4.4.2 Estimates on solutions to the Riemann-Hilbert problem

Using Corollary 9, we obtain solvability of the Riemann–Hilbert problem (4.116). Indeed, the abstract Riemann–Hilbert problem (4.128) is derived for two versions of $G_\pm$ and $F_\pm$ given by (4.119). For the first version, we have

$$G_{\pm 1}(x;\lambda) := M_\pm(x;z)\tau_1(\lambda) - \tau_1(\lambda) = \left[\mu_\pm(x;z) - e_1, 2i\lambda\left(\eta_\pm(x;z) - e_2\right)\right] \tag{4.139}$$

and

$$F_1(x;\lambda) := \tau_1(\lambda)S(x;\lambda) = R(x;z)\tau_1(\lambda). \tag{4.140}$$

By Corollary 9, there is a solution $G_{\pm 1}(x;\lambda) \in L^2_z(\mathbb{R})$ of the integral Fredholm equations

$$G_{\pm 1}(x;\lambda) = \mathcal{P}^{\pm}\left(G_{-1}(x;\lambda)S(x;\lambda) + F_1(x;\lambda)\right)(z), \quad z \in \mathbb{R}. \tag{4.141}$$

Using equation (4.141) for the first column of $G_{\pm}$, we obtain

$$\mu_{\pm}(x;z) - e_1 = \mathcal{P}^{\pm}\left(M_{-}(x;\cdot)R(x;\cdot)\right)^{(1)}(z), \quad z \in \mathbb{R}, \tag{4.142}$$

where we have used the following identities:

$$(G_{-1}S + F_1)^{(1)} = (M_{-}\tau_1 S)^{(1)} = (M_{-}R\tau_1)^{(1)} = (M_{-}R)^{(1)}.$$

For the second version of the abstract Riemann–Hilbert problem (4.128), we have

$$G_{\pm 2}(x;\lambda) := M_{\pm}(x;z)\tau_2(\lambda) - \tau_2(\lambda) = \left[(2i\lambda)^{-1}\left(\mu_{\pm}(x;z) - e_1\right), \eta_{\pm}(x;z) - e_2\right] \tag{4.143}$$

and

$$F_2(x;\lambda) := \tau_2(\lambda)S(x;\lambda) = R(x;z)\tau_2(\lambda). \tag{4.144}$$

Again by Corollary 9, there is a solution $G_{\pm 2}(x;\lambda) \in L^2_z(\mathbb{R})$ of the integral Fredholm equations (4.141), where $G_{\pm 1}$ and F_1 are replaced by $G_{\pm 2}$ and F_2. Using equation (4.141) for the second column of $G_{\pm 2}$, we obtain

$$\eta_{\pm}(x;z) - e_2 = \mathcal{P}^{\pm}\left(M_{-}(x;\cdot)R(x;\cdot)\right)^{(2)}(z), \quad z \in \mathbb{R}. \tag{4.145}$$

where we have used the following identities:

$$(G_{-2}S + F_2)^{(2)} = (M_{-}\tau_2 S)^{(2)} = (M_{-}R\tau_2)^{(2)} = (M_{-}R)^{(2)}.$$

Equations (4.142) and (4.145) can be written in the form

$$M_{\pm}(x;z) = I + \mathcal{P}^{\pm}\left(M_{-}(x;\cdot)R(x;\cdot)\right)(z), \quad z \in \mathbb{R}, \tag{4.146}$$

which represents the solution to the Riemann–Hilbert problem (4.116) on the real line. The analytic continuation of functions $M_{\pm}(x;\cdot)$ in $\mathbb{C}^{\pm}$ is given by the Cauchy operators

$$M_{\pm}(x;z) = I + \mathcal{C}\left(M_{-}(x;\cdot)R(x;\cdot)\right)(z), \quad z \in \mathbb{C}^{\pm}. \tag{4.147}$$

The corresponding result on solvability of the integral equations (4.146) is given by the following lemma.

Lemma 24. *Let $r_{\pm} \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied. There is a positive constant C that only depends on $\|r_{\pm}\|_{L^\infty}$ such that the unique solution to the integral equations (4.146) enjoys the estimate for every $x \in \mathbb{R}$,*

$$\|M_{\pm}(x;\cdot) - I\|_{L^2} \leq C\left(\|r_+\|_{L^2} + \|r_-\|_{L^2}\right). \tag{4.148}$$

Proof. By Proposition 5, if $r_{\pm} \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$, then $r(\lambda) \in L^2(\mathbb{R}) \cap L^\infty_z(\mathbb{R})$.

Under these conditions, it follows from the explicit expressions (4.140) and (4.144) that $R(x;z)\tau_{1,2}(\lambda)$ belong to $L^2_z(\mathbb{R})$ for every $x\in\mathbb{R}$ and there is a positive constant C that only depends on $\|r_\pm\|_{L^\infty(\mathbb{R})}$ such that for every $x\in\mathbb{R}$,

$$\|R(x;z)\tau_{1,2}(\lambda)\|_{L^2_z} \leq C\left(\|r_+\|_{L^2} + \|r_-\|_{L^2}\right). \tag{4.149}$$

By derivation above, the integral equation (4.146) for the projection operator $\mathcal{P}^-$ is obtained from two versions of the integral equation (4.132) corresponding to $F_{1,2}(x;\lambda) := \mathcal{P}^-\left(R(x;z)\tau_{1,2}(\lambda)\right)(z)$. Therefore, each element of $M_-(x;z)$ enjoys the bound (4.134) for the corresponding row vectors of the two versions of $F_{1,2}(x;z)$. Combining the estimates (4.134) and (4.149), we obtain the bound (4.148). □

Before we continue, let us discuss the redundancy between solutions to the two versions of the Riemann–Hilbert problems (4.118). By using equation (4.141) for the second column of $G_{\pm 1}$, we obtain

$$2i\lambda\left(\eta_\pm(x;z) - e_2\right) = \mathcal{P}^\pm\left(2i\lambda\left(M_-(x;\cdot)R(x;\cdot)\right)^{(2)}\right)(z), \quad z\in\mathbb{R}. \tag{4.150}$$

By using equation (4.141) for the first column of $G_{\pm 2}$, we obtain

$$(2i\lambda)^{-1}\left(\mu_\pm(x;z) - e_1\right) = \mathcal{P}^\pm\left((2i\lambda)^{-1}\left(M_-(x;\cdot)R(x;\cdot)\right)^{(1)}\right)(z), \quad z\in\mathbb{R}. \tag{4.151}$$

Unless equations (4.150) and (4.151) are redundant in view of equations (4.142) and (4.145), the two versions of the Riemann–Hilbert problems (4.130) may seem to be inconsistent. In order to show the redundancy explicitly, we use the following result.

Proposition 9. *Let $f(\lambda)\in L^1_z(\mathbb{R})\cap L^\infty_z(\mathbb{R})$ be even in λ for all $\lambda\in\mathbb{R}\cup i\mathbb{R}$. Then*

$$\mathcal{P}^\pm_{\text{even}}\left(\lambda f(\lambda)\right)(\lambda) = \lambda\mathcal{P}^\pm_{\text{even}}(f)(\lambda), \quad \lambda\in\mathbb{R}\cup i\mathbb{R}, \tag{4.152}$$

where

$$\mathcal{P}^\pm_{\text{even}}(f)(\lambda) := \left(\int_0^{+\infty} + \int_{+i\infty}^{i0} + \int_0^{-\infty} + \int_{-i\infty}^{i0}\right)\frac{f(\lambda')d\lambda'}{\lambda' - (\lambda\pm i0)} \equiv \mathcal{P}^\pm(f(\lambda))(\lambda^2). \tag{4.153}$$

Similarly, let $g(\lambda)\in L^1_z(\mathbb{R})\cap L^2_z(\mathbb{R})$ be odd in λ for all $\lambda\in\mathbb{R}\cup i\mathbb{R}$. Then

$$\mathcal{P}^\pm_{\text{odd}}\left(\lambda g(\lambda)\right)(\lambda) = \lambda\mathcal{P}^\pm_{\text{odd}}(g)(\lambda), \quad \lambda\in\mathbb{R}\cup i\mathbb{R}, \tag{4.154}$$

where

$$\mathcal{P}^\pm_{\text{odd}}(g)(\lambda) := \left(\int_0^{+\infty} + \int_{+i\infty}^{i0} + \int_{-\infty}^{0} + \int_{i0}^{-i\infty}\right)\frac{g(\lambda')d\lambda'}{\lambda' - (\lambda\pm i0)} \equiv \mathcal{P}^\pm(g(\lambda))(\lambda^2). \tag{4.155}$$

Proof. First, we note the validity of the definition (4.153) if $f(-\lambda) = f(\lambda)$:

$$\begin{aligned}
\mathcal{P}^{\pm}(f(\lambda))(\lambda^2) &= \int_{-\infty}^{\infty} \frac{f(\lambda')2\lambda' d\lambda'}{(\lambda')^2 - (\lambda^2 \pm i0)} \\
&= \left(\int_0^{+\infty} + \int_{+i\infty}^{i0} \right) f(\lambda') \left[\frac{1}{\lambda' - (\lambda \pm i0)} + \frac{1}{\lambda' + (\lambda \pm i0)} \right] d\lambda' \\
&=: \mathcal{P}^{\pm}_{\text{even}}(f)(\lambda).
\end{aligned}$$

Then, relation (4.152) is established from the trivial result

$$\left(\int_0^{+\infty} + \int_{+i\infty}^{i0} + \int_0^{-\infty} + \int_{-i\infty}^{i0} \right) f(\lambda') d\lambda' = 0,$$

which is justified if $f(\lambda) \in L^1_z(\mathbb{R})$ and even in λ. The relation (4.154) is proved similarly, thanks to the changes in the definition (4.155). □

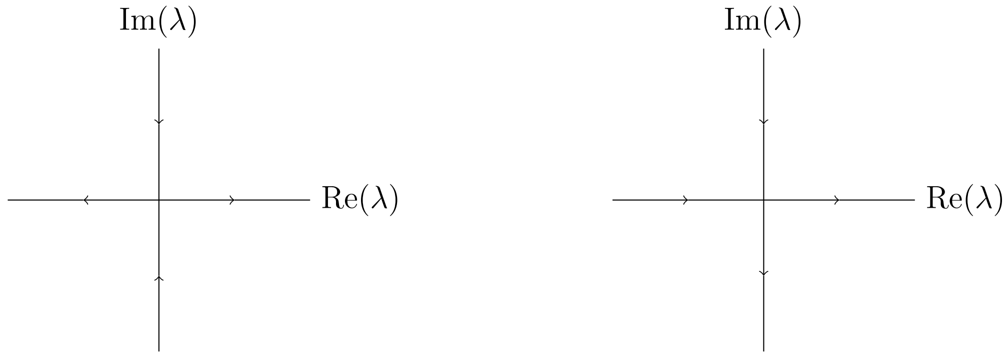


Figure 4.3: The left and right panels show the direction of contours used for $\mathcal{P}^{\pm}_{\text{even}}$ and $\mathcal{P}^{\pm}_{\text{odd}}$

Figure 4.3 shows the contours of integration used in the definitions of $\mathcal{P}^{\pm}_{\text{even}}$ and $\mathcal{P}^{\pm}_{\text{odd}}$ in (4.153) and (4.155). The following corollary of Proposition 9 specifies the redundancy between the two different versions of the Riemann–Hilbert problems (4.118).

Corollary 10. *Consider two unique solutions to the Riemann–Hilbert problems (4.118) in Corollary 9. Then, for every $x \in \mathbb{R}$, we have*

$$G_{\pm 1}(x;\lambda) = 2i\lambda \text{sign}(\lambda) G_{\pm 2}(x;\lambda), \quad \lambda \in \mathbb{R} \cup i\mathbb{R}, \tag{4.156}$$

where the sign function returns the sign of either real or imaginary part of λ.

Proof. We note the relation $\tau_2^{-1}(\lambda)\tau_1(\lambda) = 2i\lambda I$, where I is the identity 2-by-2 matrix. From here, the relation (4.156) follows for $\lambda \in \mathbb{R}^+ \cup i\mathbb{R}^+$. To consider the continuation of this relation to $\lambda \in \mathbb{R}^- \cup i\mathbb{R}^-$, we apply Proposition 9 with the explicit parametrization of the contours of integrations as on Figure 4.3. We choose the even function f and the odd function g in the form

$$f(\lambda) := \left(M_-(x;\lambda^2) R(x;\lambda^2) \right)^{(2)}, \quad g(\lambda) := (2i\lambda)^{-1} \left(M_-(x;\lambda^2) R(x;\lambda^2) \right)^{(1)}.$$

Then, equation (4.150) follows from equation (4.145), thanks to the relation (4.152), whereas equation (4.142) follows from equation (4.151) thanks to the relation (4.154). Thus, the relation (4.156) is verified for every $\lambda \in \mathbb{R} \cup i\mathbb{R}$. To ensure that the integrations (4.153) and (4.155) returns $\mathcal{P}^{\pm}$ for $\lambda f(\lambda)$ and $\lambda g(\lambda)$, the sign function is used in the relation (4.156). □

Remark 16. *Corollary 10 shows that the complex integration in the z plane in the integral equations (4.141) has to be extended in two different ways in the λ plane. For the first vector columns of the integral equation (4.141), we have to use the definition (4.155) for odd functions in λ, whereas for the second vector columns of the integral equation (4.141), we have to use the definition (4.153) for even functions in λ.*

Next, we shall obtain refined estimates on the solution to the integral equations (4.146). We start with estimates on the scattering coefficients r_+ and r_- obtained with the Fourier theory.

Proposition 10. *For every $x_0 \in \mathbb{R}^+$ and every $r_\pm \in H^1(\mathbb{R})$, we have*

$$\sup_{x\in(x_0,\infty)} \left\| \langle x \rangle \mathcal{P}^+ \left(\bar{r}_+(z) e^{-2izx} \right) \right\|_{L^2_z} \leq \|r_+\|_{H^1} \tag{4.157}$$

and

$$\sup_{x\in(x_0,\infty)} \left\| \langle x \rangle \mathcal{P}^- \left(r_-(z) e^{2izx} \right) \right\|_{L^2_z} \leq \|r_-\|_{H^1}, \tag{4.158}$$

where $\langle x \rangle := (1+x^2)^{1/2}$. In addition, if $r_\pm \in H^1(\mathbb{R})$, then

$$\sup_{x\in\mathbb{R}} \left\| \mathcal{P}^+ \left(\bar{r}_+(z) e^{-2izx} \right) \right\|_{L^\infty_z} \leq \frac{1}{\sqrt{2}} \|r_+\|_{H^1} \tag{4.159}$$

and

$$\sup_{x\in\mathbb{R}} \left\| \mathcal{P}^- \left(r_-(z) e^{2izx} \right) \right\|_{L^\infty_z} \leq \frac{1}{\sqrt{2}} \|r_-\|_{H^1}. \tag{4.160}$$

Furthermore, if $r_\pm \in L^{2,1}(\mathbb{R})$, then

$$\sup_{x\in\mathbb{R}} \left\| \mathcal{P}^+ \left(z\bar{r}_+(z) e^{-2izx} \right) \right\|_{L^2_z} \leq \|z r_+(z)\|_{L^2_z}, \tag{4.161}$$

and

$$\sup_{x\in\mathbb{R}} \left\| \mathcal{P}^- \left(z r_-(z) e^{2izx} \right) \right\|_{L^2_z} \leq \|z r_-(z)\|_{L^2_z}. \tag{4.162}$$

Proof. Recall the following elementary result from the Fourier theory. For a given function $r \in L^2(\mathbb{R})$, we use the Fourier transform $\widehat{r} \in L^2(\mathbb{R})$ with the definition $\widehat{r}(k) := \frac{1}{2\pi} \int_{\mathbb{R}} r(z) e^{-ikz} dz$, so that

$$\|r\|^2_{L^2} = 2\pi \|\hat{r}\|^2_{L^2}.$$

Then, we have $r \in H^1(\mathbb{R})$ if and only if $\widehat{r} \in L^{2,1}(\mathbb{R})$. Similarly, $r \in L^{2,1}(\mathbb{R})$ if and only if $\widehat{r} \in H^1(\mathbb{R})$.

In order to prove (4.157), we write explicitly

$$\begin{aligned}\mathcal{P}^+\left(\bar{r}_+(z)e^{-2izx}\right)(z) &= \frac{1}{2\pi i}\lim_{\epsilon\downarrow 0}\int_{\mathbb{R}}\frac{\overline{r}_+(s)e^{-2isx}}{s-(z+i\epsilon)}ds \\ &= \frac{1}{2\pi i}\int_{\mathbb{R}}\widehat{\overline{r}_+}(k)\left(\lim_{\epsilon\downarrow 0}\int_{\mathbb{R}}\frac{e^{i(k-2x)s}}{s-(z+i\epsilon)}ds\right)dk \\ &= \int_{2x}^{\infty}\widehat{\overline{r}_+}(k)e^{i(k-2x)z}dk, \qquad (4.163)\end{aligned}$$

where the following residue computation has been used:

$$\lim_{\epsilon\downarrow 0}\frac{1}{2\pi i}\int_{\mathbb{R}}\frac{e^{is(k-2x)}}{s-i\epsilon}ds = \lim_{\epsilon\downarrow 0}\begin{cases} e^{-\epsilon(k-2x)}, & \text{if } k-2x>0 \\ 0, & \text{if } k-2x<0\end{cases} = \chi(k-2x), \quad (4.164)$$

with χ being the characteristic function. The bound (4.157) is obtained from the bound (4.41) of Proposition 4 for every $x_0 \in \mathbb{R}^+$:

$$\sup_{x\in(x_0,\infty)}\left\|\langle x\rangle\int_{2x}^{\infty}\widehat{\overline{r}_+}(k)e^{i(k-2x)z}dk\right\|_{L^2_z} \leq \sqrt{2\pi}\|\widehat{r}_+\|_{L^{2,1}} = \|r_+\|_{H^1}.$$

Similarly, we use the representation (4.163) and obtain bound (4.159) for every $x \in \mathbb{R}$:

$$\|\mathcal{P}^+\left(\bar{r}_+(z)e^{-2izx}\right)(z)\|_{L^\infty_z} \leq \|\widehat{\overline{r}_+}(k)\|_{L^1_k} \leq \sqrt{\pi}\|\widehat{\overline{r}_+}(k)\|_{L^{2,1}_k} \leq \frac{1}{\sqrt{2}}\|r_+\|_{H^1}. \quad (4.165)$$

The bounds (4.158) and (4.160) are obtained similarly from the representation

$$\mathcal{P}^-\left(r_-(z)e^{2izx}\right)(z) = \frac{1}{2\pi i}\lim_{\epsilon\downarrow 0}\int_{\mathbb{R}}\frac{r_-(s)e^{2isx}}{s-(z-i\epsilon)}ds = -\int_{-\infty}^{-2x}\widehat{r_-}(k)e^{i(k+2x)z}dk.$$

The bounds (4.161) and (4.162) follow from the bound (4.123) with $C_{p=2}=1$ of Proposition 7. □

We shall use the estimates of Lemma 24 and Proposition 10 to derive useful estimates on the solutions to the Riemann–Hilbert problem (4.116). By Lemma 24, these solutions on the real line can be written in the integral Fredholm form (4.146). We only need to obtain estimates on the vector columns $\mu_- - e_1$ and $\eta_+ - e_2$. From equation (4.142), we obtain

$$\mu_-(x;z) - e_1 = \mathcal{P}^-\left(r_-(z)e^{2izx}\eta_+(x;z)\right)(z), \quad z\in\mathbb{R}, \qquad (4.166)$$

where we have used the following identities

$$(M_-R)^{(1)} = \Phi_\infty^{-1}(P_-R)^{(1)} = r_-(z)e^{2izx}\Phi_\infty^{-1}p_+ = r_-(z)e^{2izx}M_+^{(2)},$$

which follow from the representations (4.107), (4.112), and (4.115), as well as the scattering relation (4.102). From equation (4.145), we obtain

$$\eta_+(x;z) - e_2 = \mathcal{P}^+\left(\bar{r}_+(z)e^{-2izx}\mu_-(x;z)\right)(z), \quad z \in \mathbb{R}, \tag{4.167}$$

where we have used the following identities

$$(M_-R)^{(2)} = \Phi_\infty^{-1}(P_-R)^{(2)} = \bar{r}_+(z)e^{-2izx}\Phi_\infty^{-1}m_+ = \bar{r}_+(z)e^{-2izx}M_-^{(1)},$$

which also follow from the representations (4.107), (4.112), and (4.115).

Let us introduce the 2-by-2 matrix

$$M(x;z) = [\mu_-(x;z) - e_1, \quad \eta_+(x;z) - e_2] \tag{4.168}$$

and write the system of integral equations (4.166) and (4.167) in the matrix form

$$M - \mathcal{P}^+(MR_+) - \mathcal{P}^-(MR_-) = F, \tag{4.169}$$

where

$$R_+(x;z) = \begin{bmatrix} 0 & \bar{r}_+(z)e^{-2izx} \\ 0 & 0 \end{bmatrix}, \quad R_-(x;z) = \begin{bmatrix} 0 & 0 \\ r_-(z)e^{2izx} & 0 \end{bmatrix} \tag{4.170}$$

and

$$F(x;z) := \left[e_2\mathcal{P}^-(r_-(z)e^{2izx}), \quad e_1\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})\right]. \tag{4.171}$$

The inhomogeneous term F given by (4.171) isestimated by Proposition 10. The following lemma estimates solutions to the system of integral equations (4.169).

Lemma 25. *For every $x_0 \in \mathbb{R}^+$ and every $r_\pm \in H^1(\mathbb{R})$, the unique solution to the system of integral equations (4.166) and (4.167) satisfies the estimates*

$$\sup_{x\in(x_0,\infty)} \left\| \langle x\rangle \mu_-^{(2)}(x;z)\right\|_{L_z^2} \leq C\|r_-\|_{H^1} \tag{4.172}$$

and

$$\sup_{x\in(x_0,\infty)} \left\| \langle x\rangle \eta_+^{(1)}(x;z)\right\|_{L_z^2} \leq C\|r_+\|_{H^1}, \tag{4.173}$$

where C is a positive constant that depends on $\|r_\pm\|_{L^\infty}$. Moreover, if $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$, then

$$\sup_{x\in\mathbb{R}} \left\| \partial_x \mu_-^{(2)}(x;z)\right\|_{L_z^2} \leq C\left(\|r_+\|_{H^1\cap L^{2,1}} + \|r_-\|_{H^1\cap L^{2,1}}\right) \tag{4.174}$$

and

$$\sup_{x\in\mathbb{R}} \left\| \partial_x \eta_+^{(1)}(x;z)\right\|_{L_z^2} \leq C\left(\|r_+\|_{H^1\cap L^{2,1}} + \|r_-\|_{H^1\cap L^{2,1}}\right) \tag{4.175}$$

where C is another positive constant that depends on $\|r_\pm\|_{L^\infty}$.

Proof. Using the identity $\mathcal{P}^+ - \mathcal{P}^- = I$ following from relations (4.125) and the

identity
$$R_+ + R_- = (I - R_+)R,$$
which follows from the explicit form (4.113), we rewrite the inhomogeneous equation (4.169) in the matrix form
$$G - \mathcal{P}^-(GR) = F, \tag{4.176}$$
where $G := M(I - R_+)$ is given explicitly from (4.168) and (4.170) by
$$G(x;z) = \begin{bmatrix} \mu_-^{(1)}(x;z) - 1 & \eta_+^{(1)}(x;z) - \bar{r}_+(z)e^{-2izx}(\mu_-^{(1)}(x;z) - 1) \\ \mu_-^{(2)}(x;z) & \eta_+^{(2)}(x;z) - 1 - \bar{r}_+(z)e^{-2izx}\mu_-^{(2)}(x;z) \end{bmatrix}. \tag{4.177}$$

From the explicit expression (4.171) for $F(x;z)$, we can see that the second row vector of $F(x;z)$ and $F(x;z)\tau_1(\lambda)$ remains the same and is given by $[\mathcal{P}^-(r_-(z)e^{2izx}), 0]$. From the explicit expressions (4.177), the second row vector of $G(x;z)\tau_1(\lambda)$ is given by
$$\left[\mu_-^{(2)}(x;z),\ \ 2i\lambda\left(\eta_+^{(2)}(x;z) - 1 - \bar{r}_+(z)e^{-2izx}\mu_-^{(2)}(x;z)\right)\right]$$
Using bound (4.134) for the second row vector of $G(x;z)\tau_1(\lambda)$, we obtain the following bounds for every $x \in \mathbb{R}$,
$$\|\mu_-^{(2)}(x;z)\|_{L^2_z} \leq C\|\mathcal{P}^-(r_-(z)e^{2izx})\|_{L^2_z} \tag{4.178}$$
and
$$\|2i\lambda\left(\eta_+^{(2)}(x;z) - 1 - \bar{r}_+(z)e^{-2izx}\mu_-^{(2)}(x;z)\right)\|_{L^2_z} \leq C\|\mathcal{P}^-(r_-(z)e^{2izx})\|_{L^2_z}, \tag{4.179}$$
where the positive constant C only depends on $\|r_\pm\|_{L^\infty}$. By substituting bound (4.158) of Proposition 10 into (4.178), we obtain bound (4.172). Also note that since $|2i\lambda\bar{r}_+(z)| = |r(\lambda)|$ and $r(\lambda) \in L^\infty_z(\mathbb{R})$, we also obtain from (4.178) and (4.179) by the triangle inequality,
$$\|2i\lambda\left(\eta_+^{(2)}(x;z) - 1\right)\|_{L^2_z} \leq C\|\mathcal{P}^-(r_-(z)e^{2izx})\|_{L^2_z}, \tag{4.180}$$
where the positive constant C still depends on $\|r_\pm\|_{L^\infty}$ only.

