Title
Multisolitons for the cubic NLS in 1-d and their stability

Authors
Koch, Herbert
Tataru, Daniel

Publication Date
2024

DOI
10.1007/s10240-024-00148-8

MULTISOLITONS FOR THE CUBIC NLS IN 1-D AND THEIR STABILITY

HERBERT KOCH AND DANIEL TATARU

Abstract. For both the cubic Nonlinear Schrödinger Equation (NLS) as well as the modified Korteweg-de Vries (mKdV) equation in one space dimension we consider the set $\mathbf{M}_N$ of pure N-soliton states, and their associated multisoliton solutions. We prove that (i) the set $\mathbf{M}_N$ is a uniformly smooth manifold, and (ii) the $\mathbf{M}_N$ states are uniformly stable in H^s, for each $s > -\frac{1}{2}$.

One main tool in our analysis is an iterated Bäcklund transform, which allows us to nonlinearly add a multisoliton to an existing soliton free state (the soliton addition map) or alternatively to remove a multisoliton from a multisoliton state (the soliton removal map). The properties and the regularity of these maps are extensively studied.

Contents

1. Introduction

In this article we consider the focusing cubic Nonlinear Schrödinger equation (NLS)

$$iu_t + u_{xx} + 2u|u|^2 = 0, \qquad u(0) = u_0, \tag{1.1}$$

and the complex focusing modified Korteweg-de Vries equation (mKdV)

$$u_t + u_{xxx} + 6|u|^2 u_x = 0, \qquad u(0) = u_0, \tag{1.2}$$

on the real line, with real or complex solutions in one space dimension.

We are considering these equations together since they are commuting Hamiltonian flows. They are also completely integrable with a common Lax operator, and they are the first two nontrivial flows of the NLS-hierarchy of a countable number of commuting flows. Both admit soliton solutions, and the pure soliton states are common for the two equations. These can be obtained from the state

$$Q_0 = 2\,\mathrm{sech}\,2x \tag{1.3}$$

via scaling and Galilean symmetry (which we will call *spectral parameters*) and translations and phase shifts (which we will call the *scattering parameters*). The set of all single soliton states can be thought of as a four dimensional real manifold, or alternately as a two dimensional complex manifold, which we denote by $\mathbf{M}_1$. The pure soliton states are known to be orbitally stable in various topologies such as L^2 and H^1.

Our aim in this paper is to study multisoliton states and solutions, which can be thought of as the nonlinear superposition of several single soliton solutions. Precisely, for $N \geq 1$ we consider the family of pure N-soliton states, which we denote by $\mathbf{M}_N$, and we investigate its geometry as well as its stability with respect to both the mKdV flow, the NLS flow as well as all higher order flows. Some of this analysis has been carried out before by several authors, under the assumption that the spectral parameters of the N component solitons (or equivalently, the eigenvalues of the associated Lax operator) are separated. Instead, our emphasis will be on what happens when the spectral parameters are close, including the higher multiplicity case. Our two main results can be summarized as follows:

I. Regularity: The family $\mathbf{M}_N$ of N-solitons is a uniformly smooth, symplectic $4N$ dimensional submanifold of H^s for all $s > -\frac{1}{2}$.

II. Stability: The family $\mathbf{M}_N$ is uniformly stable with respect to both the NLS and the mKdV flows in H^s for all $s > -\frac{1}{2}$, in the sense that for any initial data u_0 with distance ε to $\mathbf{M}_N$ there is a pure N soliton solution v so that $\|u(t) - v(t)\|_{H^s} = O(\varepsilon)$.

The main tool in our analysis is the Bäcklund transform, which has been studied before in a context where the spectral parameters are separated. Instead, here we study the Bäcklund transform when the spectral parameters are close are equal, and the results above can be viewed to a certain extent as a consequence of this analysis. Our study of the Bäcklund transform is contained in Sections 3, 4, 5, followed by the main results in Section 6.

The aim of the rest of the introduction is to provide sufficient background in order to allow us to state more precise formulations for both results above.

1.1. **Symmetries and conservation laws.** Both the NLS equation (1.1) and the mKdV equation (1.2) are invariant with respect to translations in space and time and with respect to phase shifts, i.e. multiplication by a complex number of modulus 1.