Similarly, from the explicit expression (4.171) for $F(x;z)$, we can see that the first row vector of $F(x;z)$ and $F(x;z)\tau_2(\lambda)$ remains the same and is given by $[0, \mathcal{P}^+(\bar{r}_+(z)e^{-2izx})]$. From the explicit expressions (4.177), the first row vector of $G(x;z)\tau_2(\lambda)$ is given by
$$\left[(2i\lambda)^{-1}(\mu_-^{(1)}(x;z) - 1),\ \ \eta_+^{(1)}(x;z) - \bar{r}_+(z)e^{-2izx}(\mu_-^{(1)}(x;z) - 1)\right]$$
Using bound (4.134) for the first row vector of $G(x;z)\tau_2(\lambda)$, we obtain the following

bounds for every $x \in \mathbb{R}$,

$$\|(2i\lambda)^{-1}(\mu_-^{(1)}(x;z)-1)\|_{L^2_z} \leq C\|\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})\|_{L^2_z} \tag{4.181}$$

and

$$\|\eta_+^{(1)}(x;z)-\bar{r}_+(z)e^{-2izx}(\mu_-^{(1)}(x;z)-1)\|_{L^2_z} \leq C\|\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})\|_{L^2_z}, \tag{4.182}$$

where the positive constant C only depends on $\|r_\pm\|_{L^\infty}$. Since $|2i\lambda\bar{r}_+(z)| = |r(\lambda)|$ and $r(\lambda) \in L^\infty_z(\mathbb{R})$, we also obtain from (4.181) and (4.182) by the triangle inequality,

$$\begin{aligned}\|\eta_+^{(1)}(x;z)\|_{L^2_z} &\leq \|(2i\lambda)\bar{r}_+(z)e^{-2izx}(2i\lambda)^{-1}(\mu_-^{(1)}(x;z)-1)\|_{L^2_z} \\ &\quad +C\|\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})\|_{L^2_z} \\ &\leq C'\|\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})\|_{L^2_z},\end{aligned} \tag{4.183}$$

where the positive constant C' still depends on $\|r_\pm\|_{L^\infty}$ only. By substituting bound (4.157) of Proposition 10 into (4.183), we obtain bound (4.173).

In order to obtain bounds (4.174) and (4.175), we take derivative of the inhomogeneous equation (4.169) in x and obtain

$$\partial_x M - \mathcal{P}^+\left(\partial_x M\right)R_+ - \mathcal{P}^-\left(\partial_x M\right)R_- = \tilde{F}, \tag{4.184}$$

where

$$\begin{aligned}\tilde{F} &:= \partial_x F + \mathcal{P}^+ M\partial_x R_+ + \mathcal{P}^- M\partial_x R_- \\ &= 2i\left[e_2\mathcal{P}^-(zr_-(z)e^{2izx}), \quad e_1\mathcal{P}^+(-z\bar{r}_+(z)e^{-2izx})\right] \\ &\quad +2i\left[\begin{array}{cc} zr_-(z)\eta_+^{(1)}(x;z)e^{2izx} & -z\bar{r}_+(z)(\mu_-^{(1)}(x;z)-1)e^{-2izx} \\ zr_-(z)(\eta_+^{(2)}(x;z)-1)e^{2izx} & -z\bar{r}_+(z)\mu_-^{(2)}(x;z)e^{-2izx}\end{array}\right].\end{aligned}$$

Recall that $\lambda r_-(z) \in L^\infty_z(\mathbb{R})$ by Proposition 6. The second row vector of $\tilde{F}(x;z)\tau_1(\lambda)$ and the first row vector of $\tilde{F}(x;z)\tau_2(\lambda)$ belongs to $L^2_z(\mathbb{R})$, thanks to bounds (4.161) and (4.162) of Proposition 10, as well as bounds (4.148), (4.180), and (4.181). As a result, repeating the previous analysis, we obtain the bounds (4.174) and (4.175). □

4.4.3 Reconstruction formulas

We shall now recover the potential u of the Kaup–Newell spectral problem (4.14) from the matrices $M_\pm$, which satisfy the integral equations (4.146). This will gives us the map

$$H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R}) \ni (r_-, r_+) \mapsto u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}), \tag{4.185}$$

where r_- and r_+ are related by (4.108).

Let us recall the connection formulas between the potential u and the Jost functions of the direct scattering transform in Section 2. By Lemma 17, if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, then

$$\partial_x \left(\bar{u}(x) e^{\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2 dy} \right) = 2i \lim_{|z|\to\infty} z m_\pm^{(2)}(x;z). \tag{4.186}$$

On the other hand, by Lemma 17 and the representation (4.103), if $u \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$, then

$$u(x) e^{-\frac{1}{2i}\int_{\pm\infty}^x |u(y)|^2 dy} = -4 \lim_{|z|\to\infty} z p_\pm^{(1)}(x;z). \tag{4.187}$$

We shall now study properties of the potential u recovered by equations (4.186) and (4.187) from properties of the matrices $M_\pm$. The two choices in the reconstruction formulas (4.186) and (4.187) are useful for controlling the potential u on the positive and negative half-lines. We shall proceed separately with the estimates on the two half-lines.

Estimates on the positive half-line

By comparing (4.112) with (4.120), we rewrite the reconstruction formulas (4.186) and (4.187) for the choice of $m_+^{(2)}$ and $p_+^{(1)}$ as follows:

$$\partial_x \left(\bar{u}(x) e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} \right) = 2i e^{-\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} \lim_{|z|\to\infty} z \mu_-^{(2)}(x;z) \tag{4.188}$$

and

$$u(x) e^{-\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} = -4 e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} \lim_{|z|\to\infty} z \eta_+^{(1)}(x;z) \tag{4.189}$$

Since $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$, we have $R(x;\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ for every $x \in \mathbb{R}$, so that the asymptotic limit (4.124) in Proposition 7 is justified since $M_-(x;\cdot) - I \in L^2(\mathbb{R})$ by Lemma 24. Therefore, we use the solution representation (4.147) and rewrite the reconstruction formulas (4.188) and (4.189) in the explicit form

$$\begin{aligned} & e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} \partial_x \left(\bar{u}(x) e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} \right) \\ & \quad = -\frac{1}{\pi} \int_{\mathbb{R}} r_-(z) e^{2izx} \left[\eta_-^{(2)}(x;z) + \bar{r}_+(z) e^{-2izx} \mu_-^{(2)}(x;z) \right] dz \\ & \quad = -\frac{1}{\pi} \int_{\mathbb{R}} r_-(z) e^{2izx} \eta_+^{(2)}(x;z) dz \end{aligned} \tag{4.190}$$

and

$$u(x) e^{i\int_{+\infty}^x |u(y)|^2 dy} = \frac{2}{\pi i} \int_{\mathbb{R}} \bar{r}_+(z) e^{-2izx} \mu_-^{(1)}(x;z) dz. \tag{4.191}$$

where we have used the jump condition (4.116) for the second equality in (4.190).

If $r_+, r_- \in H^1(\mathbb{R})$, then the reconstruction formulas (4.190) and (4.191) recover u in class $H^{1,1}(\mathbb{R}^+)$. Furthermore, if $r_+, r_- \in L^{2,1}(\mathbb{R})$, then u is in class $H^2(\mathbb{R}^+)$.

Lemma 26. *Let $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied.*

Then, $u \in H^2(\mathbb{R}^+) \cap H^{1,1}(\mathbb{R}^+)$ satisfies the bound

$$\|u\|_{H^2(\mathbb{R}^+)\cap H^{1,1}(\mathbb{R}^+)} \leq C\left(\|r_+\|_{H^1\cap L^{2,1}} + \|r_-\|_{H^1\cap L^{2,1}}\right), \tag{4.192}$$

where C is a positive constant that depends on $\|r_\pm\|_{H^1\cap L^{2,1}}$.

Proof. We use the reconstruction formula (4.191) rewritten as follows:

$$\begin{aligned} u(x)e^{i\int_{+\infty}^x |u(y)|^2dy} &= \frac{2}{\pi i}\int_{\mathbb{R}} \bar{r}_+(z)e^{-2izx}dz \\ &+\frac{2}{\pi i}\int_{\mathbb{R}} \bar{r}_+(z)e^{-2izx}\left[\mu_-^{(1)}(x;z)-1\right]dz. \end{aligned} \tag{4.193}$$

The first term is controlled in $L^{2,1}(\mathbb{R})$ because $\overline{r}_+$ is in $H^1(\mathbb{R})$ and its Fourier transform $\widehat{\overline{r}_+}$ is in $L^{2,1}(\mathbb{R})$. To control the second term in $L^{2,1}(\mathbb{R}^+)$, we denote

$$I(x) := \int_{-\infty}^{\infty} \bar{r}_+(z)e^{-2izx}\left[\mu_-^{(1)}(x;z)-1\right]dz,$$

use the inhomogeneous equation (4.166), and integrate by parts to obtain

$$I(x) = -\int_{\mathbb{R}} r_-(z)\eta_+^{(1)}(x;z)e^{2izx}\mathcal{P}^+\left(\overline{r}_+(z)e^{-2izx}\right)(z)dz.$$

By bounds (4.157) in Proposition 10, bound (4.173) in Lemma 25, and the Cauchy–Schwarz inequality, we have for every $x_0 \in \mathbb{R}^+$,

$$\begin{aligned} &\sup_{x\in(x_0,\infty)} |\langle x\rangle^2 I(x)| \\ &\leq \|r_-\|_{L^\infty} \sup_{x\in(x_0,\infty)} \|\langle x\rangle \eta_+^{(1)}(x;z)\|_{L^2_z} \sup_{x\in(x_0,\infty)} \left\|\langle x\rangle\mathcal{P}^+\left(\overline{r}_+(z)e^{-2izx}\right)\right\|_{L^2_z} \\ &\leq C\|r_+\|^2_{H^1}, \end{aligned}$$

where the positive constant C only depends on $\|r_\pm\|_{L^\infty}$. By combining the estimates for the two terms with the triangle inequality, we obtain the bound

$$\|u\|_{L^{2,1}(\mathbb{R}^+)} \leq C\left(1+\|r_+\|_{H^1}\right)\|r_+\|_{H^1}. \tag{4.194}$$

On the other hand, the reconstruction formula (4.190) can be rewritten in the form

$$\begin{aligned} e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2dy}\partial_x\left(\bar{u}(x)e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2dy}\right) &= -\frac{1}{\pi}\int_{\mathbb{R}} r_-(z)e^{2izx}dz \\ &-\frac{1}{\pi}\int_{\mathbb{R}} r_-(z)e^{2izx}\left[\eta_+^{(2)}(x;z)-1\right]dz. \end{aligned} \tag{4.195}$$

Using the same analysis as above yields the bound

$$\left\|\partial_x\left(\bar{u}e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2dy}\right)\right\|_{L^{2,1}(\mathbb{R}^+)} \leq C\left(1+\|r_-\|_{H^1}\right)\|r_-\|_{H^1}, \tag{4.196}$$

where C is another positive constant that depends on $\|r_\pm\|_{L^\infty}$. Combining bounds (4.194) and (4.196), we set $v(x) := u(x)e^{-\frac{1}{2i}\int_{+\infty}^{x}|u(y)|^2dy}$ and obtain

$$\|v\|_{H^{1,1}(\mathbb{R}^+)} \leq C\left(\|r_+\|_{H^1} + \|r_-\|_{H^1}\right), \tag{4.197}$$

where C is a new positive constant that depends on $\|r_\pm\|_{H^1}$. Since $|v(x)| = |u(x)|$ and $H^1(\mathbb{R})$ is embedded into $L^6(\mathbb{R})$, the estimate (4.197) implies the bound

$$\|u\|_{H^{1,1}(\mathbb{R}^+)} \leq C\left(\|r_+\|_{H^1} + \|r_-\|_{H^1}\right), \tag{4.198}$$

where C is a positive constant that depends on $\|r_\pm\|_{H^1}$.

In order to obtain the estimate u in $H^2(\mathbb{R}^+)$ and complete the proof of the bound (4.192), we differentiate I in x, substitute the inhomogeneous equation (4.166) and its x derivative, and integrate by parts to obtain

$$\begin{aligned} I'(x) &= -2i\int_{-\infty}^{\infty} z\bar{r}_+(z)e^{-2izx}\left[\mu_-^{(1)}(x;z)-1\right]dz + \int_{-\infty}^{\infty}\bar{r}_+(z)e^{-2izx}\partial_x\mu_-^{(1)}(x;z)dz \\ &= 2i\int_{-\infty}^{\infty} r_-(z)\eta_+^{(1)}(x;z)e^{2izx}\mathcal{P}^+(z\bar{r}_+(z)e^{-2izx})(z)dz \\ &\quad -2i\int_{-\infty}^{\infty} zr_-(z)\eta_+^{(1)}(x;z)e^{2izx}\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})(z)dz \\ &\quad -\int_{-\infty}^{\infty} r_-(z)\partial_x\eta_+^{(1)}(x;z)e^{2izx}\mathcal{P}^+(\bar{r}_+(z)e^{-2izx})(z)dz. \end{aligned}$$

Using bounds (4.157), (4.159) and (4.161) in Proposition 10, bounds (4.173) and (4.175) in Lemma 25, as well as the Cauchy–Schwarz inequality, we have for every $x_0 \in \mathbb{R}^+$,

$$\begin{aligned} \sup_{x\in(x_0,\infty)}|\langle x\rangle I'(x)| &\leq 2\|r_-\|_{L^\infty}\sup_{x\in(x_0,\infty)}\|\langle x\rangle\eta_+^{(1)}(x;z)\|_{L^2_z}\sup_{x\in(x_0,\infty)}\left\|\mathcal{P}^+\left(z\bar{r}_+(z)e^{-2izx}\right)\right\|_{L^2_z} \\ &+2\|zr_-\|_{L^2}\sup_{x\in(x_0,\infty)}\|\langle x\rangle\eta_+^{(1)}(x;z)\|_{L^2_z}\sup_{x\in(x_0,\infty)}\left\|\mathcal{P}^+\left(\bar{r}_+(z)e^{-2izx}\right)\right\|_{L^\infty_z} \\ &+\|r_-\|_{L^\infty}\sup_{x\in(x_0,\infty)}\|\partial_x\eta_+^{(1)}(x;z)\|_{L^2_z}\sup_{x\in(x_0,\infty)}\left\|\langle x\rangle\mathcal{P}^+\left(\bar{r}_+(z)e^{-2izx}\right)\right\|_{L^2_z} \\ &\leq C\|r_-\|_{H^1\cap L^{2,1}}\|r_+\|_{H^1\cap L^{2,1}}\left(\|r_+\|_{H^1\cap L^{2,1}} + \|r_-\|_{H^1\cap L^{2,1}}\right), \end{aligned}$$

where C is a positive constant that only depends on $\|r_\pm\|_{L^\infty}$. This bound on $\sup_{x\in\mathbb{R}^+}|\langle x\rangle I'(x)|$ is sufficient to control I' in $L^2(\mathbb{R}^+)$ norm and hence the derivative of (4.193) in x. Using the same analysis for the derivative of (4.195) in x yields similar estimates. The proof of the bound (4.192) is complete. □

By Lemma 26, we obtain the existence of the mapping

$$H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R}) \ni (r_-, r_+) \mapsto u \in H^2(\mathbb{R}^+) \cap H^{1,1}(\mathbb{R}^+). \tag{4.199}$$

We now show that this map is Lipschitz.

Corollary 11. *Let* $r_\pm, \tilde{r}_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ *satisfy* $\|r_\pm\|_{H^1\cap L^{2,1}}, \|\tilde{r}_\pm\|_{H^1\cap L^{2,1}} \leq \rho$ *for some* $\rho > 0$. *Denote the corresponding potentials by* u *and* $\tilde{u}$ *respectively. Then, there is a positive* ρ*-dependent constant* $C(\rho)$ *such that*

$$\|u - \tilde{u}\|_{H^2(\mathbb{R}^+)\cap H^{1,1}(\mathbb{R}^+)} \leq C(\rho)\left(\|r_+ - \tilde{r}_+\|_{H^1\cap L^{2,1}} + \|r_- - \tilde{r}_-\|_{H^1\cap L^{2,1}}\right). \quad (4.200)$$

Proof. By the estimates in Lemma 26, if $r_\pm \in H^1(\mathbb{R})\cap L^{2,1}(\mathbb{R})$, then the quantities

$$v(x) := u(x)e^{i\int_{+\infty}^x |u(y)|^2 dy} \quad \text{and} \quad w(x) := \left(\partial_x u(x) + \frac{i}{2}|u(x)|^2 u(x)\right) e^{i\int_{+\infty}^x |u(y)|^2 dy},$$

are defined in function space $H^1(\mathbb{R}^+) \cap L^{2,1}(\mathbb{R}^+)$. Lipschitz continuity of the corresponding mappings follows from the reconstruction formula (4.193) and (4.195) by repeating the same estimates in Lemma 26. Since $|v| = |u|$, we can write

$$u - \tilde{u} = (v - \tilde{v})e^{-i\int_{+\infty}^x |v(y)|^2 dy} + \tilde{v}\left(e^{-i\int_{+\infty}^x |v(y)|^2 dy} - e^{-i\int_{+\infty}^x |\tilde{v}(y)|^2 dy}\right).$$

Therefore, Lipschitz continuity of the mapping $(r_+, r_-) \mapsto v \in H^1(\mathbb{R}^+) \cap L^{2,1}(\mathbb{R}^+)$ is translated to Lipschitz continuity of the mapping $(r_+, r_-) \mapsto u \in H^1(\mathbb{R}^+) \cap L^{2,1}(\mathbb{R}^+)$. Using a similar representation for $\partial_x u$ in terms of v and w, we obtain Lipschitz continuity of the mapping (4.199) with the bound (4.200). ☐

Estimates on the negative half-line

Estimates on the positive half-line were found from the reconstruction formulas (4.188) and (4.189), which only use estimates of vector columns μ_- and η_+, as seen in (4.190) and (4.191). By comparing (4.112) with (4.120), we can rewrite the reconstruction formulas (4.186) and (4.187) for the lower choice of $m_-^{(2)}$ and $p_-^{(2)}$ as follows:

$$\partial_x\left(\bar{u}(x)e^{\frac{1}{2i}\int_{-\infty}^x |u(y)|^2 dy}\right) = 2ie^{-\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy} a_\infty \lim_{|z|\to\infty} z\mu_+^{(2)}(x;z) \quad (4.201)$$

and

$$u(x)e^{-\frac{1}{2i}\int_{-\infty}^x |u(y)|^2 dy} = -4e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy}\bar{a}_\infty \lim_{|z|\to\infty} z\eta_-^{(1)}(x;z), \quad (4.202)$$

where $a_\infty := \lim_{|z|\to\infty} a(z) = e^{\frac{1}{2i}\int_{\mathbb{R}} |u(y)|^2 dy}$. If we now use the same solution representation (4.147) in the reconstruction formulas (4.201) and (4.202), we obtain the same explicit expressions (4.190) and (4.191). On the other hand, if we rewrite the Riemann–Hilbert problem (4.116) in an equivalent form, we will be able to find nontrivial representation formulas for u, which are useful on the negative half-line. To do so, we need to factorize the scattering matrix $R(x;z)$ in an equivalent form.

Let us consider the scalar Riemann–Hilbert problem

$$\begin{cases} \delta_+(z) - \delta_-(z) = \bar{r}_+(z)r_-(z)\delta_-(z), & z \in \mathbb{R}, \\ \delta_\pm(z) \to 1 \quad \text{as} \quad |z| \to \infty, \end{cases} \quad (4.203)$$

and look for analytic continuations of functions $\delta_\pm$ in $\mathbb{C}^\pm$. The solution to the scalar Riemann–Hilbert problem (4.203) and some useful estimates are reported in the following two propositions, where we recall from (4.109) that

$$\begin{cases} 1+\overline{r}_+(z)r_-(z) = 1+|r(\lambda)|^2 \geq 1, & z\in\mathbb{R}^+, \\ 1+\overline{r}_+(z)r_-(z) = 1-|r(\lambda)|^2 \geq c_0^2 > 0, & z\in\mathbb{R}^-, \end{cases}$$

where the latter inequality is due to (4.100).

Proposition 11. *Let $r_\pm \in H^1(\mathbb{R})\cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied. There exists unique analytic functions $\delta_\pm$ in $\mathbb{C}^\pm$ of the form*

$$\delta(z) = e^{\mathcal{C}\log(1+\overline{r}_+r_-)}, \quad z\in\mathbb{C}^\pm, \tag{4.204}$$

which solve the scalar Riemann–Hilbert problem (4.203) and which have the limits

$$\delta_\pm(z) = e^{\mathcal{P}^\pm\log(1+\overline{r}_+r_-)}, \quad z\in\mathbb{R}, \tag{4.205}$$

as $z\in\mathbb{C}^\pm$ approaches to a point on the real axis by any non-tangential contour in $\mathbb{C}^\pm$.

Proof. First, we prove that $\log(1+\overline{r}_+r_-)\in L^1(\mathbb{R})$. Indeed, since $r_\pm\in L^{2,1}_z(\mathbb{R})\cap L^\infty(\mathbb{R})$, we have $\overline{r}_+r_-\in L^1(\mathbb{R})$. Furthermore, it follows from the representation (4.107) as well as from Propositions 5 and 6 that

$$\langle z\rangle|r(\lambda)| \leq |r(\lambda)| + \frac{1}{2}|\lambda||r_-(z)| \leq C, \quad z\in\mathbb{R},$$

where C is a positive constant. Therefore,

$$\log(1+|r(\lambda)|^2) \leq \log(1+C^2\langle z\rangle^{-2}), \quad z\in\mathbb{R}^+, \quad \lambda\in\mathbb{R},$$

so that $\log(1+\overline{r}_+r_-)\in L^1(\mathbb{R}^+)$. On the other hand, it follows from the inequality (4.100) that

$$|\log(1-|r(\lambda)|^2)| \leq -\log(1-C^2\langle z\rangle^{-2}), \quad z\in\mathbb{R}^-, \quad \lambda\in\mathbb{R},$$

so that $\log(1+\overline{r}_+r_-)\in L^1(\mathbb{R}^-)$.

Thus, we have $\log(1+\overline{r}_+r_-)\in L^1(\mathbb{R})$. It also follows from the above estimates that $\log(1+\overline{r}_+r_-)\in L^\infty(\mathbb{R})$. By Hölder inequality, we hence obtain $\log(1+\overline{r}_+r_-)\in L^2(\mathbb{R})$. By Proposition 7 with $p=2$, the expression (4.204) defines unique analytic functions in $\mathbb{C}^\pm$, which recover the limits (4.205) and the limits at infinity: $\lim_{|z|\to\infty}\delta_\pm(z)=1$. Finally, since $\mathcal{P}^+-\mathcal{P}^-=I$, we obtain

$$\delta_+(z)\delta_-^{-1}(z) = e^{\log(1+\overline{r}_+(z)r_-(z))} = 1+\overline{r}_+(z)r_-(z), \quad z\in\mathbb{R},$$

so that $\delta_\pm$ given by (4.204) satisfy the scalar Riemann–Hilbert problem (4.203). □

Proposition 12. *Let $r_\pm\in H^1(\mathbb{R})\cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied. Then, $\delta_+\delta_-r_\pm\in H^1(\mathbb{R})\cap L^{2,1}(\mathbb{R})$.*

Proof. We first note that $\mathcal{P}^+ + \mathcal{P}^- = -i\mathcal{H}$ due to the projection formulas (4.125), where $\mathcal{H}$ is the Hilbert transform. Therefore, we write

$$\delta_+\delta_- = e^{-i\mathcal{H}\log(1+\overline{r}_+r_-)}.$$

Since $\log(1+\overline{r}_+r_-) \in L^2(\mathbb{R})$, we have $\mathcal{H}\log(1+\overline{r}_+r_-) \in L^2(\mathbb{R})$ being a real-valued function. Therefore, $|\delta_+(z)\delta_-(z)| = 1$ for almost every $z \in \mathbb{R}$. Then, $\delta_+\delta_-r_\pm \in L^{2,1}(\mathbb{R})$ follows from $r_\pm \in L^{2,1}(\mathbb{R})$.

It remains to show that $\partial_z\delta_+\delta_-r_\pm \in L^2(\mathbb{R})$. To do so, we shall prove that $\partial_z\mathcal{H}\log(1+\overline{r}_+r_-) \in L^2(\mathbb{R})$. Due to Parseval's identity and the fact $\|\mathcal{H}f\|_{L^2} = \|f\|_{L^2}$ for every $f \in L^2(\mathbb{R})$, we obtain

$$\|\partial_z\mathcal{H}\log(1+\overline{r}_+r_-)\|_{L^2} = \|\partial_z\log(1+\overline{r}_+r_-)\|_{L^2}.$$

The right-hand side is bounded since $\partial_z\log(1+\overline{r}_+r_-) = \frac{\partial_z(\overline{r}_+r_-)}{1+\overline{r}_+r_-} \in L^2(\mathbb{R})$ under the conditions of the proposition. The assertion $\partial_z\delta_+\delta_-r_\pm \in L^2(\mathbb{R})$ is proved. □

Next, we factorize the scattering matrix $R(x;z)$ in an equivalent form:

$$\begin{bmatrix}\delta_-(z) & 0\\ 0 & \delta_-^{-1}(z)\end{bmatrix}[I+R(x;z)]\begin{bmatrix}\delta_+^{-1}(z) & 0\\ 0 & \delta_+(z)\end{bmatrix}$$
$$= \begin{bmatrix}1 & \delta_-(z)\delta_+(z)\overline{r}_+(z)e^{-2izx}\\ \overline{\delta}_+(z)\overline{\delta}_-(z)r_-(z)e^{2izx} & 1+\overline{r}_+(z)r_-(z)\end{bmatrix},$$

where we have used $\delta_-^{-1}\delta_+^{-1} = \overline{\delta_-\delta_+}$. Let us now define new jump matrix

$$\tilde{R}_\delta(x;z) := \begin{bmatrix}0 & \overline{r}_{+,\delta}(z)e^{-2ixz}\\ r_{-,\delta}(z)e^{2ixz} & \overline{r}_{+,\delta}(z)r_{-,\delta}(z)\end{bmatrix},$$

associated with new scattering data

$$r_{\pm,\delta}(z) := \overline{\delta}_+(z)\overline{\delta}_-(z)r_\pm(z).$$

By Proposition 12, we have $r_{\pm,\delta} \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ similarly to the scattering data $r_\pm$.