Moreover the NLS equation is invariant with respect to scaling

$$u(x,t) \to \lambda u(\lambda x, \lambda^2 t),$$

whereas the mKdV equation is invariant with respect to

$$u(x,t) \to \lambda u(\lambda x, \lambda^3 t).$$

The initial data for the two problems scales in the same way,

$$u_0(x) \to \lambda u_0(\lambda x), \tag{1.4}$$

and so does the Sobolev space $\dot{H}^{-\frac{1}{2}}$, which one may view as the critical Sobolev space.

The NLS and the mKdV flows are in effect part of an hierarchy of an infinite number of commuting Hamiltonian flows with respect to the symplectic form

$$\omega(u,v) = 2\,\mathrm{Im} \int u\bar{v}\; dx,$$

hence the Hamiltonian equations of the Hamiltonian H are

$$\frac{d}{ds} H(u+sv)\Big|_{s=0} = i \int \dot{u}v - \overline{\dot{u}}v dx.$$

We consider Hamiltonians which are integrals over densities which we write as sums over products of u and $\bar{u}$ and their derivatives, which are invariant under exchanging u and $\bar{u}$. The Hamiltonian equations are

$$\dot{u} = \frac{1}{i} \frac{\delta H(u)}{\delta \bar{u}}. \tag{1.5}$$

Each of these Hamiltonians of the hierarchy yields joint conservation laws for all of these flows. The first several energies are as follows:

$$
\begin{aligned}
H_0 &= \int |u|^2 \, dx, \\
H_1 &= \frac{1}{i} \int u \partial_x \bar{u} \, dx, \\
H_2 &= \int |u_x|^2 - |u|^4 dx, \qquad (1.6)\\
H_3 &= \frac{1}{i} \int \ u_x \partial_x \overline{u}_x - 3|u|^2 u \partial_x \bar{u} \, dx, \\
H_4 &= \int |u_{xx}|^2 - ||u|^2_x|^2 - \frac{3}{2}|(u^2)_x|^2 + 2|u|^6 \, dx.
\end{aligned}
$$

The even ones are even with respect to complex conjugation and have a positive definite principal part, and we will refer to them as energies. The odd ones are odd under the replacement of u by its complex conjugate, and we will refer to them as momenta. With respect to the symplectic form above, With respect to the symplectic form above, these commuting Hamiltonians generate flows as follows: H_0 generates the phase shifts $e^{-it}u_0$, H_1 generates the group of translations $u_0(x+t)$, H_2 the NLS flow, H_3 the mKdV flow, etc. We denote the respective flows by Φ_n, as operators acting on suitable Sobolev spaces. Because of the commuting property, one can also easily consider combinations of these flows. To denote these combined flows in a compact fashion, for an arbitrary real polynomial

$$
P(z) = \sum_{j=0}^{n} \beta_j z^j,
$$

we define

$$
\Phi(2^{-1}P) = \Phi_0(\beta_0) \circ \Phi_1(2^{-1}\beta_1) \cdots \circ \Phi_n(2^{-n}\beta_n). \qquad (1.7)
$$

Here the factors of 2 arise as a consequence of a mismatch between the standard notations for the Fourier transform and for the scattering transform.

The integrable structure manifests itself in the existence of Lax pairs with the Lax operator

$$
\mathcal{L}(u) = i \begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix}
$$

which we will discuss below. The Lax operator for NLS was first introduced by Zakharov-Shabat [31], following the circle of ideas initiated by Kruskal-Gardner-Green-Miura [24] and Lax [25]. Together with the associated scattering transform, it was studied later by many authors, including [1], [6]. The above flows define an evolution of the Lax operator and associated similarity transforms between those Lax operators.

Here we note that if $u = Q_0$, the pure soliton in (1.3), then

$$
\mathcal{L}(Q_0) \begin{pmatrix} \operatorname{sech}(2x)e^{-x} \\ -\operatorname{sech}(2x)e^{x} \end{pmatrix} = i \begin{pmatrix} \operatorname{sech}(2x)e^{-x} \\ -\operatorname{sech}(2x)e^{x} \end{pmatrix}, \qquad \mathcal{L}(Q_0) \begin{pmatrix} \operatorname{sech}(2x)e^{x} \\ \operatorname{sech}(2x)e^{-x} \end{pmatrix} = -i \begin{pmatrix} \operatorname{sech}(2x)e^{x} \\ \operatorname{sech}(2x)e^{-x} \end{pmatrix},
$$

and hence $z = \pm i$ are eigenvalues of $\mathcal{L}(Q_0)$. The full spectrum of $\mathcal{L}(Q_0)$ consists of these two eigenvalues together with the real axis, which represents its continuous spectrum.