By using the functions $M_\pm(x;z)$ and $\delta_\pm(z)$, we define functions

$$M_{\pm,\delta}(x;z) := M_\pm(x;z)\begin{bmatrix}\delta_\pm^{-1}(z) & 0\\ 0 & \delta_\pm(z)\end{bmatrix}. \tag{4.206}$$

By Proposition 11, the new functions $M_{\pm,\delta}(x;\cdot)$ are analytic in $\mathbb{C}^\pm$ and have the same limit I as $|z| \to \infty$. On the real axis, the new functions satisfy the jump condition associated with the jump matrix $\tilde{R}_\delta(x;z)$. All together, the new Riemann–Hilbert problem

$$\begin{cases} M_{+,\delta}(x;z) - M_{-,\delta}(x;z) = M_{-,\delta}(x;z)\tilde{R}_\delta(x;z), \quad z \in \mathbb{R},\\ \lim_{|z|\to\infty} M_{\pm,\delta}(x;z) = I, \end{cases} \tag{4.207}$$

follows from the previous Riemann–Hilbert problem (4.116). By Corollary 9 and analysis preceding Lemma 24, the Riemann–Hilbert problem (4.207) admits a unique solution, which is given by the Cauchy operators in the form:

$$M_{\pm,\delta}(x;z) = I + \mathcal{C}\left(M_{-,\delta}(x;\cdot)\tilde{R}_\delta(x;\cdot)\right)(z), \quad z \in \mathbb{C}^\pm. \tag{4.208}$$

Let us denote the vector columns of $M_{\pm,\delta}$ by $M_{\pm,\delta} = [\mu_{\pm,\delta}, \eta_{\pm,\delta}]$. What is nice in the construction of $M_{\pm,\delta}$ that

$$\lim_{|z|\to\infty} z\mu^{(2)}_{\pm,\delta}(x;z) = \lim_{|z|\to\infty} z\mu^{(2)}_{\pm}(x;z) \quad \text{and} \quad \lim_{|z|\to\infty} z\eta^{(1)}_{\pm,\delta}(x;z) = \lim_{|z|\to\infty} z\eta^{(1)}_{\pm}(x;z).$$

Since $r_{\pm,\delta} \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$, we have $\tilde{R}_\delta(x;\cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ for every $x \in \mathbb{R}$, so that the asymptotic limit (4.124) in Proposition 7 is justified for the integral representation (4.208). As a result, the reconstruction formulas (4.201) and (4.202) can be rewritten in the explicit form:

$$e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy}\partial_x\left(\bar{u}(x)e^{\frac{1}{2i}\int_{+\infty}^x |u(y)|^2 dy}\right) = -\frac{1}{\pi}\int_{\mathbb{R}} r_-(z)e^{2izx}\eta^{(2)}_{-,\delta}(x;z)dz \tag{4.209}$$

and

$$\begin{aligned} u(x)e^{i\int_{+\infty}^x |u(y)|^2 dy} &= \frac{2}{\pi i}\int_{\mathbb{R}} \bar{r}_{+,\delta}(z)e^{-2izx}\left[\mu^{(1)}_{-,\delta}(x;z) + r_{-,\delta}(z)e^{2izx}\eta^{(1)}_{-,\delta}(x;z)\right]dz \\ &= \frac{2}{\pi i}\int_{\mathbb{R}} \bar{r}_{+,\delta}(z)e^{-2izx}\mu^{(1)}_{+,\delta}(x;z)dz, \end{aligned} \tag{4.210}$$

where we have used the first equation of the Riemann–Hilbert problem (4.207) for the second equality in (4.210).

The reconstruction formulas (4.209) and (4.210) can be studied similarly to the analysis in the previous subsection. First, we obtain the system of integral equations for vectors $\mu_{+,\delta}$ and $\eta_{-,\delta}$ from projections of the solution representation (4.208) to the real line:

$$\mu_{+,\delta}(x;z) = e_1 + \mathcal{P}^+\left(r_{-,\delta}e^{2izx}\eta_{-,\delta}(x;\cdot)\right)(z), \tag{4.211}$$
$$\eta_{-,\delta}(x;z) = e_2 + \mathcal{P}^-\left(\bar{r}_{+,\delta}e^{-2izx}\mu_{+,\delta}(x;\cdot)\right)(z). \tag{4.212}$$

The integral equations above can be written as

$$G_\delta - \mathcal{P}^-(G_\delta R_\delta) = F_\delta, \tag{4.213}$$

where

$$G_\delta(x;z) := [\mu_{+,\delta}(x;z) - e_1, \eta_{-,\delta}(x;z) - e_2]\begin{bmatrix} 1 & 0 \\ -r_{-,\delta}(z)e^{2izx} & 1 \end{bmatrix}$$

and

$$F_\delta(x;z) := \left[e_2\mathcal{P}^+(r_{-,\delta}(z)e^{2izx}),\ e_1\mathcal{P}^-(\bar{r}_{+,\delta}(z)e^{-2izx})\right].$$

The estimates of Proposition 10, Lemma 25, Lemma 26, and Corollary 11 apply to the system of integral equations (4.211) and (4.212) with the only change: $x_0 \in \mathbb{R}^+$ is replaced by $x_0 \in \mathbb{R}^-$ because the operators $\mathcal{P}^+$ and $\mathcal{P}^-$ swap their places in comparison with the system (4.176). As a result, we extend the statements of Lemma 26 and Corollary 11 to the negative half-line. This construction yields existence and Lipschitz continuity of the mapping

$$H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R}) \ni (r_-, r_+) \mapsto u \in H^2(\mathbb{R}^-) \cap H^{1,1}(\mathbb{R}^-). \tag{4.214}$$

Lemma 27. *Let $r_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ such that the inequality (4.100) is satisfied. Then, $u \in H^2(\mathbb{R}^-) \cap H^{1,1}(\mathbb{R}^-)$ satisfies the bound*

$$\|u\|_{H^2(\mathbb{R}^-)\cap H^{1,1}(\mathbb{R}^-)} \leq C \left(\|r_{+,\delta}\|_{H^1\cap L^{2,1}} + \|r_{-,\delta}\|_{H^1\cap L^{2,1}}\right), \tag{4.215}$$

where C is a positive constant that depends on $\|r_{\pm,\delta}\|_{H^1\cap L^{2,1}}$.

Corollary 12. *Let $r_\pm, \tilde{r}_\pm \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ satisfy $\|r_\pm\|_{H^1\cap L^{2,1}}, \|\tilde{r}_\pm\|_{H^1\cap L^{2,1}} \leq \rho$ for some $\rho > 0$. Denote the corresponding potentials by u and $\tilde{u}$ respectively. Then, there is a positive ρ-dependent constant $C(\rho)$ such that*

$$\|u - \tilde{u}\|_{H^2(\mathbb{R}^-)\cap H^{1,1}(\mathbb{R}^-)} \leq C(\rho) \left(\|r_+ - \tilde{r}_+\|_{H^1\cap L^{2,1}} + \|r_- - \tilde{r}_-\|_{H^1\cap L^{2,1}}\right). \tag{4.216}$$

Remark 17. *Since Corollaries 11 and 12 yield Lipschitz continuity of the mappings (4.199) and (4.214) for every $r_\pm$, $\tilde{r}_\pm$ in a ball of a fixed (but possibly large) radius ρ, the mappings (4.199) and (4.214) are one-to-one for every $r_\pm$ in the ball.*

4.5 Proof of the main result

Thanks to the local well-posedness theory in [106, 107] and the weighted estimates in [44, 45], there exists a local solution $u(t,\cdot) \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ to the Cauchy problem (4.1) with an initial data $u_0 \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ for $t \in [0,T]$ for some finite $T > 0$.

For every $t \in [0,T]$, we define fundamental solutions

$$\psi(t,x;\lambda) := e^{-i2\lambda^4 t - i\lambda^2 x}\varphi_\pm(t,x;\lambda)$$

and

$$\psi(t,x;\lambda) := e^{i2\lambda^4 t + i\lambda^2 x}\phi_\pm(t,x;\lambda)$$

to the Kaup–Newell spectral problem (4.9) and the time-evolution problem (4.10) associated with the potential $u(t,x)$ that belongs to $C([0,T], H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R}))$. By Corollaries 5 and 6, the bounded Jost functions $\varphi_\pm(t,x;\lambda)$ and $\psi_\pm(t,x;\lambda)$ have the same analytic property in λ plane and satisfy the same boundary conditions

$$\begin{cases} \varphi_\pm(t,x;\lambda) \to e_1 \\ \phi_\pm(t,x;\lambda) \to e_2 \end{cases} \quad \text{as} \quad x \to \pm\infty$$

for every $t \in [0,T]$. From linear independence of two solutions to the Kaup–Newell spectral problem (4.9), the bounded Jost functions satisfy the scattering relation

$$\varphi_-(t,x;\lambda) = a(\lambda)\varphi_+(t,x;\lambda) + b(\lambda)e^{2i\lambda^2 x + 4i\lambda^4 t}\phi_+(t,x;\lambda), \quad x \in \mathbb{R}, \quad \lambda \in \mathbb{R} \cup i\mathbb{R}, \tag{4.217}$$

where the scattering coefficients $a(\lambda)$ and $b(\lambda)$ are independent of (t,x) due to the fact that the matrices of the linear system (4.9) and (4.10) have zero trace. Indeed, in this case, the Wronskian determinants are independent of (t,x), so that we have

$$\begin{aligned} a(\lambda) &= W(\varphi_-(t,x;\lambda)e^{-i2\lambda^4 t - i\lambda^2 x}, \phi_+(t,x;\lambda)e^{i2\lambda^4 t + i\lambda^2 x}) \\ &= W(\varphi_-(0,0;\lambda), \phi_+(0,0;\lambda)), \\ b(\lambda) &= W(\varphi_+(t,x;\lambda)e^{-i2\lambda^4 t - i\lambda^2 x}, \varphi_-(t,x;\lambda)e^{-i2\lambda^4 t - i\lambda^2 x}) \\ &= W(\varphi_+(0,0;\lambda), \varphi_-(0,0;\lambda)). \end{aligned}$$

By Lemma 19 and assumptions on zeros of a in the λ plane, we can define the time-dependent scattering data

$$r_+(t;z) = -\frac{b(\lambda)e^{4i\lambda^4 t}}{2i\lambda a(\lambda)}, \quad r_-(t;z) = \frac{2i\lambda b(\lambda)e^{4i\lambda^4 t}}{a(\lambda)}, \quad z \in \mathbb{R}, \tag{4.218}$$

so that the scattering relation (4.217) becomes equivalent to the first scattering relation in (4.72). Thus, we define

$$r_\pm(t;z) = r_\pm(0;z)e^{4iz^2 t}, \tag{4.219}$$

where $r_\pm(0;\cdot)$ are initial spectral data found from the initial condition $u(0,\cdot)$ and the direct scattering transform in Section 2. By Lemma 19 and Corollary 7, under the condition that $u_0 \in H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ admits no resonances of the linear equation (4.9), the scattering data $r_\pm(0;\cdot)$ is defined in $H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ and is a Lipschitz continuous function of u_0.

Now the time evolution (4.219) implies that $r_\pm(t;\cdot)$ remains in $H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ for every $t \in [0,T]$. Indeed, we have

$$\|r_\pm(t;\cdot)\|_{L^{2,1}} = \|r_\pm(0;\cdot)\|_{L^{2,1}} \quad \text{and} \quad \|\partial_z r_\pm(t;\cdot) + 4itzr_\pm(t;\cdot)\|_{L^2} = \|\partial_z r(0;\cdot)\|_{L^2}.$$

Hence, $r(t;\cdot) \in H^1(\mathbb{R}) \cap L^{2,1}(\mathbb{R})$ for every $t \in [0,T]$. Moreover, the constraint (4.100) and the relation (4.108) remain valid for every $t \in [0,T]$.

The potential $u(t,\cdot)$ is recovered from the scattering data $r_\pm(t;\cdot)$ with the inverse scattering transform in Section 4. By Lemmas 26, 27 and Corollaries 11, 12, the potential $u(t,\cdot)$ is defined in $H^2(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$ for every $t \in [0,T]$ and is a Lipschitz continuous function of $r(t;\cdot)$. Thus, for every $t \in [0,T)$ we have proved that

$$\begin{aligned}\|u(t,\cdot)\|_{H^2\cap H^{1,1}} &\leq C_1\left(\|r_+(t;\cdot)\|_{H^1\cap L^{2,1}}+\|r_-(t;\cdot)\|_{H^1\cap L^{2,1}}\right)\\ &\leq C_2\left(\|r_+(0;\cdot)\|_{H^1\cap L^{2,1}}+\|r_-(0;\cdot)\|_{H^1\cap L^{2,1}}\right)\\ &\leq C_3\|u_0\|_{H^2\cap H^{1,1}}, \qquad (4.220)\end{aligned}$$

where the positive constants C_1, C_2, and C_3 depends on $\|r_\pm(t;\cdot)\|_{H^1\cap L^{2,1}}$, $(T, \|r_\pm(0;\cdot)\|_{H^1\cap L^{2,1}})$, and $(T, \|u_0\|_{H^2\cap H^{1,1}})$ respectively. Moreover, the map $H^2(\mathbb{R})\cap H^{1,1}(\mathbb{R}) \ni u_0 \mapsto u \in C([0,T], H^2(\mathbb{R})\cap H^{1,1}(\mathbb{R}))$ is Lipschitz continuous.

Since $\|r(t;\cdot)\|_{H^1}$ may grow at most linearly in t and constants C_1, C_2, C_3 in (4.221) depends polynomially on their respective norms, we have

$$\|u(t,\cdot)\|_{H^2\cap H^{1,1}} \leq C(T)\|u_0\|_{H^2\cap H^{1,1}}, \quad t\in[0,T], \qquad (4.221)$$

where the positive constant $C(T)$ (that also depends on $\|u_0\|_{H^2\cap H^{1,1}}$) may grow at most polynomially in T but it remains finite for every $T>0$. From here, we derive a contradiction on the assumption that the local solution $u\in C([0,T], H^2(\mathbb{R})\cap H^{1,1}(\mathbb{R}))$ blows up in a finite time. Indeed, if there exists a maximal existence tim $T_{\max}>0$ such that $\lim_{t\uparrow T_{\max}}\|u(t;\cdot)\|_{H^2\cap H^{1,1}}=\infty$, then the bound (4.221) is violated as $t\uparrow T$, which is impossible. Therefore, the local solution $u\in C([0,T], H^2(\mathbb{R})\cap H^{1,1}(\mathbb{R}))$ can be continued globally in time for every $T>0$. This final argument yields the proof of Theorem 4.

Figure 4.4 illustrates the proof of Theorem 4 and summarizes the main ingredients of our results.

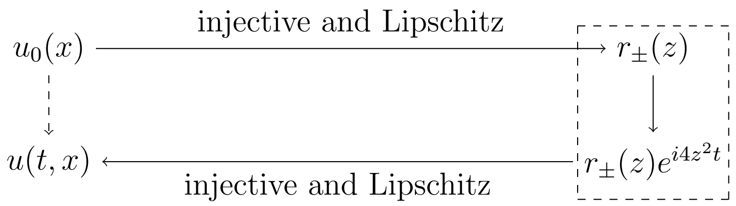


Figure 4.4: The scheme behind the proof of Theorem 4.

Chapter 5

Transverse instability of line solitary waves in massive Dirac equations

5.1 Background

Starting with pioneer contributions of V.E. Zakharov and his school [119], studies of transverse instabilities of line solitary waves in various nonlinear evolution equations have been developed in many different contexts. With the exception of the Kadometsev–Petviashvili-II (KP-II) equation, line solitary waves in many evolution equations are spectrally unstable with respect to transverse periodic perturbations [60].

More recently, it was proved for the prototypical model of the KP-I equation that the line solitary waves under the transverse perturbations of sufficiently small periods remain spectrally and orbitally stable [94]. Similar thresholds on the period of transverse instability exist in other models such as the elliptic nonlinear Schrödinger (NLS) equation [115] and the Zakharov–Kuznetsov (ZK) equation [85]. Nevertheless, this conclusion is not universal and the line solitary waves can be spectrally unstable for all periods of the transverse perturbations, as it happens for the hyperbolic NLS equation [86].

Conclusions on the transverse stability or instability of line solitary waves may change in the presence of the periodic potentials. In the two-dimensional problems with square periodic potentials, it was found numerically in [49, 58, 116] that line solitary waves are spectrally stable with respect to periodic transverse perturbations if they bifurcate from the so-called X point of the Brillouin zone. Line solitary waves remain spectrally unstable if they bifurcate from the Γ point of the Brillouin zone. These numerical results were rigorously justified in [93] from the analysis of the two-dimensional discrete NLS equation, which models the tight-binding limit of the periodic potentials [92].

For the one-dimensional periodic (stripe) potentials, similar stabilization of the line solitary waves was observed numerically in [117]. In the contrast to these results, it was proven within the tight-binding limit in [93] that transverse instabilities of line solitary waves persist for any parameter configurations of the discrete NLS equation. One of the motivations for our present work is to inspect if the line solitary waves become spectrally stable with respect to the periodic transverse perturbations in periodic stripe potentials far away from the tight–binding limit.

In particular, we employ the massive Dirac equations also known as the coupled-mode equations, which have been derived and justified in the reduction of the Gross–Pitaevskii equation with small periodic potentials [97]. Similar models were also introduced in the context of the periodic stripe potentials in [33], where the primary focus was on the existence and stability of fully localized two-dimensional solitary waves. From the class of massive Dirac models, we will be particularly interested in a generalization of the massive Thirring model [104], for which orbital stability of one-dimensional solitons was proved in our previous work with the help of conserved quantities [88] and auto–Bäcklund transformation [23]. In the present work, we prove analytically that the line solitary waves of the massive Thirring model in two spatial dimensions are spectrally unstable with respect to the periodic transverse perturbations of large periods. The spectral instability is induced by the spatial translation of the line solitary waves. We also show numerically that the instability persists for smaller periods of transverse perturbations.

In the context of numerical results in [117], we now confirm that line solitary waves in the periodic stripe potential remain spectrally unstable with respect to periodic transverse perturbations both in the tight-binding and small-potential limits. The numerical results in [117] are observed apparently in a narrow interval of the existence domain for the line solitary waves supported by the periodic stripe potential.

Different versions of the massive Dirac equations were derived recently in the context of hexagonal potentials. The corresponding systems generalize the massive Gross–Neveu model (also known as the Soler model in $(1 + 1)$ dimensions) [42]. These equations were derived formally in [2, 3] and were justified recently in [36, 37, 38]. Extending the scope of our work, we prove analytically that the line solitary waves of the massive Gross–Neveu model in two spatial dimensions are also spectrally unstable with respect to the periodic perturbations of large periods. The spectral instability is induced by the gauge rotation. Numerical results indicate that the instability exhibits a finite threshold on the period of the transverse perturbations.

The method we employ in our work is relatively old [119] (see review in [60]), although it has not been applied to the class of massive Dirac equations even at the formal level. We develop analysis at the rigorous level of arguments. Our work relies on the resolvent estimates for the spectral stability problem in $(1+1)$ dimensions, where the zero eigenvalue is disjoint from the continuous spectrum, whereas the eigenfunctions for the zero eigenvalue are known from the translational and gauge symmetries of the massive Dirac equations. When the transverse wave number is nonzero but small, the multiple zero eigenvalue split and one can rigorously justify if this splitting induces the spectral instability or not. It becomes notoriously more difficult to prove persistence of instabilities for large transverse wave numbers (small periods), hence, we have to retreat to numerical computations for such studies of the corresponding transverse stability problem.

The approach we undertake in this paper is complementary to the computations based on the Evans function method [52, 53]. Although both approaches stand on rigorous theory based on the implicit function theorem, we believe that the perturbative computations are shorter and provide the binary answer on the spectral stability or instability of the line solitary wave with respect to periodic transverse perturbations in a simple and concise way.

The structure of this paper is as follows. Section 2 introduces two systems of the massive Dirac equations and their line solitary waves in the context of stripe and hexagonal potentials. Section 3 presents the analytical results and gives details of algorithmic computations of the perturbation theory for the massive Thirring and Gross–Neveu models in two spatial dimensions. Section 4 contains numerical approximations of eigenvalues of the spectral stability problem. Transverse instabilities of small-amplitude line solitary waves in more general massive Dirac models are discussed in Section 5.

5.2 Massive Dirac equations

The class of massive Dirac equations on the line can be written in the following general form [17, 84],

$$\begin{cases} i(u_t + u_x) + v = \partial_{\bar{u}} W(u, v, \bar{u}, \bar{v}), \\ i(v_t - v_x) + u = \partial_{\bar{v}} W(u, v, \bar{u}, \bar{v}), \end{cases} \quad x \in \mathbb{R}, \tag{5.1}$$

where the subscripts denote partial differentiation, (u, v) are complex-valued amplitudes in spatial x and temporal t variables, and W is the real function of $(u, v, \bar{u}, \bar{v})$, which is symmetric with respect to u and v and satisfies the gauge invariance

$$W(e^{i\alpha}u, e^{i\alpha}v, e^{-i\alpha}\bar{u}, e^{-i\alpha}\bar{v}) = W(u, v, \bar{u}, \bar{v}) \quad \text{for every } \alpha \in \mathbb{R}.$$

As it is shown in [17], under the constraints on W, it can be expressed in terms of variables $(|u|^2 + |v|^2)$, $|u|^2|v|^2$, and $(\bar{u}v + u\bar{u})$. For the cubic Dirac equations, W is a homogeneous quartic polynomial in u and v, which is written in the most general form as

$$W = c_1(|u|^2 + |v|^2)^2 + c_2|u|^2|v|^2 + c_3(|u|^2 + |v|^2)(\bar{u}v + u\bar{u}) + c_4(\bar{u}v + u\bar{u})^2,$$

where c_1, c_2, c_3, and c_4 are real coefficients. In this case, a family of stationary solitary waves of the massive Dirac equations can be found in the explicit form [17] (see also [73]).

Among various nonlinear Dirac equations, the following particular cases have profound significance in relativity theory:

- $W = |u|^2|v|^2$ - the massive Thirring model [104];
- $W = \frac{1}{2}(\bar{u}v + u\bar{v})^2$ - the massive Gross–Neveu model [42].

Global well-posedness of the massive Thirring model was proved both in $H^s(\mathbb{R})$ for $s > \frac{1}{2}$ [98] and in $L^2(\mathbb{R})$ [15]. Recently, global well-posedness of the massive Gross–Neveu equations was proved both in $H^s(\mathbb{R})$ for $s > \frac{1}{2}$ [48] and in $L^2(\mathbb{R})$ [120].

When the massive Dirac equations are used in modeling of the Gross–Pitaevskii equation with small periodic potentials, the realistic nonlinear terms are typically different from the two particular cases of the massive Thirring and Gross–Neveu models. (In this context, the nonlinear Dirac equations are also known as the coupled-mode equations.) In the following two subsections, we describe the connection of the generalized massive Thirring and Gross–Neveu models in two spatial dimensions to physics of nonlinear states of the Gross–Pitaevskii equation trapped in periodic potentials.

5.2.1 Periodic stripe potentials

In the context of one-dimensional periodic (stripe) potentials, the massive Dirac equations (5.1) can be derived in the following form [33],

$$\left\{ \begin{array}{l} i(u_t + u_x) + v + u_{yy} = (\alpha_1 |u|^2 + \alpha_2 |v|^2) u, \\ i(v_t - v_x) + u + v_{yy} = (\alpha_2 |u|^2 + \alpha_1 |v|^2) v, \end{array} \right. \quad (x,y) \in \mathbb{R}^2, \tag{5.2}$$

where y is a new coordinate in the transverse direction to the stripe potential, the complex-valued amplitudes (u, v) correspond to two counter-propagating resonant Fourier modes interacting with the small periodic potential, and (α_1, α_2) are real-valued parameters. For the stripe potentials, the parameters satisfy the constraint $\alpha_2 = 2\alpha_1$.

To illustrate the derivation of the massive Dirac equations (5.2), we can consider a two-dimensional Gross–Pitaevskii equation with a small periodic potential

$$i\psi_t = -\psi_{xx} - \psi_{yy} + 2\epsilon \cos(x)\psi + |\psi|^2\psi, \tag{5.3}$$

and apply the Fourier decomposition

$$\psi(x,y,t) = \sqrt{\epsilon}\left[u(\epsilon x, \sqrt{\epsilon} y, \epsilon t) e^{\frac{i}{2}x - \frac{i}{4}t} + v(\epsilon x, \sqrt{\epsilon} y, \epsilon t) e^{-\frac{i}{2}x - \frac{i}{4}t} + \epsilon R(x,y,t)\right], \tag{5.4}$$

where ϵ is a small parameter and R is the remainder term. From the condition that R is bounded in variables (x, y, t), it can be obtained from (5.3) and (5.4) that (u, v) satisfy the nonlinear Dirac equations (5.2) with $\alpha_1 = 1$ and $\alpha_2 = 2$. Justification of the Fourier decomposition (5.4) and the nonlinear Dirac equations (5.2) in the context of the Gross–Pitaevskii equation (5.3) has been reported for y-independent perturbations in [97]. Transverse modulations can be taken into account in the same justification procedure, since the error R is bounded in the supremum norm, whereas the solution of the massive Dirac equations (5.2) and the solution of the Gross–Pitaevskii equation (5.3) can be defined in Sobolev spaces of sufficiently high regularity (see Chapter 2.2 in the book [83]).

The stationary y-independent solitary waves of the massive Dirac equations

(5.2) are referred to as the line solitary waves. According to the analysis in [17, 73], the corresponding solutions can be represented in the form

$$u(x,t) = U_\omega(x)e^{i\omega t}, \quad v(x,t) = \bar{U}_\omega(x)e^{i\omega t}, \tag{5.5}$$

where $\omega \in (-1,1)$ is taken in the gap between two branches of the linear wave spectrum of the massive Dirac equations (5.2). The complex-valued amplitude U_ω satisfies the first-order differential equation

$$iU'_\omega - \omega U_\omega + \overline{U}_\omega = (\alpha_1 + \alpha_2)|U_\omega|^2 U_\omega. \tag{5.6}$$

In terms of physical applications, the line solitary wave (5.5) of the massive Dirac equations (5.2) corresponds to a localized mode (the so-called gap soliton) trapped by the periodic stripe potential [83].

In our work, we perform transverse spectral stability analysis of the line solitary waves (5.5) for the particular configuration $\alpha_1 = 0$ and $\alpha_2 = 1$, which correspond to the massive Thirring model on the line [104]. If $\alpha_1 = 0$ and $\alpha_2 = 1$, the solitary wave solution of the differential equation (5.6) exists for every $\omega \in (-1,1)$ in the explicit form

$$U_\omega(x) = \sqrt{2}\mu \frac{\sqrt{1+\omega}\cosh(\mu x) - i\sqrt{1-\omega}\sinh(\mu x)}{\omega + \cosh(2\mu x)}, \tag{5.7}$$

where $\mu = \sqrt{1-\omega^2}$. The solitary wave solution of the differential equation (5.6) is unique up to the translational and gauge transformation. As $\omega \to 1$, the family of solitary waves (5.7) approaches the NLS profile $U_{\omega\to 1}(x) \to \mu \text{sech}(\mu x)$. As $\omega \to -1$, it degenerates into the algebraic profile

$$U_{\omega=-1}(x) = \frac{2(1-2ix)}{1+4x^2}.$$

When y-independent perturbations are considered, solitary waves (5.5) and (5.7) are orbitally stable in the time evolution of the massive Thirring model on the line for every $\omega \in (-1,1)$. The corresponding results were obtained in our previous works [88] in $H^1(\mathbb{R})$ and [23] in a weighted subspace of $L^2(\mathbb{R})$. Note that the solitary waves in more general nonlinear Dirac equations (5.2) are spectrally unstable for y-independent perturbations if $\alpha_1 \neq 0$ but the instability region and the number of unstable eigenvalues depend on the parameter ω [17].

We will show (see Theorem 5 below) that the line solitary waves (5.5) and (5.7) for $\alpha_1 = 0$ and $\alpha_2 = 1$ are spectrally unstable with respect to long periodic transverse perturbations for every $\omega \in (-1,1)$. In the more general massive Dirac equations (5.2), we also show (see Section 5.1 below) that the instability conclusion remains true at least in the small-amplitude limit (when either $\omega \to 1$ or $\omega \to -1$) if $\alpha_1 + \alpha_2 \neq 0$.

5.2.2 Hexagonal potentials

In the context of the hexagonal potentials in two spatial dimensions, the massive Dirac equations can be derived in a different form [38],

$$\left\{ \begin{array}{l} i\partial_t\varphi_1 + i\partial_x\varphi_2 - \partial_y\varphi_2 + \varphi_1 = (\beta_1|\varphi_1|^2 + \beta_2|\varphi_2|^2)\varphi_1, \\ i\partial_t\varphi_2 + i\partial_x\varphi_1 + \partial_y\varphi_1 - \varphi_2 = (\beta_2|\varphi_1|^2 + \beta_1|\varphi_2|^2)\varphi_2, \end{array} \right. \quad (x,y) \in \mathbb{R}^2, \tag{5.8}$$

where (φ_1, φ_2) are complex-valued amplitudes for two resonant Floquet–Bloch modes in the hexagonal lattice and (β_1, β_2) are real-valued positive parameters. The nonlinear Dirac equations (5.8) correspond to equations (4.4)–(4.5) in [38]. Derivation of these equations can also be found in [2, 3]. Justification of the linear part of these equations is performed by Fefferman and Weinstein [37].

To transform the nonlinear Dirac equations (5.8) to the form (5.1), we use the change of variables,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix},$$

and obtain

$$\left\{ \begin{array}{l} i(u_t + u_x) + v + v_y = \beta_1(u|u|^2 + \overline{u}v^2 + 2u|v|^2) + \beta_2\overline{u}(u^2 - v^2), \\ i(v_t - v_x) + u - u_y = \beta_1(v|v|^2 + \overline{v}u^2 + 2v|u|^2) + \beta_2\overline{v}(v^2 - u^2). \end{array} \right. \tag{5.9}$$

In comparison with the nonlinear Dirac equations (5.2), we note that both the cubic nonlinearities and the y-derivative diffractive terms are different.

For the family of line solitary waves (5.5), the complex-valued amplitude U_ω satisfies the first-order differential equation

$$iU_\omega' - \omega U_\omega + \overline{U}_\omega = (3\beta_1 + \beta_2)U_\omega|U_\omega|^2 + (\beta_1 - \beta_2)\overline{U}_\omega^3. \tag{5.10}$$

In terms of physical applications, the line solitary wave (5.5) of the massive Dirac equations (5.9) corresponds to a localized mode trapped by the deformed hexagonal potential with broken Dirac points [3, 38].

In what follows, we perform the transverse spectral stability analysis of the line solitary waves (5.5) for the particular configuration $\beta_1 = -\beta_2 = \frac{1}{2}$, which corresponds to the massive Gross–Neveu model on the line [42]. If $\beta_1 = -\beta_2 = \frac{1}{2}$, the solitary wave solution of the differential equation (5.10) exists for every $\omega \in (0,1)$ in the explicit form

$$U_\omega(x) = \mu\frac{\sqrt{1+\omega}\cosh(\mu x) - i\sqrt{1-\omega}\sinh(\mu x)}{1 + \omega\cosh(2\mu x)}, \tag{5.11}$$

where $\mu = \sqrt{1-\omega^2}$. Again the solitary wave solution of the differential equation (5.10) is unique up to the translational and gauge transformation. The family of solitary waves (5.11) diverges at infinity as $\omega \to 0$ and can not be continued for $\omega \in (-1, 0)$ [8]. As $\omega \to 1$, the family approaches the NLS profile $U_{\omega\to 1}(x) \to 2^{-1/2}\mu\text{sech}(\mu x)$.

When y-independent perturbations are considered, solitary waves (5.5) and

(5.11) are orbitally stable in $H^1(\mathbb{R})$ in the time evolution of the massive Gross–Neveu model for $\omega \approx 1$ [13]. Regarding spectral stability, two numerical studies exist, which show contradictory results to each other. A numerical approach based on the Evans function computation leads to the conclusion on the spectral stability of solitary waves for all $\omega \in (0,1)$ [8, 9]. However, another approach based on the finite-difference discretization indicates existence of $\omega_c \approx 0.6$ such that the family of solitary waves is spectrally stable for $\omega \in (\omega_c, 1)$ and unstable for $\omega \in (0, \omega_c)$ [73, 99]. The presence of additional unstable eigenvalues in the case of y-independent perturbations, if they exist, is not an obstacle in our analysis of transverse stability of line solitary waves.

Our work concerns both y-independent and y-dependent perturbations. In the case of y-independent perturbations, we show numerically (see Section 4.2 below) that the solitary waves of the massive Gross–Neveu model are spectrally stable for every $\omega \in (0,1)$ thus supporting the numerical results of [8, 9] with an independent numerical method based on the Chebyshev interpolation method. In the case of y-periodic perturbations, we show analytically (see Theorem 5 below) that the line solitary waves (5.5) and (5.11) for $\beta_1 = -\beta_2 = \frac{1}{2}$ are spectrally unstable with respect to long periodic transverse perturbations for every $\omega \in (0,1)$. In the more general massive Dirac equations (5.9), we also show (see Section 5.2 below) that the instability conclusion remains true at least in the small-amplitude limit (when either $\omega \to 1$ or $\omega \to -1$) if $\beta_1 \neq 0$.

5.3 Transverse of line solitary waves

We consider two versions (5.2) and (5.9) of the nonlinear Dirac equations for spatial variables (x, y) in the domain $\mathbb{R} \times \mathbb{T}$, where $\mathbb{T} = \mathbb{R}/(L\mathbb{Z})$ is the one dimensional torus and $L \in \mathbb{R}$ is the period of the transverse perturbation. To study stability of the line solitary wave (5.5) under periodic transverse perturbations, we use the Fourier series and write

$$u(x,y,t) = e^{i\omega t}\left[U_\omega(x) + \sum_{n\in\mathbb{Z}} \hat{f}_n(x,t) e^{\frac{2\pi n i y}{L}}\right]. \tag{5.12}$$

In the setting of the spectral stability theory, we are going to use the linear superposition principle and consider just one Fourier mode with continuous parameter $p \in \mathbb{R}$. In the context of the Fourier series (5.12), the parameter p takes the countable set of values $\{\frac{2\pi n}{L}\}_{n\in\mathbb{Z}}$. The limit $p \to 0$ corresponds to the limit of long periodic perturbations with $L \to \infty$.

For each $p \in \mathbb{R}$, we separate the time evolution of the linearized system and introduce the spectral parameter λ in the decomposition $\hat{f}_n(x,t) = \hat{F}_n(x)e^{\lambda t}$. This decomposition reduces the linearized equations for $\hat{f}_n$ to the eigenvalue problem for $\hat{F}_n$ and λ. Performing similar manipulations with four components of the nonlinear

Dirac equations, we set the transverse perturbation in the form

$$u(x,y,t) = e^{i\omega t}[U_\omega(x) + u_1(x)e^{\lambda t+ipy}], \qquad \overline{u}(x,y,t) = e^{-i\omega t}[\overline{U}_\omega(x) + u_2(x)e^{\lambda t+ipy}],$$
$$v(x,y,t) = e^{i\omega t}[\overline{U}_\omega(x) + v_1(x)e^{\lambda t+ipy}], \qquad \overline{v}(x,y,t) = e^{-i\omega t}[U_\omega(x) + v_2(x)e^{\lambda t+ipy}].$$

Remark 18. *Since the perturbation to the line solitary wave is just one linear mode, the component* (u_2, v_2) *are not the complex conjugate of* (u_1, v_1)*. However, given a solution* (u_1, u_2, v_1, v_2) *of the eigenvalue problem for* λ *and* p*, there exists another solution* $(\bar{u}_2, \bar{u}_1, \bar{v}_2, \bar{v}_1)$ *of the same eigenvalue problem for* $\bar{\lambda}$ *and* $-p$*.*

Let $\mathbf{F} = (u_1, u_2, v_1, v_2)^t$. The eigenvalue problem for $\mathbf{F}$ and λ can be written in the form

$$i\lambda\sigma\mathbf{F} = (D_\omega + E_p + W_\omega)\mathbf{F}, \tag{5.13}$$

where

$$D_\omega = \begin{bmatrix} -i\partial_x + \omega & 0 & -1 & 0 \\ 0 & i\partial_x + \omega & 0 & -1 \\ -1 & 0 & i\partial_x + \omega & 0 \\ 0 & -1 & 0 & -i\partial_x + \omega \end{bmatrix}, \quad \sigma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

whereas matrices E_p and W_ω depend on the particular form of the nonlinear Dirac equations. For the model (5.2) with $\alpha_1 = 0$ and $\alpha_2 = 1$, we obtain $E_p = p^2 I$ with

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad W_\omega = \begin{bmatrix} |U_\omega|^2 & 0 & U_\omega^2 & |U_\omega|^2 \\ 0 & |U_\omega|^2 & |U_\omega|^2 & \overline{U}_\omega^2 \\ \overline{U}_\omega^2 & |U_\omega|^2 & |U_\omega|^2 & 0 \\ |U_\omega|^2 & U_\omega^2 & 0 & |U_\omega|^2 \end{bmatrix}, \tag{5.14}$$

where U_ω is given by (5.7). For the model (5.9) with $\beta_1 = -\beta_2 = \frac{1}{2}$, we obtain $E_p = -ipJ$ with

$$J = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad W_\omega = \begin{bmatrix} |U_\omega|^2 & \bar{U}_\omega^2 & U_\omega^2 + 2\overline{U}_\omega^2 & |U_\omega|^2 \\ U_\omega^2 & |U_\omega|^2 & |U_\omega|^2 & 2U_\omega^2 + \overline{U}_\omega^2 \\ 2U_\omega^2 + \overline{U}_\omega^2 & |U_\omega|^2 & |U_\omega|^2 & U_\omega^2 \\ |U_\omega|^2 & U_\omega^2 + 2\overline{U}_\omega^2 & \overline{U}_\omega^2 & |U_\omega|^2 \end{bmatrix}, \tag{5.15}$$

where U_ω is now given by (5.11).

Remark 19. *Let us denote the existence interval for the line solitary wave (5.5) of the nonlinear Dirac equations (5.1) by* $\Omega \subset (-1, 1)$*. For the model (5.2) with* $\alpha_1 = 0$ *and* $\alpha_2 = 1$*, we have* $\Omega = (-1, 1)$*. For the model (5.9) with* $\beta_1 = -\beta_2 = \frac{1}{2}$*, we have* $\Omega = (0, 1)$*.*

The linear operator $D_\omega + E_p + W_\omega$ is self-adjoint in $L^2(\mathbb{R}, \mathbb{C}^4)$ with the domain in $H^1(\mathbb{R}, \mathbb{C}^4)$ thanks to the boundness of the potential term W_ω. We shall use the notation $\langle \cdot, \cdot \rangle_{L^2}$ for the inner product in $L^2(\mathbb{R}, \mathbb{C}^4)$ and the notation $\| \cdot \|_{L^2}$ for the induced norm. Our convention is to apply complex conjugation to the element at the first position of the inner product $\langle\cdot,\cdot\rangle_{L^2}$.

The next elementary result shows that the zero eigenvalue is isolated from the continuous spectrum of the spectral stability problem (5.13) both for $E_p = p^2 I$ and $E_p = -ipJ$ if the real parameter p is sufficiently small.

Proposition 13. *Assume that $W_\omega(x) \to 0$ as $|x| \to \infty$ according to an exponential rate. For every $\omega \in \Omega$ and every $p \in \mathbb{R}$, the continuous spectrum of the stability problem (5.13) is located along the segments $\pm i\Lambda_1$ and $\pm i\Lambda_2$, where for with $E_p = p^2 I$,*

$$\Lambda_1 := \left\{\sqrt{1+k^2}+\omega+p^2, \quad k \in \mathbb{R}\right\}, \quad \Lambda_2 := \left\{\sqrt{1+k^2}-\omega-p^2, \quad k \in \mathbb{R}\right\}, \tag{5.16}$$

whereas for $E_p = -ipJ$,

$$\Lambda_1 := \left\{\sqrt{1+p^2+k^2}+\omega, \quad k \in \mathbb{R}\right\}, \quad \Lambda_2 := \left\{\sqrt{1+p^2+k^2}-\omega, \quad k \in \mathbb{R}\right\}. \tag{5.17}$$

Proof. By Weyl's lemma, the continuous spectrum of the stability problem (5.13) coincides with the purely continuous spectrum of the same problem with $W_\omega \equiv 0$, thanks to the exponential decay of the potential terms W_ω to zero as $|x| \to \infty$. If $W_\omega \equiv 0$, we solve the spectral stability problem (5.13) with the Fourier transform in x, which means that we simply replace ∂_x in the operator D_ω with ik for $k \in \mathbb{R}$ and denote the resulting matrix by $D_{\omega,k}$. As a result, we obtain the matrix eigenvalue problem

$$(D_{\omega,k} + E_p - i\lambda\sigma)\mathbf{F} = 0.$$

After elementary algebraic manipulations, the characteristic equation for this linear system yields four solutions for λ given by $\pm i\Lambda_1$ and $\pm i\Lambda_2$, where the explicit expressions for Λ_1 and Λ_2 are given by (5.16) and (5.17) for $E_p = p^2 I$ and $E_p = -ipJ$, respectively. □

Remark 20. *We note the different role of the matrix E_p in the location of the continuous spectrum for larger values of the real parameter p. If $E_p = p^2 I$, then the two bands $\pm i\Lambda_2$ touches each other for $|p| = p_\omega := \sqrt{1-\omega}$ and overlap for $|p| > p_\omega$. If $E_p = -ipJ$, all the four bands do not overlap for all values of $p \in \mathbb{R}$ and the zero point $\lambda = 0$ is always in the gap between the branches of the continuous spectrum.*

The next result shows that if $p = 0$, then the spectral stability problem (5.13) admits the zero eigenvalue of quadruple multiplicity. The zero eigenvalue is determined by the symmetries of the nonlinear Dirac equations with respect to the spatial translation and the gauge rotation.

Proposition 14. *For every $\omega \in \Omega$ and $p = 0$, the stability problem (5.13) admits exactly two eigenvectors in $H^1(\mathbb{R})$ for the eigenvalue $\lambda = 0$ given by*

$$\mathbf{F}_t = \partial_x \mathbf{U}_\omega, \quad \mathbf{F}_g = i\sigma\mathbf{U}_\omega, \tag{5.18}$$

where $\mathbf{U}_\omega = (U_\omega, \bar{U}_\omega, \bar{U}_\omega, U_\omega)^t$. *For each eigenvector* $\mathbf{F}_{t,g}$, *there exists a generalized eigenvector* $\tilde{\mathbf{F}}_{t,g}$ *in* $H^1(\mathbb{R})$ *from solutions of the inhomogeneous problem*

$$(D_\omega + W_\omega)\mathbf{F} = i\sigma \mathbf{F}_{t,g}, \tag{5.19}$$

in fact, in the explicit form,

$$\tilde{\mathbf{F}}_t = i\omega x\sigma \mathbf{U}_\omega - \frac{1}{2}\tilde{\sigma}\mathbf{U}_\omega, \quad \tilde{\mathbf{F}}_g = \partial_\omega \mathbf{U}_\omega, \tag{5.20}$$

where $\tilde{\sigma} = \mathrm{diag}(1,1,-1,-1)$. *Moreover, if* $\langle \mathbf{F}_{t,g}, \sigma\tilde{\mathbf{F}}_{t,g}\rangle_{L^2} \neq 0$, *no solutions of the inhomogeneous problem*

$$(D_\omega + W_\omega)\mathbf{F} = i\sigma \tilde{\mathbf{F}}_{t,g} \tag{5.21}$$

exist in $H^1(\mathbb{R})$.

Proof. Existence of the eigenvectors (5.18) follows from the two symmetries of the massive Dirac equations and is checked by elementary substitution as $(D_\omega + W_\omega)\mathbf{F}_{t,g} = \mathbf{0}$. Because $(D_\omega + W_\omega)$ is a self-adjoint operator of the fourth order and solutions of $(D_\omega + W_\omega)\mathbf{F} = \mathbf{0}$ have constant Wronskian determinant in x, there exists at most two spatially decaying solutions of these homogeneous equations, which means that the stability problem (5.13) with $p = 0$ admits exactly two eigenvectors in $H^1(\mathbb{R})$ for $\lambda = 0$. Since

$$\langle \mathbf{F}_{t,g}, \sigma\mathbf{F}_{t,g}\rangle_{L^2} = \langle \mathbf{F}_{t,g}, \sigma\mathbf{F}_{g,t}\rangle_{L^2} = 0$$

there exist solutions of the inhomogeneous problem (5.19) in $H^1(\mathbb{R})$. Existence of the generalized eigenvectors (5.20) is checked by elementary substitution. Finally, under the condition $\langle \mathbf{F}_{t,g}, \sigma\tilde{\mathbf{F}}_{t,g}\rangle_{L^2} \neq 0$, no solutions of the inhomogeneous problem (5.21) exist in $H^1(\mathbb{R})$ by Fredholm's alternative. □

Our main result is formulated in the following theorem. The theorem guarantees spectral instability of the line solitary waves with respect to the transverse perturbations of sufficiently large period both for the massive Thirring model and the massive Gross–Neveu model in two spatial dimensions.

Theorem 5. *For every* $\omega \in \Omega$, *there exists* $p_0 > 0$ *such that for every* p *in* $0 < |p| < p_0$, *the spectral stability problem* (5.13) *with either (5.14) or (5.15) admits at least one eigenvalue* λ *with* $\mathrm{Re}(\lambda) > 0$. *Moreover, up to a suitable normalization, as* $p \to 0$, *the corresponding eigenvector* $\mathbf{F}$ *converges in* $L^2(\mathbb{R})$ *to* $\mathbf{F}_t$ *for (5.13) and (5.14) and to* $\mathbf{F}_g$ *for (5.13) and (5.15).*

Simultaneously, there exists at least one pair of purely imaginary eigenvalues λ *of the spectral stability problem* (5.13) *and the corresponding eigenvector* $\mathbf{F}$ *converges as* $p \to 0$ *to the other eigenvector of Proposition 14.*

The proof of Theorem 5 is based on the perturbation theory for the Jordan block associated with the zero eigenvalue of the spectral problem (5.13) existing for $p = 0$, according to Proposition 14. The zero eigenvalue is isolated from the continuous spectrum, according to Proposition 13. Consequently, we do not have to deal with bifurcations from the continuous spectrum (unlike the difficult tasks of the recent work [13]), but can develop straightforward perturbation expansions based on a modification of the Lyapunov–Schmidt reduction method.

A useful technical approach to the perturbation theory for the spectral stability problem (5.13) is based on the block diagonalization of the 4×4 matrix operator into two 2×2 Dirac operators. This block diagonalization technique was introduced in [17] and used for numerical approximations of eigenvalues of the spectral stability problem for the massive Dirac equations. After the block diagonalization, each Dirac operator has a one-dimensional kernel space induced by either translational or gauge symmetries. It enables us to uncouple the invariant subspaces associated with the Jordan block for the zero eigenvalue of the spectral stability problem (5.13) with $p = 0$.

Using the self-similarity transformation matrix

$$T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix}$$

and setting $\mathbf{F} = T\mathbf{V}$, we can rewrite the spectral stability problem (5.13) in the following form:

$$i\lambda T^t \sigma T \mathbf{V} = T^t (D_\omega + E_p + W_\omega) T \mathbf{V}, \tag{5.22}$$

where

$$T^t D_\omega T = \begin{bmatrix} -i\partial_x + \omega & -1 & 0 & 0 \\ -1 & i\partial_x + \omega & 0 & 0 \\ 0 & 0 & -i\partial_x + \omega & 1 \\ 0 & 0 & 1 & i\partial_x + \omega \end{bmatrix}, \quad T^t \sigma T = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \tag{5.23}$$

whereas the transformation of matrices E_p and W_ω depend on the particular form of the nonlinear Dirac equations. For the model (5.2) with $\alpha_1 = 0$ and $\alpha_2 = 1$, we obtain

$$T^t E_p T = p^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T^t W_\omega T = \begin{bmatrix} 2|U_\omega|^2 & U_\omega^2 & 0 & 0 \\ \overline{U}_\omega^2 & 2|U_\omega|^2 & 0 & 0 \\ 0 & 0 & 0 & -U_\omega^2 \\ 0 & 0 & -\overline{U}_\omega^2 & 0 \end{bmatrix}. \tag{5.24}$$

For the model (5.9) with $\beta_1 = -\beta_2 = \frac{1}{2}$, we obtain

$$T^t E_p T = ip \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix},$$

$$T^t W_\omega T = \begin{bmatrix} 2|U_\omega|^2 & U_\omega^2 + 3\bar{U}_\omega^2 & 0 & 0 \\ 3U_\omega^2 + \bar{U}_\omega^2 & 2|U_\omega|^2 & 0 & 0 \\ 0 & 0 & 0 & -U_\omega^2 - \bar{U}_\omega^2 \\ 0 & 0 & -U_\omega^2 - \bar{U}_\omega^2 & 0 \end{bmatrix}. \tag{5.25}$$

Let us apply the self-similarity transformation to the eigenvectors and generalized eigenvectors of Proposition 14. Using $\mathbf{F} = T\mathbf{V}$, the eigenvectors (5.18) become

$$\mathbf{V}_t = \begin{pmatrix} U'_\omega \\ \overline{U}'_\omega \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{V}_g = i \begin{pmatrix} 0 \\ 0 \\ U_\omega \\ -\overline{U}_\omega \end{pmatrix}, \tag{5.26}$$

whereas the generalized eigenvectors (5.20) become

$$\tilde{\mathbf{V}}_t = i\omega x \begin{pmatrix} 0 \\ 0 \\ U_\omega \\ -\overline{U}_\omega \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ U_\omega \\ \overline{U}_\omega \end{pmatrix} \quad \text{and} \quad \tilde{\mathbf{V}}_g = \partial_\omega \begin{pmatrix} U_\omega \\ \overline{U}_\omega \\ 0 \\ 0 \end{pmatrix}. \tag{5.27}$$

Setting $\Phi_V = [\mathbf{V}_t, \mathbf{V}_g, \tilde{\mathbf{V}}_t, \tilde{\mathbf{V}}_g]$ and denoting $\mathcal{S} = T^t \sigma T$, we compute elements of the matrix of skew-symmetric inner products between eigenvectors and generalized eigenvectors:

$$\langle \Phi_V, \mathcal{S}\Phi_V \rangle_{L^2} = \begin{bmatrix} 0 & 0 & \langle \mathbf{V}_t, \mathcal{S}\tilde{\mathbf{V}}_t \rangle_{L^2} & 0 \\ 0 & 0 & 0 & \langle \mathbf{V}_g, \mathcal{S}\tilde{\mathbf{V}}_g \rangle_{L^2} \\ \langle \tilde{\mathbf{V}}_t, \mathcal{S}\mathbf{V}_t \rangle_{L^2} & 0 & 0 & 0 \\ 0 & \langle \tilde{\mathbf{V}}_g, \mathcal{S}\mathbf{V}_g \rangle_{L^2} & 0 & 0 \end{bmatrix}, \tag{5.28}$$

where only nonzero elements are included. Verification of (5.28) is straightforward except for the term

$$\langle \tilde{\mathbf{V}}_t, \mathcal{S}\tilde{\mathbf{V}}_g \rangle_{L^2} = -i\omega \int_{\mathbb{R}} x \partial_\omega |U_\omega|^2 dx - \frac{1}{2} \int_{\mathbb{R}} \left(\bar{U}_\omega \partial_\omega U_\omega - U_\omega \partial_\omega \bar{U}_\omega \right) dx = 0. \tag{5.29}$$

Both integrals in (5.29) are zero because $x\partial_\omega |U_\omega|^2$ and $\mathrm{Im}(\bar{U}_\omega \partial_\omega U_\omega)$ are odd functions of x. As for the nonzero elements, we compute them explicitly from (5.26) and (5.27):

$$\langle \mathbf{V}_t, \mathcal{S}\tilde{\mathbf{V}}_t \rangle_{L^2} = -i\omega \int_{\mathbb{R}} |U_\omega|^2 dx + \frac{1}{2} \int_{\mathbb{R}} \left(\bar{U}_\omega U'_\omega - U_\omega \bar{U}'_\omega \right) dx \tag{5.30}$$

and

$$\langle \mathbf{V}_g, \mathcal{S}\tilde{\mathbf{V}}_g \rangle_{L^2} = -i\frac{d}{d\omega}\int_{\mathbb{R}} |U_\omega|^2 dx. \tag{5.31}$$

Remark 21. *In further analysis, we obtain explicit expressions for (5.30) and (5.31) and show that they are nonzero for every $\omega \in \Omega$. Consequently, the assumption $\langle \mathbf{F}_{t,g}, \sigma\tilde{\mathbf{F}}_{t,g}\rangle_{L^2} \neq 0$ in Proposition 14 is verified for either (5.14) or (5.15) in the spectral stability problem (5.13).*

We shall now proceed separately with the proof of Theorem 5 for the massive Thirring and Gross–Neveu models in two spatial dimensions. Moreover, we derive explicit asymptotic expressions for the eigenvalues mentioned in Theorem 5.

5.3.1 Perturbation theory for the massive Thirring model

The block-diagonalized system (5.22) with (5.23) and (5.24) can be rewritten in the explicit form

$$\begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix} \mathbf{V} + p^2 \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_0 \end{pmatrix} \mathbf{V} = i\lambda \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix} \mathbf{V}, \tag{5.32}$$

where

$$H_+ = \begin{pmatrix} -i\partial_x + \omega + 2|U_\omega|^2 & -1 + U_\omega^2 \\ -1 + \overline{U}_\omega^2 & i\partial_x + \omega + 2|U_\omega|^2 \end{pmatrix}, \quad H_- = \begin{pmatrix} -i\partial_x + \omega & 1 - U_\omega^2 \\ 1 - \overline{U}_\omega^2 & i\partial_x + \omega \end{pmatrix}, \tag{5.33}$$

and the following Pauli matrices are used throughout our work:

$$\sigma_0 = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \sigma_1 = \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right), \quad \sigma_3 = \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right). \tag{5.34}$$

Note that H_+ and H_- are self-adjoint operators in $L^2(\mathbb{R}, \mathbb{C}^2)$ with the domain in $H^1(\mathbb{R}, \mathbb{C}^2)$. The operators $H_\pm$ satisfy the symmetry

$$\sigma_1 H_\pm = \bar{H}_\pm \sigma_1, \tag{5.35}$$

whereas the Pauli matrices satisfy the relation

$$\sigma_1\sigma_1 = \sigma_3\sigma_3 = \sigma_0, \quad \sigma_1\sigma_3 + \sigma_3\sigma_1 = 0, \tag{5.36}$$

Before proving the main result of the perturbation theory for the massive Thirring model in two spatial dimensions, we note the following elementary result.

Proposition 15. *For every $\omega \in (-1, 1)$ and every $p \in \mathbb{R}$, eigenvalues λ of the spectral problem (5.32) are symmetric about the real and imaginary axes in the complex plane.*

Proof. It follows from symmetries (5.35) and (5.36) that if λ is an eigenvalue of the spectral problem (5.32). with the eigenvector $\mathbf{V} = (v_1, v_2, v_3, v_4)^t$, then $\bar{\lambda}$, $-\lambda$, and

$-\bar{\lambda}$ are also eigenvalues of the same problem with the eigenvectors $(\bar{v}_2, \bar{v}_1, \bar{v}_4, \bar{v}_3)^t$, $(v_1, v_2, -v_3, -v_4)^t$, and $(\bar{v}_2, \bar{v}_1, -\bar{v}_4, -\bar{v}_3)^t$. Consequently, we have the following:

- if λ is a simple real nonzero eigenvalue, then the eigenvector $\mathbf{V}$ can be chosen to satisfy the reduction $v_1 = \bar{v}_2$, $v_3 = \bar{v}_4$, whereas $-\lambda$ is also an eigenvalue with the eigenvector $(v_1, v_2, -v_3, -v_4)^t = (\bar{v}_2, \bar{v}_1, -\bar{v}_4, -\bar{v}_3)^t$;
- if λ is a simple purely imaginary nonzero eigenvalue, then the eigenvector $\mathbf{V}$ can be chosen to satisfy the reduction $v_1 = \bar{v}_2$, $v_3 = -\bar{v}_4$, whereas $\bar{\lambda}$ is also an eigenvalue with the eigenvector $(\bar{v}_2, \bar{v}_1, \bar{v}_4, \bar{v}_3)^t = (v_1, v_2, -v_3, -v_4)^t$;
- if a simple eigenvalue λ occurs in the first quadrant, then the symmetry generates eigenvalues in all other quadrants and all four eigenvectors generated by the symmetry are linearly independent.

The symmetry between eigenvalues also applies to multiple nonzero eigenvalues and the corresponding eigenvectors of the associated Jordan blocks. □

For the sake of simplicity, we denote

$$\mathcal{H} = \begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix}, \quad \mathcal{I} = \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_0 \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}.$$

It follows from Proposition 14 and the explicit expressions (5.26) and (5.27) that

$$\mathcal{H}\mathbf{V}_{t,g} = \mathbf{0}, \quad \mathcal{H}\tilde{\mathbf{V}}_{t,g} = i\mathcal{S}\mathbf{V}_{t,g}. \tag{5.37}$$

Setting $\Phi_V = [\mathbf{V}_t, \mathbf{V}_g, \tilde{\mathbf{V}}_t, \tilde{\mathbf{V}}_g]$ as earlier, we note that

$$\langle \Phi_V, \mathcal{I}\Phi_V \rangle_{L^2} = \begin{bmatrix} \|\mathbf{V}_t\|^2_{L^2} & 0 & 0 & 0 \\ 0 & \|\mathbf{V}_g\|^2_{L^2} & 0 & 0 \\ 0 & 0 & \|\tilde{\mathbf{V}}_t\|^2_{L^2} & 0 \\ 0 & 0 & 0 & \|\tilde{\mathbf{V}}_g\|^2_{L^2} \end{bmatrix}, \tag{5.38}$$

where only nonzero terms are included. Again, it is straightforward to verify (5.38) from (5.26) and (5.27), except for the elements

$$\langle \mathbf{V}_t, \tilde{\mathbf{V}}_g \rangle_{L^2} = \int_{\mathbb{R}} \left(\bar{U}'_\omega \partial_\omega U_\omega + U'_\omega \partial_\omega \bar{U}_\omega \right) dx = 0$$

and

$$\langle \mathbf{V}_g, \tilde{\mathbf{V}}_t \rangle_{L^2} = 2\omega \int_{\mathbb{R}} x|U_\omega|^2 dx = 0.$$

These elements are zero because $x|U_\omega|^2$ and $\mathrm{Re}(\bar{U}'_\omega \partial_\omega U_\omega)$ are odd functions of x.

The following result gives the outcome of the perturbation theory associated with the generalized null space of the spectral stability problem (5.32). The result is equivalent to the part of Theorem 5 corresponding to the spectral stability problem (5.13) with (5.14). The asymptotic expressions Λ_r and Λ_i of the real and imaginary eigenvalues λ at the leading order in p versus parameter ω are shown on Fig. 5.1a.

Lemma 28. *For every $\omega \in (-1,1)$, there exists $p_0 > 0$ such that for every p with $0 < |p| < p_0$, the spectral stability problem* (5.32) *admits a pair of real eigenvalues λ with the eigenvectors $\mathbf{V} \in H^1(\mathbb{R})$ such that*

$$\lambda = \pm p\Lambda_r(\omega) + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_t \pm p\Lambda_r(\omega)\tilde{\mathbf{V}}_t + \mathcal{O}_{H^1}(p^2) \quad \text{as} \quad p \to 0, \tag{5.39}$$

where $\Lambda_r = (1-\omega^2)^{-1/4}\|U_\omega'\|_{L^2} > 0$. Simultaneously, it admits a pair of purely imaginary eigenvalues λ with the eigenvector $\mathbf{V} \in H^1(\mathbb{R})$ such that

$$\lambda = \pm ip\Lambda_i(\omega) + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_g \pm ip\Lambda_i(\omega)\tilde{\mathbf{V}}_g + \mathcal{O}_{H^1}(p^2) \quad \text{as} \quad p \to 0, \tag{5.40}$$

where $\Lambda_i = \sqrt{2}(1-\omega^2)^{1/4}\|U_\omega\|_{L^2} > 0$.

Before proving Lemma 28, we give formal computations of the perturbation theory, which recover expansions (5.39) and (5.40) with explicit expressions for $\Lambda_r(\omega)$ and $\Lambda_i(\omega)$. Consider the following formal expansions

$$\lambda = p\Lambda_1 + p^2\Lambda_2 + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_0 + p\Lambda_1\mathbf{V}_1 + p^2\mathbf{V}_2 + \mathcal{O}_{H^1}(p^3), \tag{5.41}$$

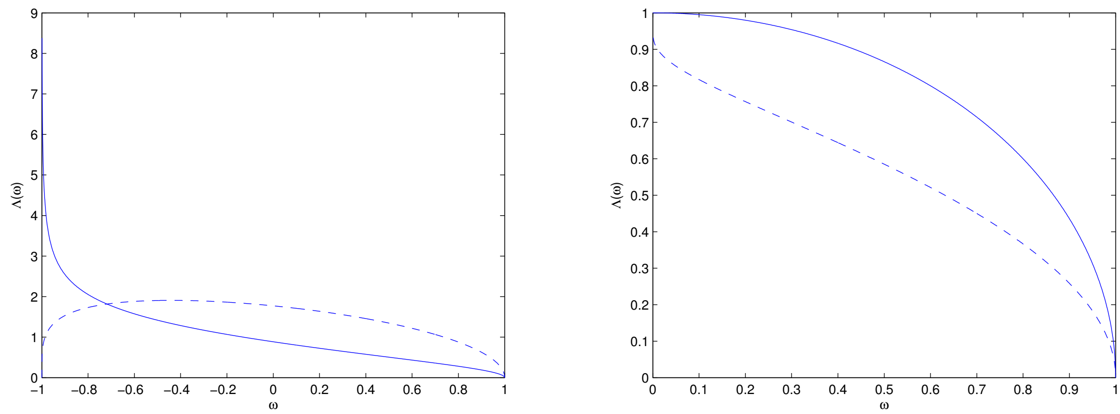
where $\mathbf{V}_0$ is spanned by the eigenvectors (5.26), $\mathbf{V}_1$ is spanned by the generalized eigenvectors (5.27), and $\mathbf{V}_2$ satisfies the linear inhomogeneous equation

$$\mathcal{H}\mathbf{V}_2 = -\mathcal{I}\mathbf{V}_0 + i\Lambda_1^2\mathcal{S}\mathbf{V}_1 + i\Lambda_2\mathcal{S}\mathbf{V}_0. \tag{5.42}$$

By Fredholm's alternative, there exists a solution $\mathbf{V}_2 \in H^1(\mathbb{R})$ of the linear inhomogeneous equation (5.42) if and only if Λ_1 is found from the quadratic equation

$$i\Lambda_1^2\langle \mathbf{W}_0, \mathcal{S}\mathbf{V}_1\rangle_{L^2} = \langle \mathbf{W}_0, \mathbf{V}_0\rangle_{L^2}, \tag{5.43}$$

where $\mathbf{W}_0$ is spanned by the eigenvectors (5.26) independently of $\mathbf{V}_0$. Because of the block diagonalization of the projection matrices in (5.28) and (5.38), the



(a) (b)

Figure 5.1: Asymptotic expressions Λ_r (solid line) and Λ_i (dashed line) versus parameter ω for the massive Thirring (left) and Gross–Neveu (right) models.

2-by-2 matrix eigenvalue problem (5.43) is diagonal and we can proceed separately with precise computations for each eigenvector in $\mathbf{V}_0$.

Selecting $\mathbf{V}_0 = \mathbf{W}_0 = \mathbf{V}_t$ and $\mathbf{V}_1 = \tilde{\mathbf{V}}_t$, we rewrite the solvability condition (5.43) as the following quadratic equation

$$\Lambda_1^2 \int_{\mathbb{R}} \left(\omega |U_\omega|^2 + \frac{i}{2} \left(\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega' \right) \right) dx = 2 \int_{\mathbb{R}} |U_\omega'|^2 dx,$$

where we have used relation (5.30). Substituting the exact expression (5.7), we obtain

$$\int_{\mathbb{R}} \left(\omega |U_\omega|^2 + \frac{i}{2} \left(\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega' \right) \right) dx = 2\sqrt{1-\omega^2} \tag{5.44}$$

and

$$\int_{\mathbb{R}} |U_\omega'|^2 dx = -4\omega\sqrt{1-\omega^2} + 4(1+\omega^2) \arctan\left(\sqrt{\frac{1-\omega}{1+\omega}} \right),$$

which yields the expression $\Lambda_1^2 = (1-\omega^2)^{-1/2} \|U_\omega'\|_{L^2}^2 = \Lambda_r(\omega)^2$.

Selecting now $\mathbf{V}_0 = \mathbf{W}_0 = \mathbf{V}_g$ and $\mathbf{V}_1 = \tilde{\mathbf{V}}_g$, we rewrite the solvability condition (5.43) as the following quadratic equation

$$\Lambda_1^2 \frac{d}{d\omega} \int_{\mathbb{R}} |U_\omega|^2 dx = 2 \int_{\mathbb{R}} |U_\omega|^2 dx,$$

where we have used relation (5.31). Substituting the exact expression (5.7), we obtain

$$\int_{\mathbb{R}} |U_\omega|^2 dx = 4 \arctan\left(\sqrt{\frac{1-\omega}{1+\omega}} \right)$$

and

$$\frac{d}{d\omega} \int_{\mathbb{R}} |U_\omega|^2 dx = -\frac{1}{\sqrt{1-\omega^2}}, \tag{5.45}$$

which yields the expression for $\Lambda_1^2 = -2(1-\omega^2)^{1/2} \|U_\omega\|_{L^2}^2 = -\Lambda_i(\omega)^2$. Note that the nonzero values in (5.44) and (5.45) verify the nonzero values in (5.30) and (5.31), hence the assumption $\langle \mathbf{F}_{t,g}, \sigma \tilde{\mathbf{F}}_{t,g} \rangle_{L^2} \neq 0$ in Proposition 14, according to Remark 21.

We shall now justify the asymptotic expansions (5.39) and (5.40) to give the proof of Lemma 28. Note that Λ_2 in (5.41) is not determined in the linear equation (5.42). Nevertheless, we will show in the proof of Lemma 28 that $\Lambda_2 = 0$, see (5.47), (5.55), and (5.57) below.

I have made a change from proof1 to proof

Proof of Lemma 28. Consider the linearized operator for the spectral problem (5.32):

$$\mathcal{A}_{\lambda,p} = \mathcal{H} + p^2 \mathcal{I} - i\lambda \mathcal{S} : H^1(\mathbb{R}) \to L^2(\mathbb{R}).$$

This operator is self-adjoint if $\lambda \in i\mathbb{R}$ and nonself-adjoint if $\lambda \notin i\mathbb{R}$.

Since $\mathcal{S}\mathcal{S} = \mathcal{I}$, it follows from Proposition 14 and the computations (5.37) that $\mathcal{S}\mathcal{H}$ has the four-dimensional generalized null space $X_0 \subset L^2(\mathbb{R})$ spanned by the vectors in Φ_V. By Propositions 13, the rest of spectrum of $\mathcal{SH}$ is bounded away from zero. By Fredholm's theory, the range of $\mathcal{SH}$ is orthogonal with respect to the generalized null space $Y_0 \subset L^2(\mathbb{R})$ of the adjoint operator $\mathcal{HS}$, which is spanned by the vectors in $\mathcal{S}\Phi_V$.

The inhomogeneous equation $(\mathcal{H}-i\lambda\mathcal{S})g = f$ for $f \in L^2(\mathbb{R})$ is equivalent to the inhomogeneous equation $(\mathcal{SH} - i\lambda)g = \mathcal{S}f$. By Fredholm's alternative, for $\lambda = 0$, a solution $g \in H^1(\mathbb{R})$ exists if and only if $\mathcal{S}f$ is orthogonal to the generalized kernel of $\mathcal{HS}$, which means that $\mathcal{S}f \in Y_0^\perp$ or equivalently, $f \in X_0^\perp$. For $\lambda \neq 0$ but small, it is natural to define the solution $g \in H^1(\mathbb{R})$ uniquely by the constraint $g \in Y_0^\perp$.

Consequently, there is $\lambda_0 > 0$ sufficiently small such that $\mathcal{A}_{\lambda,0}$ with $|\lambda| < \lambda_0$ is invertible on $X_0^\perp$ with a bounded inverse in $Y_0^\perp$. Since $p^2\mathcal{I}$ is a bounded self-adjoint perturbation to $\mathcal{H}$, there exist positive constants λ_0, p_0, and C_0 such that for all $|\lambda| < \lambda_0$, $|p| < p_0$, and all $\mathbf{f} \in X_0^\perp \subset L^2(\mathbb{R})$, there exists a unique $\mathcal{A}_{\lambda,p}^{-1}\mathbf{f} \in Y_0^\perp$ satisfying

$$\|\mathcal{A}_{\lambda,p}^{-1}\mathbf{f}\|_{L^2} \leq C_0\|\mathbf{f}\|_{L^2}. \tag{5.46}$$

Moreover, $\mathcal{A}_{\lambda,p}^{-1}\mathbf{f} \in H^1(\mathbb{R})$.

Let us now use the method of the Lyapunov–Schmidt reduction. We apply the partition of Φ_V as $\Phi_V^{(0)} = [\mathbf{V}_t, \mathbf{V}_g]$ and $\Phi_V^{(1)} = [\tilde{\mathbf{V}}_t, \tilde{\mathbf{V}}_g]$. Given the computations above, we consider the decomposition of the solution of the spectral problem (5.32) in the form

$$\begin{cases} \lambda = p(\Lambda + \mu_p), \\ \mathbf{V} = \Phi_V^{(0)}\vec{\alpha}_p + p\Phi_V^{(1)}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p) + \mathbf{V}_p, \end{cases} \tag{5.47}$$

where $\Lambda \in \mathbb{C}$ is p-independent, whereas $\mu_p \in \mathbb{C}$, $\vec{\alpha}_p \in \mathbb{C}^2$, $\vec{\gamma}_p \in \mathbb{C}^2$, and $\mathbf{V}_p \in H^1(\mathbb{R})$ may depend on p. For uniqueness of the decomposition, we use the Fredholm theory and require that the correction term $\mathbf{V}_p$ satisfy the orthogonality conditions:

$$\langle \Phi_V, \mathcal{S}\mathbf{V}_p \rangle_{L^2} = 0, \tag{5.48}$$

which ensures that $\mathbf{V}_p \in H^1(\mathbb{R}) \cap Y_0^\perp$. Substituting expansions (5.47) into the spectral problem (5.32), we obtain

$$\begin{aligned} &\left(\mathcal{H} + p^2\mathcal{I} - ip(\Lambda + \mu_p)\mathcal{S}\right)\mathbf{V}_p + p^2\left(\Phi_V^{(0)}\vec{\alpha}_p + p\Phi_V^{(1)}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p)\right) \\ &\quad = ip^2(\Lambda + \mu_p)\mathcal{S}\Phi_V^{(1)}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p) - ip\mathcal{S}\Phi_V^{(0)}\vec{\gamma}_p. \end{aligned} \tag{5.49}$$

In order to solve equation (5.49) for $\mathbf{V}_p$ in $H^1(\mathbb{R}) \cap Y_0^\perp$, we project the equation to $X_0^\perp$. It makes sense to do so separately for $\Phi_V^{(0)}$ and $\Phi_V^{(1)}$. Using the projection matrices (5.28) and (5.38) as well as the orthogonality conditions (5.48), we obtain

$$\begin{aligned} &p^2\langle \Phi_V^{(0)}, \mathbf{V}_p \rangle_{L^2} + p^2\langle \Phi_V^{(0)}, \Phi_V^{(0)} \rangle_{L^2}\vec{\alpha}_p \\ &\quad = ip^2(\Lambda + \mu_p)\langle \Phi_V^{(0)}, \mathcal{S}\Phi_V^{(1)} \rangle_{L^2}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p) \end{aligned} \tag{5.50}$$

and

$$p^2\langle \Phi_V^{(1)}, \mathbf{V}_p \rangle_{L^2} + p^3\langle \Phi_V^{(1)}, \Phi_V^{(1)} \rangle_{L^2}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p) = -ip\langle \Phi_V^{(1)}, \mathcal{S}\Phi_V^{(0)} \rangle_{L^2}\vec{\gamma}_p. \tag{5.51}$$

Under the constraints (5.50) and (5.51), the right-hand-side of equation (5.49) belongs to $X_0^\perp$. The resolvent estimate (5.46) implies that the operator $A_{\lambda,p}$ can be inverted with a bounded inverse in $Y_0^\perp$. By the inverse function theorem, there are positive numbers $p_1 \leq p_0$, μ_1, and C_1 such that for every $|p| < p_1$ and $|\mu_p| < \mu_1$, there exists a unique solution of equation (5.49) for $\mathbf{V}_p$ in $H^1(\mathbb{R}) \cap Y_0^\perp$ satisfying the estimate

$$\|\mathbf{V}_p\|_{L^2} \leq C_1 \left(p^2\|\vec{\alpha}_p\| + |p|\|\vec{\gamma}_p\|\right). \tag{5.52}$$

Substituting this solution to the projection equations (5.50) and (5.51), we shall be looking for values of Λ, μ_p, $\vec{\alpha}_p$, and $\vec{\gamma}_p$ for $|p| < p_1$ sufficiently small. Using the estimate (5.52), we realize that the leading order of equation (5.50) is

$$\langle \Phi_V^{(0)}, \Phi_V^{(0)} \rangle_{L^2} \vec{c} = i\Lambda^2 \langle \Phi_V^{(0)}, \mathcal{S}\Phi_V^{(1)} \rangle_{L^2} \vec{c}, \quad \vec{c} \in \mathbb{C}^2. \tag{5.53}$$

This equation is diagonal and admits two eigenvalues for Λ^2 given by $\Lambda_r(\omega)^2$ and $-\Lambda_i(\omega)^2$, so that

$$\|\mathbf{V}_t\|_{L^2}^2 = i\Lambda_r(\omega)^2 \langle \mathbf{V}_t, \mathcal{S}\tilde{\mathbf{V}}_t \rangle_{L^2}, \quad \|\mathbf{V}_g\|_{L^2}^2 = -i\Lambda_i(\omega)^2 \langle \mathbf{V}_g, \mathcal{S}\tilde{\mathbf{V}}_g \rangle_{L^2}.$$

Choosing Λ^2 being equal to one of the two eigenvalues (which are distinct), we obtain a rank-one coefficient matrix for equation (5.50) at the leading order. In what follows, we omit the argument ω from Λ_r and Λ_i

For simplicity, let us choose $\Lambda^2 = \Lambda_r^2$ (the other case is considered similarly) and represent $\vec{\alpha}_p = (\alpha_p, \beta_p)^t$ and $\vec{\gamma}_p = (\gamma_p, \delta_p)^t$. In this case, α_p can be normalized to unity independently of p, after which equation (5.50) divided by p^2 is rewritten in the following explicit form

$$\begin{bmatrix} \|\mathbf{V}_t\|_{L^2}^2 & 0 \\ 0 & \|\mathbf{V}_g\|_{L^2}^2 \end{bmatrix} \begin{bmatrix} \left(1 + \frac{\mu_p}{\Lambda_r}\right)^2 - 1 + \frac{\Lambda_r + \mu_p}{\Lambda_r^2}\gamma_p \\ -\frac{\Lambda_r^2}{\Lambda_i^2}\left(1 + \frac{\mu_p}{\Lambda_r}\right)^2 \beta_p - \beta_p - \frac{\Lambda_r+\mu_p}{\Lambda_i^2}\delta_p \end{bmatrix} = \langle \Phi_V^{(0)}, \mathbf{V}_p \rangle_{L^2}. \tag{5.54}$$

We invoke the implicit function theorem for vector functions. It follows from the estimate (5.52) that there are positive numbers $p_2 \leq p_1$ and C_2 such that for every $|p| < p_2$, there exists a unique solution of equation (5.54) for μ_p and β_p satisfying the estimate

$$|\mu_p| + |\beta_p| \leq C_2 \left(\|\vec{\gamma}_p\| + \|\mathbf{V}_p\|_{L^2}\right) \leq C_2 \left(\|\vec{\gamma}_p\| + p^2\right), \tag{5.55}$$

where the last inequality with a modified value of constant C_2 is due to the estimate (5.52).

Finally, we divide equation (5.51) by p and rewrite it in the form

$$-i\langle \Phi_V^{(1)}, \mathcal{S}\Phi_V^{(0)} \rangle_{L^2} \vec{\gamma}_p = p\langle \Phi_V^{(1)}, \mathbf{V}_p \rangle_{L^2} + p^2 \langle \Phi_V^{(1)}, \Phi_V^{(1)} \rangle_{L^2}((\Lambda + \mu_p)\vec{\alpha}_p + \vec{\gamma}_p). \tag{5.56}$$

Thanks to the estimates (5.52) and (5.55), equation (5.56) can be solved for $\vec{\gamma}_p$ by the implicit function theorem, if p is sufficiently small and $\vec{V}_p$, μ_p, and $\vec{\alpha}_p$ are substituted from solutions of the previous equations. As a result, there are positive numbers $p_3 \leq p_2$ and C_3 such that for every $|p| < p_3$, there exists a unique solution of equation (5.56) for $\vec{\gamma}_p$ satisfying the estimate

$$\|\vec{\gamma}_p\| \leq C_3 \left(p^2 + p\|\mathbf{V}_p\|_{L^2}\right) \leq C_3 p^2, \tag{5.57}$$

where the last inequality with a modified value of constant C_3 is due to the estimate (5.52).

Decomposition (5.47) and estimates (5.52), (5.55), and (5.57) justify the asymptotic expansion (5.39). It remains to prove that the eigenvalue $\lambda = p(\Lambda_r + \mu_p)$ is purely real. Since Λ_r is real, the result holds if μ_p is real. Assume that μ_p has a nonzero imaginary part. By Proposition 15, there exists another distinct eigenvalue of the spectral problem (5.32) given by $\lambda = (p\Lambda_r + \bar{\mu}_p)$ such that $\bar{\mu}_p = \mathcal{O}(p^2)$ as $p \to 0$. However, the existence of this distinct eigenvalue contradicts the uniqueness of constructing of μ_p and all terms in the decomposition (5.47). Therefore, $\bar{\mu}_p = \mu_p$, so that $\lambda = p(\Lambda_r + \mu_p)$ is real.

The asymptotic expansion (5.40) is proved similarly with the normalization $\beta_p = 1$ and the choice $\Lambda^2 = -\Lambda_i^2$ among eigenvalues of the reduced eigenvalue problem (5.53). □

5.3.2 Perturbation theory for the massive Gross–Neveu model

The block-diagonalized system (5.22) with (5.23) and (5.25) can be rewritten in the explicit form

$$\begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix} \mathbf{V} + ip \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} \mathbf{V} = i\lambda \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix} \mathbf{V}, \tag{5.58}$$

where σ_1 and σ_3 are the Pauli matrices, whereas

$$H_+ = \begin{pmatrix} -i\partial_x + \omega + 2|U_\omega|^2 & -1 + U_\omega^2 + 3\overline{U}_\omega^2 \\ -1 + \overline{U}_\omega^2 + 3U_\omega^2 & i\partial_x + \omega + 2|U_\omega|^2 \end{pmatrix},$$

$$H_- = \begin{pmatrix} -i\partial_x + \omega & 1 - U_\omega^2 - \overline{U}_\omega^2 \\ 1 - U_\omega^2 - \overline{U}_\omega^2 & i\partial_x + \omega \end{pmatrix}.$$

We note again the symmetry relation (5.35), which applies to the Dirac operators $H_\pm$ for the massive Gross–Neveu model as well. From this symmetry, we derive the result, which is similar to Proposition 15 and is proved directly.

Proposition 16. *For every $\omega \in (0, 1)$, if λ is an eigenvalue of the spectral problem (5.58) with $p \in \mathbb{R}$ and the eigenvector $\mathbf{V} = (v_1, v_2, v_3, v_4)^t$, then $-\bar{\lambda}$ is also an eigenvalue of the same problem with the eigenvector $(\bar{v}_2, \bar{v}_1, -\bar{v}_4, -\bar{v}_3)^t$, whereas $\bar{\lambda}$ and $-\lambda$ are eigenvalues of the spectral problem (5.58) with $-p \in \mathbb{R}$ and the eigenvectors $(\bar{v}_2, \bar{v}_1, \bar{v}_4, \bar{v}_3)^t$ and $(v_1, v_2, -v_3, -v_4)^t$, respectively. Consequently, for every $p \in \mathbb{R}$, eigenvalues λ of the spectral problem (5.58) are symmetric about the imaginary axis.*

For the sake of simplicity, we use again the notations

$$\mathcal{H} = \begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix}, \quad \mathcal{P} = i \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}.$$

The relations (5.37) hold true for this case as well. Besides the eigenvectors (5.26) and the generalized eigenvectors (5.27), we need solutions of the linear inhomogeneous equations

$$\mathcal{H}\mathbf{V} = -\mathcal{P}\mathbf{V}_{t,g}, \tag{5.59}$$

which are given by

$$\check{\mathbf{V}}_t = -\frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ \overline{U}_\omega \\ -U_\omega \end{pmatrix} \quad \text{and} \quad \check{\mathbf{V}}_g = -\frac{1}{2\omega} \begin{pmatrix} \overline{U}_\omega \\ -U_\omega \\ 0 \\ 0 \end{pmatrix}. \tag{5.60}$$

The existence of these explicit expressions is checked by elementary substitution.

We apply again the partition of Φ_V as $\Phi_V^{(0)} = [\mathbf{V}_t, \mathbf{V}_g]$ and $\Phi_V^{(1)} = [\tilde{\mathbf{V}}_t, \tilde{\mathbf{V}}_g]$. In addition, we augment the matrix Φ_V with $\Phi_V^{(2)} = [\check{\mathbf{V}}_t, \check{\mathbf{V}}_g]$ and compute the missing entries in the projection matrices:

$$\langle \Phi_V^{(0)}, \mathcal{S}\Phi_V^{(2)} \rangle_{L^2} = \langle \Phi_V^{(2)}, \mathcal{S}\Phi_V^{(2)} \rangle_{L^2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{5.61}$$

and

$$\langle \Phi_V^{(1)}, \mathcal{S}\Phi_V^{(2)} \rangle_{L^2} = \begin{bmatrix} 0 & 0 \\ \langle \tilde{\mathbf{V}}_g, \mathcal{S}\check{\mathbf{V}}_t \rangle_{L^2} & 0 \end{bmatrix}. \tag{5.62}$$

Indeed, in addition to the matrix elements, which are trivially zero, we check that

$$\langle \mathbf{V}_g, \mathcal{S}\check{\mathbf{V}}_g \rangle_{L^2} = \frac{i}{2\omega} \int_{\mathbb{R}} \left(\bar{U}_\omega^2 - U_\omega^2 \right) dx = 0, \tag{5.63}$$

because $\mathrm{Im}(U_\omega^2)$ is an odd function of x, and

$$\langle \tilde{\mathbf{V}}_t, \mathcal{S}\check{\mathbf{V}}_g \rangle_{L^2} = \frac{i}{2} \int_{\mathbb{R}} x \left(\bar{U}_\omega^2 - U_\omega^2 \right) dx + \frac{1}{4\omega} \int_{\mathbb{R}} (\bar{U}_\omega^2 + U_\omega^2) dx = 0, \tag{5.64}$$

where the exact expression (5.11) is used. On the other hand, we have

$$\begin{aligned} \langle \tilde{\mathbf{V}}_g, \mathcal{S}\check{\mathbf{V}}_t \rangle_{L^2} &= -\frac{1}{4} \frac{d}{d\omega} \int_{\mathbb{R}} \left(\bar{U}_\omega^2 + U_\omega^2 \right) dx \\ &= -\frac{1}{2} \frac{d}{d\omega} \log \left(\frac{1 + \omega + \sqrt{1 - \omega^2}}{1 + \omega - \sqrt{1 - \omega^2}} \right) = \frac{1}{2\omega\sqrt{1 - \omega^2}}, \end{aligned} \tag{5.65}$$

which is nonzero.

Similarly, we compute the zero projection matrices

$$\langle \Phi_V^{(0)}, \mathcal{P}\Phi_V^{(0)} \rangle_{L^2} = \langle \Phi_V^{(0)}, \mathcal{P}\Phi_V^{(1)} \rangle_{L^2} = \langle \Phi_V^{(1)}, \mathcal{P}\Phi_V^{(2)} \rangle_{L^2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{5.66}$$

and the nonzero projection matrices

$$\langle \Phi_V^{(1)}, \mathcal{P}\Phi_V^{(1)} \rangle_{L^2} = \begin{bmatrix} 0 & \langle \tilde{\mathbf{V}}_t, \mathcal{P}\tilde{\mathbf{V}}_g \rangle_{L^2} \\ \langle \tilde{\mathbf{V}}_g, \mathcal{P}\tilde{\mathbf{V}}_t \rangle_{L^2} & 0 \end{bmatrix}, \tag{5.67}$$

$$\langle \Phi_V^{(0)}, \mathcal{P}\Phi_V^{(2)} \rangle_{L^2} = \begin{bmatrix} \langle \mathbf{V}_t, \mathcal{P}\check{\mathbf{V}}_t \rangle_{L^2} & 0 \\ 0 & \langle \mathbf{V}_g, \mathcal{P}\check{\mathbf{V}}_g \rangle_{L^2} \end{bmatrix}, \tag{5.68}$$

and

$$\langle \Phi_V^{(2)}, \mathcal{P}\Phi_V^{(2)} \rangle_{L^2} = \begin{bmatrix} 0 & \langle \check{\mathbf{V}}_t, \mathcal{P}\check{\mathbf{V}}_g \rangle_{L^2} \\ \langle \check{\mathbf{V}}_g, \mathcal{P}\check{\mathbf{V}}_t \rangle_{L^2} & 0 \end{bmatrix}. \tag{5.69}$$

Indeed, the first matrix in (5.66) is zero because the Fredholm conditions for the inhomogeneous linear systems (5.59) are satisfied. The second matrix in (5.66) is zero because

$$\langle \mathbf{V}_t, \mathcal{P}\tilde{\mathbf{V}}_t \rangle_{L^2} = \frac{\omega}{2} \int_{\mathbb{R}} \left(U_\omega^2 - \bar{U}_\omega^2 \right) dx = 0 \tag{5.70}$$

and

$$\langle \mathbf{V}_g, \mathcal{P}\tilde{\mathbf{V}}_g \rangle_{L^2} = \frac{1}{2} \frac{d}{d\omega} \int_{\mathbb{R}} \left(U_\omega^2 - \bar{U}_\omega^2 \right) dx = 0. \tag{5.71}$$

The third matrix in (5.66) is zero because

$$\langle \tilde{\mathbf{V}}_t, \mathcal{P}\check{\mathbf{V}}_g \rangle_{L^2} = -\int_{\mathbb{R}} x |U_\omega|^2 dx = 0 \tag{5.72}$$

and

$$\langle \tilde{\mathbf{V}}_g, \mathcal{P}\check{\mathbf{V}}_t \rangle_{L^2} = \frac{i}{2} \int_{\mathbb{R}} \left(U_\omega \partial_\omega \bar{U}_\omega - \bar{U}_\omega \partial_\omega U_\omega \right) dx = 0. \tag{5.73}$$

For the projection matrices (5.67), (5.68), and (5.69), we compute the nonzero elements explicitly:

$$\begin{aligned}
\langle \tilde{\mathbf{V}}_t, \mathcal{P}\tilde{\mathbf{V}}_g \rangle_{L^2} &= \frac{i}{4} \frac{d}{d\omega} \int_{\mathbb{R}} \left(U_\omega^2 + \bar{U}_\omega^2 \right) dx + \frac{\omega}{2} \frac{d}{d\omega} \int_{\mathbb{R}} x \left(U_\omega^2 - \bar{U}_\omega^2 \right) dx, && (5.74) \\
\langle \mathbf{V}_t, \mathcal{P}\check{\mathbf{V}}_t \rangle_{L^2} &= \frac{i}{2} \int_{\mathbb{R}} \left(U_\omega \bar{U}_\omega' - \bar{U}_\omega U_\omega' \right) dx, && (5.75) \\
\langle \mathbf{V}_g, \mathcal{P}\check{\mathbf{V}}_g \rangle_{L^2} &= -\frac{1}{\omega} \int_{\mathbb{R}} |U_\omega|^2 dx, && (5.76) \\
\langle \check{\mathbf{V}}_t, \mathcal{P}\check{\mathbf{V}}_g \rangle_{L^2} &= \frac{i}{4\omega} \int_{\mathbb{R}} \left(U_\omega^2 + \bar{U}_\omega^2 \right) dx. && (5.77)
\end{aligned}$$

The following result gives the outcome of the perturbation theory associated with the generalized null space of the spectral stability problem (5.58). The result is equivalent to the part of Theorem 5 corresponding to the spectral stability problem (5.13) with (5.15). The asymptotic expressions Λ_r and Λ_i for the corresponding eigenvalues λ at the leading order in p versus parameter ω are shown on Fig. 5.1b.

Lemma 29. *For every $\omega \in (0,1)$, there exists $p_0 > 0$ such that for every p with $0 < |p| < p_0$, the spectral stability problem* (5.58) *admits a pair of purely imaginary eigenvalues λ with the eigenvectors $\mathbf{V} \in H^1(\mathbb{R})$ such that*

$$\lambda = \pm ip\Lambda_i(\omega) + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_t \pm ip\Lambda_i(\omega)\tilde{\mathbf{V}}_t + p\check{\mathbf{V}}_t + p\beta\mathbf{V}_g + \mathcal{O}_{H^1}(p^2) \quad \text{as} \quad p \to 0, \tag{5.78}$$

where $\Lambda_i(\omega) = \sqrt{\frac{I(\omega)}{1+I(\omega)}} > 0$ with $I(\omega) > 0$ given by the explicit expression (5.87) below and β is uniquely defined in (5.94) *below.*

Simultaneously, the spectral stability problem (5.58) *admits a pair of eigenvalues λ with* $\mathrm{Re}(\lambda) \neq 0$ *symmetric about the imaginary axis, and the eigenvector $\mathbf{V} \in H^1(\mathbb{R})$ such that*

$$\lambda = \pm p\Lambda_r(\omega) + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_g \pm p\Lambda_r(\omega)\tilde{\mathbf{V}}_g + p\check{\mathbf{V}}_g + p\alpha\mathbf{V}_t + \mathcal{O}_{H^1}(p^2) \quad \text{as} \quad p \to 0, \tag{5.79}$$

where $\Lambda_r = (1-\omega^2)^{1/2} > 0$ and α is uniquely defined in (5.93) *below.*

We proceed with formal expansions, which are similar to the expansions (5.41). However, because the $\mathcal{O}(p)$ terms appear explicitly in the spectral stability problem (5.58), we introduce the modified expansions as follows,

$$\lambda = p\Lambda_1 + p^2\Lambda_2 + \mathcal{O}(p^3), \quad \mathbf{V} = \mathbf{V}_0 + p(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0') + p^2\mathbf{V}_2 + \mathcal{O}_{H^1}(p^3), \tag{5.80}$$

where $\mathbf{V}_0$ and $\mathbf{V}_0'$ are spanned independently by the eigenvectors (5.26), $\mathbf{V}_1$ is spanned by the generalized eigenvectors (5.27), $\check{\mathbf{V}}_1$ is spanned by the vectors (5.60), and $\mathbf{V}_2$ satisfies the linear inhomogeneous equation

$$\mathcal{H}\mathbf{V}_2 = (i\Lambda_1\mathcal{S} - \mathcal{P})(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0') + i\Lambda_2\mathcal{S}\mathbf{V}_0. \tag{5.81}$$

By Fredholm's alternative, there exists a solution $\mathbf{V}_2 \in H^1(\mathbb{R})$ of the linear inhomogeneous equation (5.81) if and only if Λ_1 is found from the quadratic equation

$$\langle \mathbf{W}_0, (i\Lambda_1\mathcal{S} - \mathcal{P})(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0')\rangle_{L^2} = 0, \tag{5.82}$$

where $\mathbf{W}_0$ is again spanned by the eigenvectors (5.26) independently of $\mathbf{V}_0$. Similar to (5.43), the matrix eigenvalue problem (5.82) is diagonal with respect to the translational and gauge symmetries. As a result, subsequent computations can be constructed independently for the two corresponding eigenvectors.

Selecting $\mathbf{V}_0 = \mathbf{W}_0 = \mathbf{V}_g$, $\mathbf{V}_1 = \tilde{\mathbf{V}}_g$, $\check{\mathbf{V}}_1 = \check{\mathbf{V}}_g$, and $\mathbf{V}_0' = \alpha\mathbf{V}_t$, we use (5.28), (5.31), (5.61), (5.66), (5.68), and (5.76) in the solvability condition (5.82) and obtain the quadratic equation for Λ_1 in the explicit form

$$\Lambda_1^2 \frac{d}{d\omega}\int_{\mathbb{R}} |U_\omega|^2 dx + \frac{1}{\omega}\int_{\mathbb{R}} |U_\omega|^2 dx = 0. \tag{5.83}$$

Using the explicit expression (5.11), we obtain

$$\int_{\mathbb{R}} |U_\omega|^2 dx = \frac{\sqrt{1-\omega^2}}{\omega}, \quad \frac{d}{d\omega}\int_{\mathbb{R}} |U_\omega|^2 dx = -\frac{1}{\omega^2\sqrt{1-\omega^2}}, \tag{5.84}$$

which yield $\Lambda_1^2 = 1-\omega^2 = \Lambda_r(\omega)^2$. Correction terms Λ_2 and α are not determined up to this order of the asymptotic expansion.

Selecting now $\mathbf{V}_0 = \mathbf{W}_0 = \mathbf{V}_t$, $\mathbf{V}_1 = \tilde{\mathbf{V}}_t$, $\check{\mathbf{V}}_1 = \check{\mathbf{V}}_t$, and $\mathbf{V}_0' = \beta \mathbf{V}_g$, we use (5.28), (5.30), (5.61), (5.66), (5.68), and (5.75) in the solvability condition (5.82) and obtain the quadratic equation for Λ_1 in the explicit form

$$\Lambda_1^2 \int_{\mathbb{R}} \left[\omega|U_\omega|^2 + \frac{i}{2}(\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega')\right] dx + \frac{i}{2}\int_{\mathbb{R}} (\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega') dx = 0. \tag{5.85}$$

Expressing

$$\frac{i}{2}\int_{\mathbb{R}} (\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega') dx = \int_{\mathbb{R}} \frac{(1-\omega^2)^2}{(1+\omega\cosh(2\mu x))^2} dx = \sqrt{1-\omega^2}\, I(\omega),$$

and

$$\int_{\mathbb{R}} \left[\omega|U_\omega|^2 + \frac{i}{2}(\bar{U}_\omega U_\omega' - U_\omega \bar{U}_\omega')\right] dx = \sqrt{1-\omega^2}\left[1 + I(\omega)\right], \tag{5.86}$$

where

$$I(\omega) := (1-\omega^2)\int_0^\infty \frac{dz}{(1+\omega\cosh(z))^2} = 1 - \frac{1}{\sqrt{1-\omega^2}} \log\left(\frac{1-\sqrt{1-\omega^2}}{\omega}\right) > 0, \tag{5.87}$$

we obtain $\Lambda_1^2 = -\frac{I(\omega)}{1+I(\omega)} = -\Lambda_i(\omega)^2$. Again, correction terms Λ_2 and β are not determined up to this order of the asymptotic expansion.

Note again that the nonzero values in (5.84) and (5.86) verify the nonzero values in (5.30) and (5.31), hence the assumption $\langle \mathbf{F}_{t,g}, \sigma \tilde{\mathbf{F}}_{t,g}\rangle_{L^2} \neq 0$ in Proposition 14, according to Remark 21.

Justification of the formal expansion (5.80) and the proof of Lemma 29 is achieved by exactly the same argument as in the proof of Lemma 28. The proof relies on the resolvent estimate (5.46), which is valid for the massive Gross–Neveu model, because by Propositions 13 and 14, the zero eigenvalue of the operator $\mathcal{SH}$ (which has algebraic multiplicity four) is isolated from the rest of the spectrum.

Persistence of eigenvalues is proved with the symmetry in Proposition 16. If an eigenvalue is expressed as $\lambda = p(i\Lambda_i(\omega)+\mu_p)$ with unique $\mu_p = \mathcal{O}(p)$ and $\Lambda_i(\omega) > 0$, then nonzero real part of μ_p would contradict the symmetry of eigenvalues about the imaginary axis. Therefore, $\text{Re}(\mu_p) = 0$ and the eigenvalues in the expansion (5.78) remain on the imaginary axis. On the other hand, if another eigenvalue is expressed as $\lambda = p(\Lambda_r(\omega) + \mu_p)$ with unique $\mu_p = \mathcal{O}(p)$ and $\Lambda_r(\omega) > 0$, then μ_p may have in general a nonzero imaginary part, as it does not contradict the symmetry of Proposition 16 for a fixed $p \neq 0$. This is why the statement of Lemma 29 does not guarantee that the corresponding eigenvalues in the expansion (5.79) are purely real.

In the end of this section, we will show that $\mu_p = \mathcal{O}(p^2)$, which justifies the $\mathcal{O}(p^3)$ bound for the eigenvalues in the asymptotic expansions (5.78) and (5.79). In this procedure, we will uniquely determine the parameters β and α in the same asymptotic expansions. Extending the expansion (5.80) to $p^3\Lambda_3$ and $p^3\mathbf{V}_3$ terms, we obtain the linear inhomogeneous equation

$$\mathcal{H}\mathbf{V}_3 = (i\Lambda_1\mathcal{S} - \mathcal{P})\mathbf{V}_2 + i\Lambda_2\mathcal{S}(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0') + i\Lambda_3\mathcal{S}\mathbf{V}_0. \tag{5.88}$$

The Fredholm solvability condition

$$\langle \mathbf{W}_0, (i\Lambda_1\mathcal{S} - \mathcal{P})\mathbf{V}_2 + i\Lambda_2\mathcal{S}(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0')\rangle_{L^2} = 0 \tag{5.89}$$

determines the correction terms Λ_2, β, and α uniquely. Indeed, using (5.28) and (5.61), we rewrite the solvability condition (5.89) in the form

$$\begin{aligned} i\langle \mathbf{W}_0, \mathcal{S}\mathbf{V}_1\rangle_{L^2}\Lambda_2\Lambda_1 &= -\langle \mathbf{W}_0, (i\Lambda_1\mathcal{S} - \mathcal{P})\mathbf{V}_2\rangle_{L^2} \\ &= -\langle (-i\bar{\Lambda}_1\mathcal{S} - \mathcal{P})\mathbf{W}_0, \mathbf{V}_2\rangle_{L^2} \\ &= -\langle \mathcal{H}(-\bar{\Lambda}_1\mathbf{W}_1 + \check{\mathbf{W}}_1), \mathbf{V}_2\rangle_{L^2} \\ &= -\langle (-\bar{\Lambda}_1\mathbf{W}_1 + \check{\mathbf{W}}_1), \mathcal{H}\mathbf{V}_2\rangle_{L^2} \\ &= -\langle (-\bar{\Lambda}_1\mathbf{W}_1 + \check{\mathbf{W}}_1), i\Lambda_2\mathcal{S}\mathbf{V}_0 \\ &\qquad + (i\Lambda_1\mathcal{S} - \mathcal{P})(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0')\rangle_{L^2}, \end{aligned}$$

where we have used the linear inhomogeneous equation (5.81) and have introduced $\mathbf{W}_1$ and $\check{\mathbf{W}}_1$ from solutions of the inhomogeneous equations $\mathcal{H}\mathbf{W}_1 = i\mathcal{S}\mathbf{W}_0$ and $\mathcal{H}\check{\mathbf{W}}_1 = -\mathcal{P}\mathbf{W}_0$. Using

$$\langle \mathbf{W}_1, i\mathcal{S}\mathbf{V}_0\rangle_{L^2} = \langle \mathbf{W}_1, \mathcal{H}\mathbf{V}_1\rangle_{L^2} = \langle \mathcal{H}\mathbf{W}_1, \mathbf{V}_1\rangle_{L^2} = \langle i\mathcal{S}\mathbf{W}_0, \mathbf{V}_1\rangle_{L^2} = -i\langle \mathbf{W}_0, \mathcal{S}\mathbf{V}_1\rangle_{L^2}$$

and

$$\begin{aligned} \langle \check{\mathbf{W}}_1, i\mathcal{S}\mathbf{V}_0\rangle_{L^2} = \langle \check{\mathbf{W}}_1, \mathcal{H}\mathbf{V}_1\rangle_{L^2} &= \langle \mathcal{H}\check{\mathbf{W}}_1, \mathbf{V}_1\rangle_{L^2} \\ &= -\langle \mathcal{P}\mathbf{W}_0, \mathbf{V}_1\rangle_{L^2} = -\langle \mathbf{W}_0, \mathcal{P}\mathbf{V}_1\rangle_{L^2} = 0, \end{aligned}$$

where the last equality is due to (5.66), we rewrite the solvability equation in the form

$$2i\langle \mathbf{W}_0, \mathcal{S}\mathbf{V}_1\rangle_{L^2}\Lambda_2\Lambda_1 = -\langle (-\bar{\Lambda}_1\mathbf{W}_1 + \check{\mathbf{W}}_1), (i\Lambda_1\mathcal{S} - \mathcal{P})(\Lambda_1\mathbf{V}_1 + \check{\mathbf{V}}_1 + \mathbf{V}_0')\rangle_{L^2}. \tag{5.90}$$

Removing zero entries by using (5.28), (5.61), and (5.66), we rewrite equation (5.90) in the form

$$\begin{aligned} 2i\langle \mathbf{W}_0, \mathcal{S}\mathbf{V}_1\rangle_{L^2}\Lambda_2\Lambda_1 = \Lambda_1^2 \big(& i\langle \mathbf{W}_1, \mathcal{S}\mathbf{V}_0'\rangle_{L^2} + i\langle \mathbf{W}_1, \mathcal{S}\check{\mathbf{V}}_1\rangle_{L^2} - i\langle \check{\mathbf{W}}_1, \mathcal{S}\mathbf{V}_1\rangle_{L^2} \\ & -\langle \mathbf{W}_1, \mathcal{P}\mathbf{V}_1\rangle_{L^2} + \langle \check{\mathbf{W}}_1, \mathcal{P}\check{\mathbf{V}}_1\rangle_{L^2} + \langle \check{\mathbf{W}}_1, \mathcal{P}\mathbf{V}_0'\rangle_{L^2}. \end{aligned} \tag{5.91}$$

We shall now write equation (5.91) explicitly as the 2-by-2 matrix equation by using $\mathbf{V}_0 = \mathbf{W}_0 = \Phi_V^{(0)}$, $\mathbf{V}_1 = \mathbf{W}_1 = \Phi_V^{(1)}$, $\check{\mathbf{V}}_1 = \check{\mathbf{W}}_1 = \Phi_V^{(2)}$, and

$$\mathbf{V}_0' = \Phi_V^{(0)} \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix} = [\beta \mathbf{V}_g, \alpha \mathbf{V}_t].$$

Using (5.28), (5.62), (5.67), (5.68), and (5.69), we rewrite equation (5.91) in the matrix form

$$\begin{aligned} 2i \begin{bmatrix} \langle \mathbf{V}_t, \mathcal{S}\tilde{\mathbf{V}}_t\rangle_{L^2} & 0 \\ 0 & \langle \mathbf{V}_g, \mathcal{S}\tilde{\mathbf{V}}_g\rangle_{L^2} \end{bmatrix} \Lambda_2\Lambda_1 &= i\Lambda_1^2 \begin{bmatrix} \langle \tilde{\mathbf{V}}_t, \mathcal{S}\mathbf{V}_t\rangle_{L^2} & 0 \\ 0 & \langle \tilde{\mathbf{V}}_g, \mathcal{S}\mathbf{V}_g\rangle_{L^2} \end{bmatrix} \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix} \\ &+ \begin{bmatrix} \langle \check{\mathbf{V}}_t, \mathcal{P}\mathbf{V}_t\rangle_{L^2} & 0 \\ 0 & \langle \check{\mathbf{V}}_g, \mathcal{P}\mathbf{V}_g\rangle_{L^2} \end{bmatrix} \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix} \\ &+ i\Lambda_1^2 \begin{bmatrix} 0 & -\langle \check{\mathbf{V}}_t, \mathcal{S}\check{\mathbf{V}}_g\rangle_{L^2} \\ \langle \tilde{\mathbf{V}}_g, \mathcal{S}\check{\mathbf{V}}_t\rangle_{L^2} & 0 \end{bmatrix} \\ &- \Lambda_1^2 \begin{bmatrix} 0 & \langle \tilde{\mathbf{V}}_t, \mathcal{P}\tilde{\mathbf{V}}_g\rangle_{L^2} \\ \langle \tilde{\mathbf{V}}_g, \mathcal{P}\tilde{\mathbf{V}}_t\rangle_{L^2} & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & \langle \check{\mathbf{V}}_t, \mathcal{P}\check{\mathbf{V}}_g\rangle_{L^2} \\ \langle \check{\mathbf{V}}_g, \mathcal{P}\check{\mathbf{V}}_t\rangle_{L^2} & 0 \end{bmatrix}, \end{aligned} \tag{5.92}$$

where Λ_1 is defined uniquely from either solution of the quadratic equations (5.83) and (5.85). Because the 2-by-2 matrix on the right-hand side of equation (5.92) is anti-diagonal, we obtain $\Lambda_2 = 0$ for every choice of Λ_1.

Now, we check that the coefficients α and β are uniquely determined from the right-hand side of the matrix equation (5.92). The coefficient α is determined for $\Lambda_1^2 = \Lambda_r(\omega)^2 > 0$ from the anti-diagonal entry

$$i\Lambda_1^2 \langle \tilde{\mathbf{V}}_t, \mathcal{S}\mathbf{V}_t\rangle_{L^2} + \langle \check{\mathbf{V}}_t, \mathcal{P}\mathbf{V}_t\rangle_{L^2} = i\langle \tilde{\mathbf{V}}_t, \mathcal{S}\mathbf{V}_t\rangle_{L^2} \left(\Lambda_r(\omega)^2 + \Lambda_i(\omega)^2\right),$$

which is nonzero for every $\omega \in (0,1)$. Therefore, we obtain from (5.92) the unique expression for α:

$$\alpha = \frac{\Lambda_r(\omega)^2 \left(\langle \tilde{\mathbf{V}}_t, \mathcal{P}\tilde{\mathbf{V}}_g\rangle_{L^2} + i\langle \check{\mathbf{V}}_t, \mathcal{S}\check{\mathbf{V}}_g\rangle_{L^2}\right) - \langle \check{\mathbf{V}}_t, \mathcal{P}\check{\mathbf{V}}_g\rangle_{L^2}}{i\langle \tilde{\mathbf{V}}_t, \mathcal{S}\mathbf{V}_t\rangle_{L^2} \left(\Lambda_r(\omega)^2 + \Lambda_i(\omega)^2\right)}. \tag{5.93}$$

Similarly, the coefficient β is determined for $\Lambda_1^2 = -\Lambda_i(\omega)^2 < 0$ from the anti-diagonal entry

$$i\Lambda_1^2 \langle \tilde{\mathbf{V}}_g, \mathcal{S}\mathbf{V}_g\rangle_{L^2} + \langle \check{\mathbf{V}}_g, \mathcal{P}\mathbf{V}_g\rangle_{L^2} = -i\langle \tilde{\mathbf{V}}_g, \mathcal{S}\mathbf{V}_g\rangle_{L^2} \left(\Lambda_i(\omega)^2 + \Lambda_r(\omega)^2\right),$$

which is nonzero for every $\omega \in (0,1)$. Therefore, we obtain from (5.92) the unique expression for β:

$$\beta = \frac{\Lambda_i(\omega)^2 \left(i\langle \tilde{\mathbf{V}}_g, \mathcal{S}\check{\mathbf{V}}_t\rangle_{L^2} - \langle \tilde{\mathbf{V}}_g, \mathcal{P}\tilde{\mathbf{V}}_t\rangle_{L^2}\right) - \langle \check{\mathbf{V}}_g, \mathcal{P}\check{\mathbf{V}}_t\rangle_{L^2}}{-i\langle \tilde{\mathbf{V}}_g, \mathcal{S}\mathbf{V}_g\rangle_{L^2} \left(\Lambda_i(\omega)^2 + \Lambda_r(\omega)^2\right)}. \tag{5.94}$$

These computations justify the $\mathcal{O}(p^3)$ terms in the expansions (5.78) and (5.79) for the eigenvalues λ.

5.4 Numerical approximations

We approximate eigenvalues of the spectral stability problems (5.32) and (5.58) with the Chebyshev interpolation method. This method has been already applied to the linearized Dirac system in one dimension in [17]. The block diagonalized systems in (5.32) and (5.58) are discretized on the grid points

$$x_j = L \tanh^{-1}(z_j), \quad j = 0, 1, \dots, N,$$

where $z_j = \cos\left(\frac{j\pi}{N}\right)$ is the Chebyshev node and a scaling parameter L is chosen suitably so that the grid points are concentrated in the region, where the solitary wave U_ω changes fast. Note that $x_0 = \infty$ and $x_N = -\infty$.

According to the standard Chebyshev interpolation method [105], the first derivative that appears in the systems (5.32) and (5.58) is constructed from the scaled Chebyshev differentiation matrix $\widetilde{D}_N$ of the size $(N+1) \times (N+1)$, whose each element at i^{th} row and j^{th} column is given by

$$[\widetilde{D}_N]_{ij} = \frac{1}{L} \operatorname{sech}^2\left(\frac{x_i}{L}\right) [D_N]_{ij},$$

where D_N is the standard Chebyshev differentiation matrix (see page 53 of [105]) and the chain rule $\frac{du}{dx} = \frac{dz}{dx}\frac{du}{dz}$ has been used. Denoting I_N as an identity matrix of the size $(N+1) \times (N+1)$, we replace each term in the systems (5.32) and (5.58) as follows:

$$\partial_x \to \widetilde{D}_N, \quad 1 \to I_N, \quad U_\omega \to \operatorname{diag}(U_\omega(x_0), U_\omega(x_1), \cdots, U_\omega(x_N)),$$

Due to the decay of the solitary wave U_ω to zero at infinity, we have $U_\omega(x_0) = U_\omega(x_N) = 0$.

The resulting discretized systems from (5.32) and (5.58) are of the size $4(N+1) \times 4(N+1)$. Boundary conditions are naturally built into this formulation, because the elements of the first and last rows of the matrix $[\widetilde{D}_N]_{ij}$ are zero. As a result, eigenvalues from the first and last rows of the linear discretized system are nothing but the end points of the continuous spectrum in Proposition 13, whereas the boundary values of the vector $\mathbf{V}$ at the end points x_0 and x_N are identically zero for all other eigenvalues of the linear discretized system.

Throughout all our numerical results, we pick the value of a scaling parameter L to be $L = 10$. This choice ensures that the solitary wave solutions U_ω for all values of ω used in our numerical experiments remain nonzero up to 16 decimals on all interior grid points x_j with $1 \leq j \leq N-1$.

5.4.1 Eigenvalue computations for the massive Thirring model

Figure 5.2 shows eigenvalues of the spectral stability problem (5.32) for the solitary wave (5.7) of the massive Thirring model. We set $\omega = 0$ and display eigenvalues λ in the complex plane for different values of p. The subfigure at $p = 0.2$ demonstrates our analytical result in Lemma 28, which predicts splitting of the zero eigenvalue of algebraic multiplicity four into two pairs of real and imaginary eigenvalues. Increasing the value of p further, we observe emergence of imaginary eigenvalues from the edges of the continuous spectrum branches, as seen at $p = 0.32$. A pair of imaginary eigenvalues coalesces and bifurcates into the complex plane with nonzero real parts, as seen at $p = 0.36$, and later absorbs back into the continuous spectrum branches, seen in the next subfigures. We can also see emergence of a pair of imaginary eigenvalues from the edges of the continuous spectrum branches at $p = 0.915$. The pair bifurcates along the real axis after coalescence at the origin, as seen at $p = 0.93$. The gap of the continuous spectrum closes up at $p = 1$. For a larger value of p, two pairs of real eigenvalues are seen to approach each other.

Figure 5.3 show how the positive imaginary and real eigenvalues bifurcating from the zero eigenvalue depends on p for $\omega = 0.5, 0, -0.5$, respectively at each row. Red solid lines show asymptotic approximations established in Lemma 28 for $\lambda = \Lambda_r(\omega)p$ and $\lambda = i\Lambda_i(\omega)p$. Green filled regions in Figures (5.3a), (5.3c), and (5.3e) denote the location of the continuous spectrum. Symbols $*$ and $+$ in Figures (5.3b), (5.3d), and (5.3f) denote purely real eigenvalues and eigenvalues with nonzero imaginary part.

Numerical results suggest the persistence of transverse instability for any period p because of purely real eigenvalues, which come close to each other and persist for a large p. We observe a stronger instability for a larger solitary wave with $\omega = -0.5$ than for a smaller solitary wave with $\omega = 0.5$. We notice that an imaginary eigenvalue does not reach the edge of the continuous spectrum for $\omega = 0.5$ and $\omega = 0$ due to colliding with other imaginary eigenvalue coming from the edge of the continuous spectrum. On the other hand, an imaginary eigenvalue for $\omega = -0.5$ gets absorbed in the edge of the continuous spectrum. This is explained by the movement of the two branches of the continuous spectrum in the opposite directions: up and down as the value of p varies. Moving-down branch on $\mathrm{Im}(\lambda) > 0$, as seen in $\omega = 0.5$ and $\omega = 0$, expels an eigenvalue from its edge that makes collision with the other imaginary eigenvalue, while moving-up branch on $\mathrm{Im}(\lambda) > 0$, as seen in $\omega = -0.5$, absorbs an imaginary eigenvalue approaching the edge.

To verify a reasonable accuracy of the numerical method, we measure the maximum real part of eigenvalues along the imaginary axis with $|\mathrm{Im}(\lambda)| < 10$. This quantity shows the level of spurious parts of the eigenvalues and it is known to be large in the finite-difference methods applied to the linearized Dirac systems (see discussion in [17]). Table 5.1 shows values of $\max\limits_{|\operatorname{Im}\lambda|<10} |\operatorname{Re}\lambda|$ for three values of ω and three values of the number N of the Chebyshev points. In all numerical

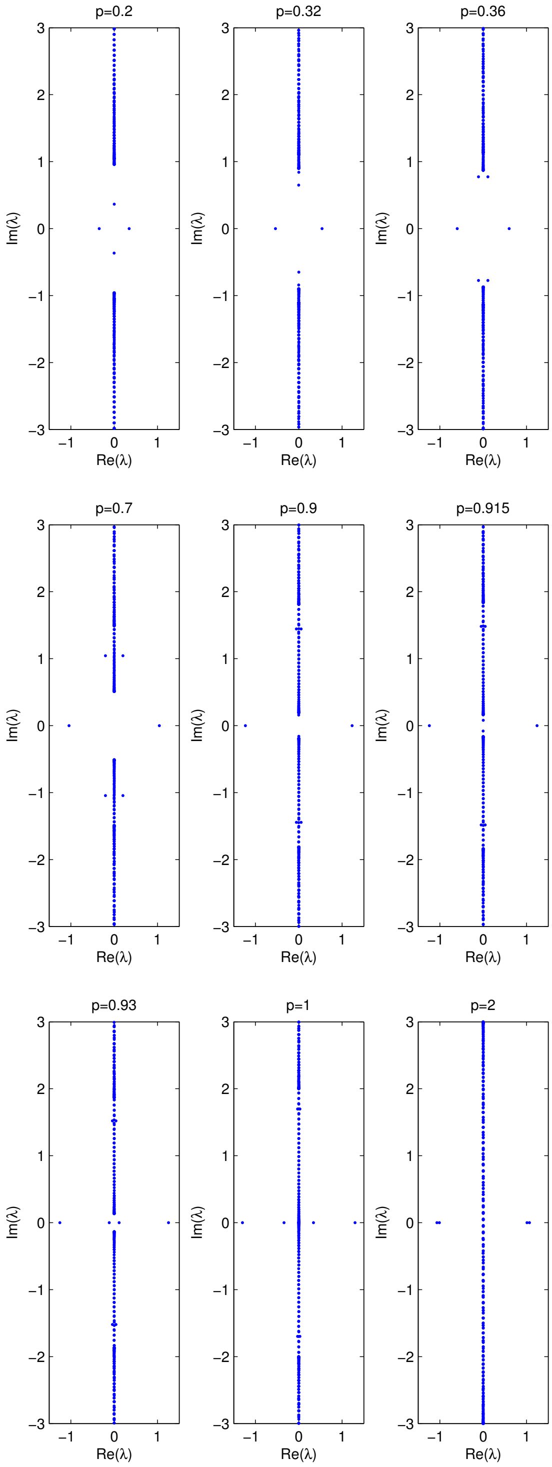
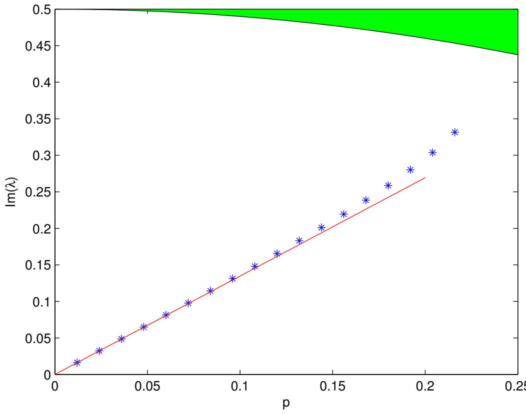
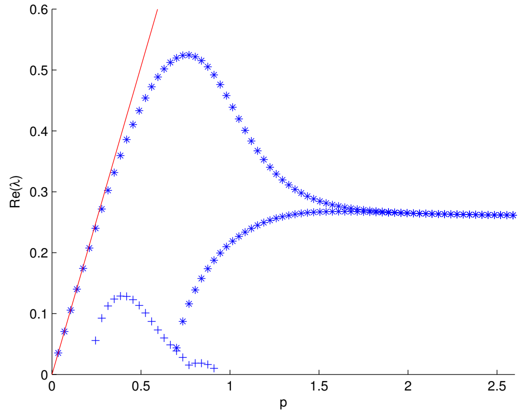


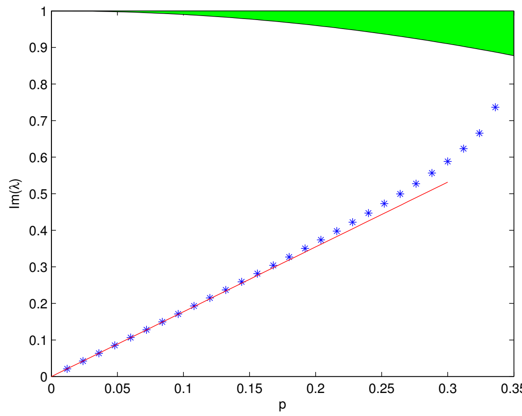
Figure 5.2: Numerical approximations for the spectral problem (5.32) associated with the solitary wave (5.7) of the massive Thirring model at $\omega = 0$.



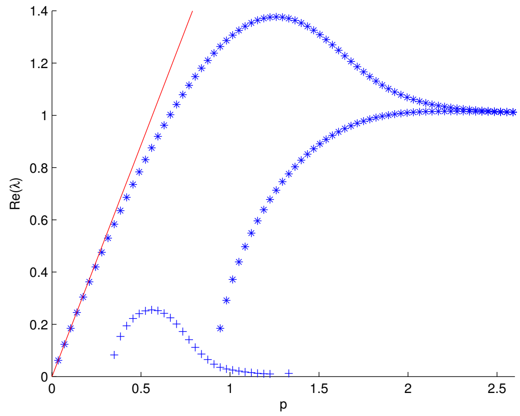
(a) $\omega = 0.5$



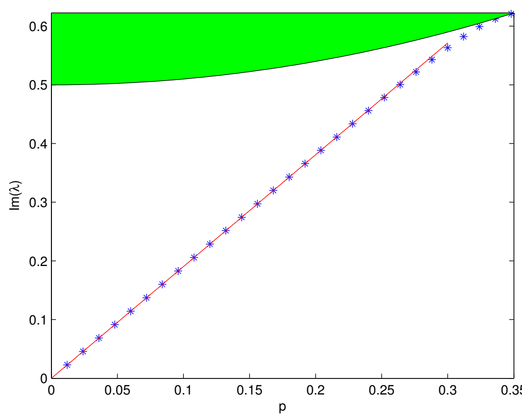
(b) $\omega = 0.5$



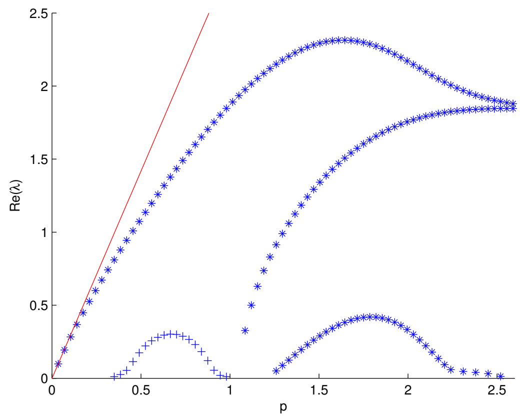
(c) $\omega = 0$



(d) $\omega = 0$

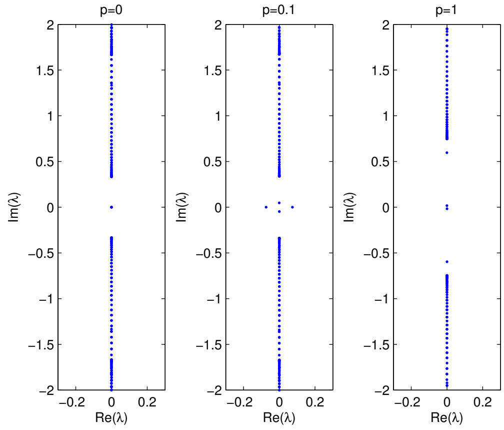


(e) $\omega = -0.5$

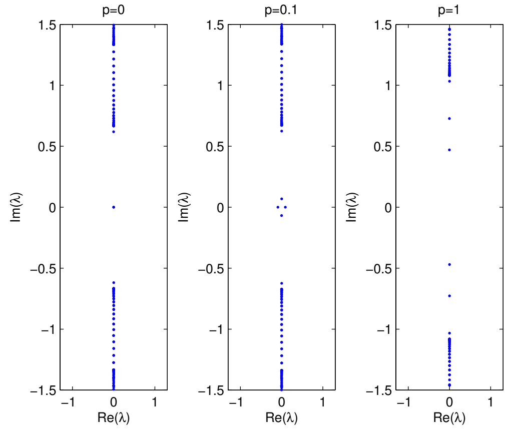


(f) $\omega = -0.5$

Figure 5.3: Numerical approximations of isolated eigenvalues of the spectral problem (5.32) versus parameter p.

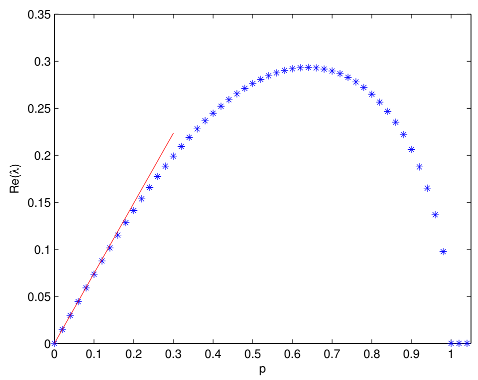


(a) $\omega = 2/3$

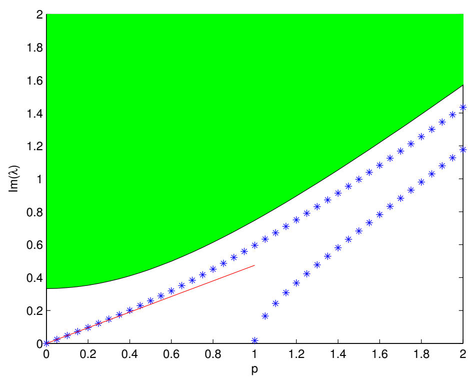


(b) $\omega = 1/3$

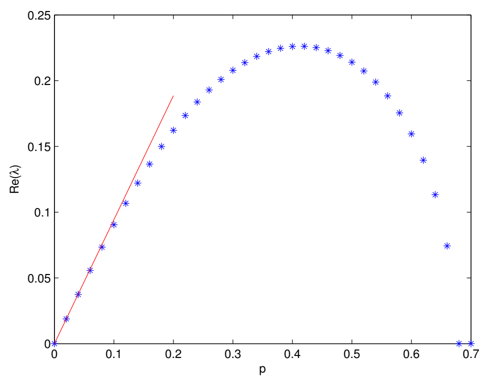
Figure 5.4: Numerical approximations for the spectral problem (5.58) associated with the solitary wave (5.11) of the massive Gross-Neveu model.



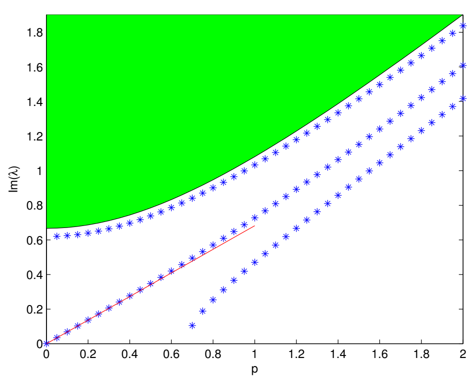
(a) $\omega = 2/3$



(b) $\omega = 2/3$



(c) $\omega = 1/3$



(d) $\omega = 1/3$

Figure 5.5: Numerical approximations of isolated eigenvalues of the spectral problem (5.58) versus parameter p.

computations reported on Figures 5.2 and 5.3, we choose $N = 300$, in this way spurious eigenvalues are hardly visible on the figures.

	$\omega = -0.5$	$\omega = 0$	$\omega = 0.5$
$N = 100$	1.96×10^{-1}	2.57×10^{-1}	1.16×10^{-1}
$N = 300$	1.36×10^{-4}	2.18×10^{-4}	7.02×10^{-5}
$N = 500$	2.22×10^{-7}	8.77×10^{-5}	6.56×10^{-8}

Table 5.1: $\max_{|\operatorname{Im}\lambda|<10} |\operatorname{Re}\lambda|$ versus values of ω and N for the spectral problem (5.32) with $p = 0$.

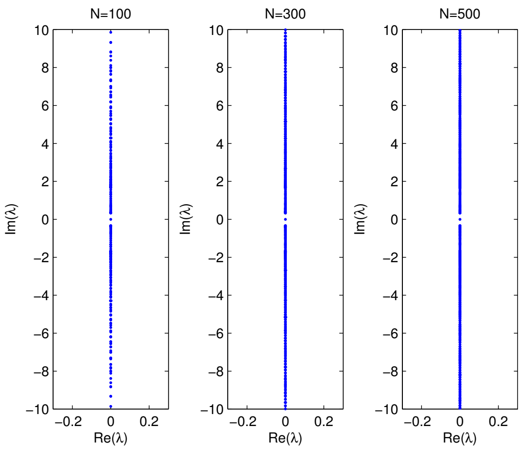
5.4.2 Eigenvalue computations for the massive Gross–Neveu model

Figures 5.4 and 5.5 show eigenvalues of the spectral stability problem (5.58) for the solitary wave (5.11) of the massive Gross–Neveu equation with parameter values $\omega = 2/3$ and $\omega = 1/3$, respectively. We confirm spectral stability of the solitary wave for $p = 0$. In agreement with numerical results in [8], we also observe that the spectrum of a linearized operator for $p = 0$ has an additional pair of imaginary eigenvalues in the case $\omega = 1/3$. (Recall that this issue was considered to be contradictory in the literature with some results reporting spectral instability of solitary waves for $\omega = 1/3$ [73, 99].)

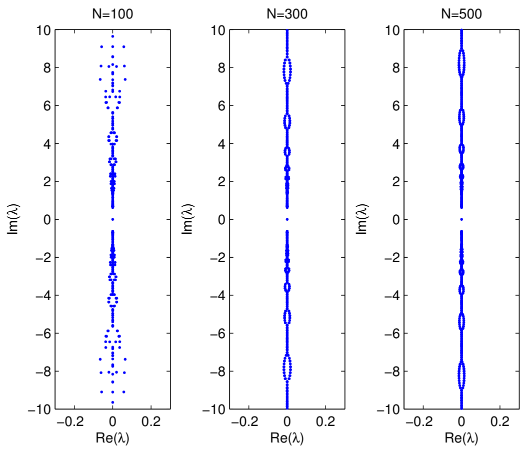
The subfigures of Figure 5.4 at $p = 0.1$ demonstrate our analytical result in Lemma 29, which predicts splitting of the zero eigenvalue of algebraic multiplicity four into two pairs of eigenvalues along the real and imaginary axes. Note that the pair along the real axis persists as the pair of real eigenvalues up to the numerical accuracy. (Recall that the statement of Lemma 29 lacks the result on the persistence of real eigenvalues.) Increasing the values of p further, we observe that the real eigenvalues move back to the origin and split along the imaginary axis, as seen on the subfigures at $p = 1$. The gap of the continuous spectrum branches around the origin is preserved for all values of parameter p. The pairs of imaginary eigenvalues persist in the gap of continuous spectrum for larger values of the parameter p.

Figure 5.5 shows real and imaginary eigenvalues versus p for the same cases $\omega = 2/3$ and $\omega = 1/3$. The green shaded region indicates the location of the continuous spectrum. Red solid lines show asymptotic approximations established in Lemma 29 for $\lambda = \Lambda_r(\omega)p$ and $\lambda = i\Lambda_i(\omega)p$. It follows from our numerical results that the transverse instability has a threshold on the p values so that the solitary waves are spectrally stable for sufficiently large values of p. These thresholds on the transverse instability were observed for other values of ω in $(0, 1)$.

To control the accuracy of the numerical method, we again compute the values of $\max_{|\operatorname{Im}\lambda|<10} |\operatorname{Re}\lambda|$ for spurious parts of eigenvalues along the imaginary axis. Table 5.2 shows the result for two values of ω and three values of N. Compared to Table



(a) $\omega = 2/3$



(b) $\omega = 1/3$

Figure 5.6: Numerically computed λ for the spectral problem (5.58) with $p = 0$ for different values of the number N of Chebyshev points.

5.1, we observe a slower convergence rate and lower accuracy of our numerical approximations.

	$\omega = 1/3$	$\omega = 2/3$
$N = 100$	6.48×10^{-2}	2.03×10^{-3}
$N = 300$	1.72×10^{-2}	1.68×10^{-3}
$N = 500$	1.38×10^{-2}	1.20×10^{-3}

Table 5.2: $\max\limits_{|\operatorname{Im}\lambda|<10} |\operatorname{Re}\lambda|$ versus values of ω and N for the spectral problem (5.58) with $p = 0$.

We found that spurious eigenvalues are more visible for smaller values of ω, in particular, for the value $\omega = 1/3$, evidenced in Figure 5.6. While spurious eigenvalues in the case of $\omega = 1/3$ in Figure 5.6 are quite visible, the maximum real part of eigenvalues with $|\operatorname{Im}\lambda| < 2$ is much smaller for $N = 400$. As a result, the value $N = 400$ was chosen for numerical approximations reported on Figures 5.4 and 5.5, this choice guarantees that spurious eigenvalues are hardly visible on the figures.

Bibliography

[1] M. J. Ablowitz, B. Prinari, A.D. Trubatch, "Discrete and continuous nonlinear Schrödinger systems" (Cambridge University Press, Cambridge, 2004).

[2] M.J. Ablowitz and Y. Zhu, "Nonlinear waves in shallow honeycomb lattices", SIAM J. Appl. Math. **72** (2012), 240–260.

[3] M.J. Ablowitz and Y. Zhu, "Nonlinear wave packets in deformed honeycomb lattice", SIAM J. Appl. Math. **73** (2013), 1959-1979.

[4] M.A. Alejo and C. Munoz, "Nonlinear stability of MKdV breathers", arXiv:1206.3157 (2012).

[5] I.V. Barashenkov, D.E. Pelinovsky, and E.V. Zemlyanaya, "Vibrations and oscillatory instabilities of gap solitons", *Phys. Rev. Lett.* **80** (1998), 5117–5120.

[6] R. Beals and R. R. Coifman, "Scattering and inverse scattering for first order systems", *Comm. Pure Appl. Math.* **37** (1984), 39–90.

[7] R. Beals and R. R. Coifman, "Inverse scattering and evolution equations", *Comm. Pure Appl. Math.* **38** (1985), 29–42.

[8] G. Berkolaiko, A. Comech, and A. Sukhtayev, "Vakhitov–Kolokolov and energy vanishing conditions for linear instability of solitary waves in models of classical self-interacting spinor fields", *Nonlinearity* **28** (2015), 577–592.

[9] G. Berkolaiko and A. Comech, "On spectral stability of solitary waves of nonlinear Dirac equations in 1D", *Math. Model. Nat. Phenom.* **7** (2012), no. 2, 13–31.

[10] H.A. Biagioni and F. Linares, "Ill-posedness for the derivative Schrödinger and generalized Benjamin-Ono equations", Trans. Amer. Math. Soc. **353** (2001), 3649–3659.

[11] N. Boussaid, "Stable directions for small nonlinear Dirac standing waves", Comm. Math. Phys. **268** (2006), no. 3, 757–817.

[12] N. Boussaid, "On the asymptotic stability of small nonlinear Dirac standing waves in a resonant case," SIAM J. Math. Anal. **40** (2008), no. 4, 1621–1670.

[13] N. Boussaid and A. Comech, "On spectral stability of the nonlinear Dirac equation", arXiv:1211.3336 (2012).

[14] N. Boussaid and S. Cuccagna, "On stabibity of standing waves of nonlinear Dirac equations", Comm. PDEs **37** (2012), no. 6, 1001–1056.

[15] T. Candy, "Global existence of an L^2-critical Dirac equation in one dimension", Adv. Diff. Eqs. **7-8** (2011), 643-666

[16] Y. Cher, G. Simpson, and C. Sulem, "Local structure of singular profiles for a derivative Schrödinger equation", to be submitted (2016).

[17] M. Chugunova, D. Pelinovsky, "Block-diagonalization of the symmetric first-order coupled-mode system," *SIAM J. Appl. Dyn. Syst.* **5** (2006), 66–83.

[18] A. Comech, M. Guan, and S. Gustafson, "On linear instability of solitary waves for the nonlinear Dirac equation", Annales de lÍnstitute Henri Poincaré **31** (2014), 639–654.

[19] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Global well-posedness for Schrödinger equations with derivative", SIAM J. Math. Anal., **33** (2001), 649–669.

[20] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "A refined global well-posedness result for Schrödinger equations with derivative", SIAM J. Math. Anal., **34** (2002), 64–86.

[21] M. Colin and M. Ohta, "Stability of solitary waves for derivative nonlinear Schrödinger equation", Ann. I.H. Poincaré-AN **23** (2006), 753–764.

[22] A. Contreras and D. E. Pelinovsky, "Stability of multi-solitons in the cubic NLS equation", J. Hyper. Differential Equations **11** (2014), 329

[23] A. Contreras, D. E. Pelinovsky, and Y. Shimabukuro, "L^2 orbital stability of Dirac solitons in the massive Thirring model", Commun. part. diff. eq. **41** (2016), 227-255

[24] S. Cuccagna and D. E. Pelinovsky, "The asymptotic stability of solitons in the cubic NLS equation on the line", Applicable Analysis **93** (2014), 791-822

[25] G. Darboux, "On a proposition relative to linear equations", Comptes Rendus Acad. Sci. (Paris) **94** (1882), 1456

[26] V. Delgado, "Global solutions of the Cauchy problem for the (classical) coupled Maxwell–Dirac and other nonlinear Dirac equations in one space dimension", Proc. AMS **69** (1978), 289–296.

[27] G. Derks and G.A. Gottwald, "A robust numerical method to study oscillatory instability of gap solitary waves", SIAM J. Appl. Dyn. Syst. **4**, 140–158 (2005).

[28] P. Deift, "Orthogonal polynomials and random matrices: a Riemann-Hilbert approach", Courant Lecture notes 3, (1999)

[29] P.A. Deift and J. Park, "Long-time asymptotics for solutions of the NLS equation with a delta potential and even initial data", Int. Math. Res. Notes **24** (2011), 5505–5624.

[30] P. Deift and X. Zhou, "A steepest descent method for oscillatory Riemann-Hilbert problems", Bull. Amer. Math. Soc. **20** (1992), 119-123

[31] P.A. Deift and X. Zhou, "Long-time asymptotics for solutions of the NLS equation with initial data in weighted Sobolev spaces", Comm. Pure Appl. Math. **56** (2003), 1029–1077.

[32] V. Delgado, "Global solutions of the Cauchy problem for the (classical) coupled Maxwell–Dirac and other nonlinear Dirac equations in one space dimension", Proc. AMS **69** (1978), 289–296

[33] T. Dohnal and A.B. Aceves, "Optical soliton bullets in (2+1)D nonlinear Bragg resonant periodic geometries", Stud. Appl. Math. **115** (2005), 209–232.

[34] J. Duoandikoetxea, *Fourier Analysis*, Graduate Studies in Mathematics **29** (AMS, Providence, 2001).

[35] P. Duren, *Theory of H^p Spaces* (Academic Press, New York, 1970).

[36] C. L. Fefferman, M. I. Weinstein, "Honeycomb lattice potentials and Dirac points", J. Am. Math. Soc. **25**, 1169-1220 (2012)

[37] C. L. Fefferman, M. I. Weinstein, "Wave packets in honeycomb structures and two-dimensional Dirac equations", Commun. Math. Phys. **326**, 251-286 (2014)

[38] C. L. Fefferman, M. I. Weinstein, "Waves in honeycomb structures", Journees Equations aux derivees partialles, Biarretz, 3-7 juin 2012, GDR 243 4 (CNRS).

[39] A. S. Fokas, A. R. Its, and A. V. Kitaev, "The isomonodromy approach to matrix models in 2D quantum gravity", Commun. Math. Phys. **147** (1992), 395-440

[40] G. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "A method for solving the Korteweg-de-Vries equation". Phys. Rev. Lett. **19** (1967), 1095

[41] M. Grillakis, J. Shatah, and W. Strauss, "Stability theory of solitary waves in the presence of symmetry", J. Funct. Anal. **74** (1987), 160–197; J. Funct. Anal. **94** (1990), 308–348.

[42] D.J. Gross and A. Neveu, "Dynamical symmetry breaking in asymptotically free field theories", Phys. Rev. D **10** (1974), 3235-3253.

[43] N. Hayashi, “The initial value problem for the derivative nonlinear Schrödinger equation in the energy space”, Nonlinear Anal. **20** (1993), 823–833

[44] N. Hayashi and T. Ozawa, “On the derivative nonlinear Schrödinger equation”, Physica D **55** (1992), 14–36.

[45] N. Hayashi and T. Ozawa, “Finite energy solution of nonlinear Schrödinger equations of derivative type”, SIAM J. Math. Anal. **25** (1994), 1488–1503.

[46] A. Hoffman and C. E. Wayne, “Orbital stability of localized structures via Backlund transformations”, Diff. Integral Eqs. **26** (2013), 303-320

[47] H. Huh, “Global solutions to Gross–Neveu equation”, Lett. Math. Phys. **103** (2013), 927–931.

[48] H. Huh and B. Moon, “Low regularity well-posedness for Gross–Neveu equations”, Comm. Pure Appl. Anal. **14** (2015), 1903–1913.

[49] Q.E. Hoq, J. Gagnon, P.G. Kevrekidis, B.A. Malomed, D.J. Franzeskakis, and R. Carretero–Gonzáles, “Extended nonlinear waves in multidimensional dynamical lattices”, Math. Comp. Simul. 80, 721–731 (2009)

[50] K. Imai, “Generalization of Kaup-Newell inverse scattering formulation and Darboux transformation”, J. Phys. Soc. Japan. **68** (1999), 355-359.

[51] M. Jimbo, T. Miwa, Y. Mori, and M. Sato, “Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent”, Physica **1D** (1980), 80-158

[52] M.A. Johnson, “The transverse instability of periodic waves in Zakharov-Kuznetsov type equations”, Stud. Appl. Math. **124** (2010), 323–345.

[53] M.A. Johnson and K. Zumbrun, “Transverse instability of periodic traveling waves in the generalized Kadomtsev-Petviashvili equation”, SIAM J. Math. Anal. **42** (2010), 2681–2702.

[54] T. Kato, “Perturbation theory for linear operators” (Springer-Verlag, Berlin, 1995)

[55] D.J. Kaup and T.I. Lakoba, ”The squared eigenfunctions of the massive Thirring model in laboratory coordinates”, J. Math. Phys. **37** (1996), 308–323.

[56] D.J. Kaup and A.C. Newell, ”On the Coleman correspondence and the solution of the Massive Thirring model”, Lett. Nuovo Cimento **20** (1977), 325–331.

[57] D. Kaup and A. Newell, “An exact solution for a derivative nonlinear Schrödinger equation”, J. Math. Phys. **19** (1978), 789–801.

[58] P.G. Kevrekidis, J. Gagnon, D.J. Franzeskakis, and B.A. Malomed, "X, Y, and Z waves: extended structures in nonlinear lattices", Phys. Rev. E 75, 016607 (2007).

[59] A.V. Kitaev and A.H. Vartanian, "Leading order temporal asymptotics of the modified nonlinear Schrödinger equation: solitonless sector", Inverse Problems **13** (1997), 1311–1339.

[60] Yu.S. Kivshar and D.E. Pelinovsky, "Self-focusing and transverse instabilities of solitary waves", Phys. Rep. **331** (2000), 117–195.

[61] A. Komech and A. Komech, "Global attraction to solitary waves for a nonlinear Dirac equation with mean field interaction", SIAM J. Math. Anal. **42** (2010), 2944–2964.

[62] E.A. Kopylova, "Weighted energy decay for 1D Dirac equation", Dynamics of PDE **8** (2011), no. 2, 113–125.

[63] E.A. Kuznetzov and A.V. Mikhailov, "On the complete integrability of the two-dimensional classical Thirring model", Theor. Math. Phys. **30** (1977), 193–200.

[64] E. A. Kuznetzov and A. V. Mikhailov, "On the complete integrability of the two-dimensional classical Thirring model", Theor. Math. Phys. **30** (1977), 193-200

[65] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves", Comm. Pure Appl. Math. **21** (1968), 467-490

[66] J.H. Lee, "Analytic properties of a Zakharov–Shabat inverse scattering problem (I)", Chinese J. Math. **14** (1986), 225–248.

[67] J.H. Lee, "Analytic properties of a Zakharov–Shabat inverse scattering problem (II)", Chinese J. Math. **16** (1988), 81–110.

[68] J.H. Lee, "Global solvability of the derivative nonlinear Schrödinger equation", Trans. AMS **314** (1989), 107–118.

[69] J. Liu, P.A. Perry, and C. Sulem, "Global existence for the derivative nonlinear Schrödinger equation by the method of inverse scattering", arXiv:1511.01173 (2016).

[70] X. Liu, G. Simpson, and C. Sulem, "Stability of solitary waves for a generalized derivative nonlinear Schrödinger equation", J. Nonlin. Science **23** (2013), 557–583.

[71] X. Liu, G. Simpson, and C. Sulem, "Focusing singularity in a derivative nonlinear Schrödinger equation", Physica D **262** (2013), 48–58.

[72] "F. Merle and L. Vega, L^2 stability of solitons for KdV equation", Int. Math. Res. Not. **13** (2003), 735-753

[73] F.G. Mertens, N.R. Quintero, F. Cooper, A. Khare, and A. Saxena, "Nonlinear Dirac equation solitary waves in external fields", Phys. Rev. E **86**, 046602 (2012).

[74] T. Mizumachi and R.L. Pego, "Asymptotic stability of Toda lattice solitons", Nonlinearity **21** (2008), 2099–2111.

[75] C. Miao, Y. Wu, and G. Xu, "Global well-posedness for Schrödinger equation with derivative in $H^{1/2}(\mathbb{R})$", J. Diff. Eqs., **251** (2011), 2164–2195.

[76] P. Miller, *Applied Asymptotic Analysis*, Graduate Studies in Mathematics **75** (AMS Publications, Providence, 2006).

[77] T. Mizumachi, "Stability of line solitons for the KP-II equation in $\mathbb{R}^2$", Memoirs of the AMS **238** (2015)

[78] T. Mizumachi, R. Pego, and J. Quintero, "Asymptotic stability of solitary waves in the Benney-Luke model of water waves", Differ. Integral. Equ. **26** (2013), no. 3-4, 253-301.

[79] T. Mizumachi and D. E. Pelinovsky, "Backlund transformation and L^2-stability of NLS solitons", Int. Math, Res. Not. **9** (2012), 2034-2067

[80] S. J. Orfanidis, "Soliton solutions of the massive Thirring model and the inverse scattering transform", Phys. Rev. D **14** (1976), 472–478.

[81] T. Ozawa, "On the nonlinear Schrödinger equations of derivative type", Indiana Univ. Math. J. **45** (1996), 137–163.

[82] R. L. Pego and M. I. Weinstein, "Asymptotic stability of solitary waves", Comm. Math. Phys. **164** (1994), 305-349.

[83] D.E. Pelinovsky, " Localization in periodic potentials: from Schrödinger operators to the Gross–Pitaevskii equation" (Cambridge University Press, Cambridge, 2011).

[84] D.E. Pelinovsky, "Survey on global existence in the nonlinear Dirac equations in one dimension", in *Harmonic Analysis and Nonlinear Partial Differential Equations* (Editors: T. Ozawa and M. Sugimoto) RIMS Kokyuroku Bessatsu **B26** (2011), 37–50

[85] D. Pelinovsky, F. Rousset, and N. Tzvetkov, "Normal form for transverse instability of line solitons of the Zakharov–Kuznetsov equation", in preparation.

[86] D. E. Pelinovsky, E. A. Ruvinskaya, O. A. Kurkina, B. Deconinck, "Short-wave transverse instabilities of line solitons of the 2-D hyperbolic nonlinear Schrödinger equation", Theor. Math. Phys. **179** (2014), 452–461.

[87] D. Pelinovsky and G. Schneider, "Bounds on the tight-binding approximation for the Gross-Pitaevskii equation with a periodic potential", J. Diff. Eqs. **248** (2010), 837–849.

[88] D. E. Pelinovsky and Y. Shimabukuro, "Orbital stability of Dirac solitons", Lett. Math. Phys. **104** (2014), 21-41

[89] D. E. Pelinovsky and Y. Shimabukuro, "Transverse instability of line solitary waves in massive Dirac equations", J. nonlinear. sci. **26** (2016), 365-403

[90] D. E. Pelinovsky and Y. Shimabukuro, "Existence of global solutions to the derivative NLS equation with the inverse scattering transform method", submitted (2016)

[91] D.E. Pelinovsky, G. Simpson, and M.I. Weinstein, "Polychromatic solitons in a periodic and nonlinear Maxwell system", SIAM J. Appl. Dynam. Syst. **11** (2012), 478–506.

[92] D.E. Pelinovsky and A. Stefanov, "Asymptotic stability of small gap solitons in nonlinear Dirac equations", J. Math. Phys. **53** (2012), 073705 (27 pages).

[93] D.E. Pelinovsky and J. Yang, "On transverse stability of discrete line solitons", Physica D **255** (2014), 1–11.

[94] F. Rousset and N. Tzvetkov, "Stability and instability of the KDV solitary wave under the KP-I flow", Commun. Math. Phys. **313** (2012), 155–173.

[95] M. Reed and B. Simon, Methods of modern mathematical physics. IV. Analysis of operators, Academic Press [Harcourt Brace Jovanovich Publishers], New York 1978

[96] C. Rogers and W. K. Schief, "Bäcklund and Darboux transformations. Geometry and modern applications in soliton Theory", Cambridge University Press, 200, xviii+413 pp.

[97] G. Schneider and H. Uecker, "Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems", Asymp. Anal. **28**, 163–180 (2001).

[98] S. Selberg and A. Tesfahun, "Low regularity well-posedness for some nonlinear Dirac equations in one space dimension", Diff. Integral Eqs. **23** (2010), 265–278.

[99] S. Shao, N.R. Quintero, F.G. Mertens, F. Cooper, A. Khare, and A. Saxena, "Stability of solitary waves in the nonlinear Dirac equations with arbitrary nonlinearity", Phys. Rev. E **90** (2014), 032915.

[100] H. Steudel, "The hierarchy of multi-soliton solutions of the derivative nonlinear Schrödinger equation", J. Phys. A **36** (2003), 1931-46

[101] H. Takaoka, "Well-posedness for the one dimensional Schrödinger equation with the derivative nonlinearity", Adv. Diff. Eq. **4** (1999), 561–680.

[102] H. Takaoka, "Global well-posedness for Schrödinger equations with derivative in a nonlinear term and data in low-order Sobolev spaces", Electronic J. Diff. Eqs. **2001** (2001), no. 42, 1–23.

[103] T. Tao, Nonlinear dispersive equations. Local and global analysis. CBMS Regional Conference Series in Mathematics, **106**. American Mathematical Society, Providence, RI, 2006

[104] W. Thirring, "A soluble relativistic field theory", Annals of Physics **3** (1958), 91-112.

[105] L. N. Trefethen, *Spectral methods in MATLAB* (SIAM, Philadelphia, 2000).

[106] M. Tsutsumi and I. Fukuda, "On solutions of the derivative nonlinear Schrödinger equation. Existence and uniqueness theorem", Funkcialaj Ekvacioj **23** (1980), 259–277.

[107] M. Tsutsumi and I. Fukuda, "On solutions of the derivative nonlinear Schrödinger equation II", Funkcialaj Ekvacioj **24** (1981), 85–94.

[108] A.H. Vartanian, "Higher order asymptotics of the modified nonlinear Schrödinger equation", Comm. PDEs **25** (2000), 1043–1098.

[109] M. Weinstein, "Nonlinear Schrödinger equations and sharp interpolation estimates", Comm. Math. Phys. **87** (1983), 567–576.

[110] Y. Wu, "Global well-posedness for the nonlinear Schrödinger equation with derivative in energy space", **6** (2013), 1989–2002.

[111] Y. Wu, "Global well-posedness on the derivative nonlinear Schrödinger equation revisited", arXiv:1404.5159 (2014).

[112] J. Xu and E.G. Fan, "Long-time asymptotic for the derivative nonlinear Schrödinger equation with decaying initial value", arXiv:1209.4245 (2012).

[113] J. Xu, E.G. Fan, and Y. Chen, "Long-time asymptotic for the derivative nonlinear Schrödinger equation with step-like initial value", Math. Phys. Anal. Geom. **16** (2013), 253–288.

[114] S. Xu, J. He, and L. Wang, "The Darboux transformation of the derivative nonlinear Schrödinger equation", J. Phys. A **44** (2011), 305203

[115] Y. Yamazaki, "Stability of line standing waves near the bifurcation point for nonlinear Schrödinger equations," Kodai Math. J. **38** (2015), 65–96.

[116] J. Yang, "Transversely stable soliton trains in photonic lattice", Phys. Rev. A 84, 033840 (2011).

[117] J. Yang, D. Gallardo, A. Miller, and Z. Chen, "Elimination of transverse instability in stripe solitons by one-dimensional lattices", Opt. Lett. 37, 1571–1573 (2012).

[118] V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one dimensional self-modulation of waves in nonlinear media", Zh. Eksp. Teor. Fiz. **61** (1971), 118-134

[119] V.E. Zakharov and A.M. Rubenchik, "Instability of waveguides and solitons in nonlinear media", Sov. Phys. JETP **38** (1974), 494-500 (1974).

[120] Y. Zhang and Q. Zhao, "Global solution to nonlinear Dirac equation for Gross–Neveu model in $1+1$ dimensions, Nonlin. Anal.: Th. Meth. Appl. **118** (2015), 82–96.

[121] X. Zhou, "L^2-Sobolev space bijectivity of the scattering and inverse scattering transforms", Comm. Pure Appl. Math., **51** (1998), 697–731.

[122] X. Zhou, "The Riemann–Hilbert problem and inverse scattering", SIAM J. Math. Anal. **20** (1989), 966–986.