All equations of the hierarchy are well-posed for initial data in the Schwartz space, see Zhou [33]: A modified inverse scattering transform maps

$$
H^{l,k} = \{u \in H^l, x^k u \in L^2\}
$$

bijectively to the scattering data (more precisely some Riemann-Hilbert data) in a space denoted by $H^k \cap L^2(1+|z|^{2l}|dz|)$ for any $k \geq 1$ and $l \geq 0$ on a contour which is the union of the real line with a large circle. For more details, see Theorem 1.8 in [33] for the map, and Theorem 2.7 and Theorem 2.11 for the inverse map for the simpler cases $l = 0, 1$; The case of large l require only an understanding of the situation of large z, which is the same as for small data and has been understood already in [6]. The Hamiltonian evolution of the j-th Hamiltonian is effectively linear on the scattering data r: It maps

$$r \to e^{i\frac{t}{2}(2z)^j} r$$

which defines a strongly continuous group on $H^k \cap L^2(1+|z|^{2jk})$ and a unique evolution on Schwartz functions. This result is much older in the case of generic data, see Beals and Coifman [6].

The odd order equations preserve real data. This is mKdV hierarchy. In this case stronger local existence results as well as illposedness results in the sense of failure of uniform continuity have been proven by Grünrock [18]. These flows are never smooth in H^s with respect to the initial data, no matter how large s is chosen, if the order of the equation is 5 or higher. In view of recent work by Harrop-Griffith, Killip and Visan [19], who proved well-posedness for NLS and mKdV in H^s, $s > -1/2$, one might hope that the whole hierarchy is well-posed in H^s for all $s > -1/2$ on the real line.

1.2. **The Galilean invariance and frequency shifts.** Classically the Galilean invariance is a symmetry of the NLS equation,

$$u(x,t) \to e^{i(x\xi_0 - t\xi_0^2)} u(x - 2t\xi_0, t). \tag{1.8}$$

This can be reformulated using the notation of (1.7) as

$$\begin{aligned} \Phi(2tz^2)e^{ix\xi_0}u_0 &= \Phi_2(t)(e^{ix\xi_0}u_0) = e^{ix\xi_0}\Phi_2(t)\Phi_1(-2t\xi_0)\Phi_0(\xi_0^2)u_0 \\ &= e^{ix\xi_0}\Phi(2tz^2 - 2t\xi_0 z + t\xi_0^2)u_0 = e^{ix\xi}\Phi(2t(z-\xi_0/2)^2)u_0, \end{aligned} \tag{1.9}$$

which says that phase shifts lead to linear combinations of all the lower flows. In this form it generalizes to all higher flows,

$$\Phi(P)(e^{ix\xi_0}u_0) = e^{ix\xi_0}\Phi(P(. - \xi_0/2))u_0. \tag{1.10}$$

This is a straightforward algebraic computation for the corresponding linear flows, where one can use their Fourier representation. To pass to the nonlinear flows, the easiest way is to use the scattering transform. This conjugates the nonlinear flows to the linear flows, while commuting with the phase shifts G_ξ. Such an argument applies rigorously in the Schwartz class, and extends by density to any Sobolev spaces H^s where these flows are well defined.

1.3. **Solitons and soliton parameters.** Here we start from the soliton data Q_0 given by (1.3). Then $2e^{4it}\operatorname{sech}(2x)$ is a soliton solution to NLS with initial data Q_0, and $2\operatorname{sech}(2x-8t)$ is a soliton of mKdV with the same data. We use the symmetries of the NLS equation to generate a full set of single soliton data. We divide this process into two parts:

a) Spectral parameters. These correspond to two symmetries, namely (i) modulation,

$$u(x) \to e^{-ix(2\xi)}u(x),$$

and (ii) the scale invariance,

$$u(x) \to \lambda u(\lambda x).$$

The reason for the factor 2 is again due to the mismatch between standard notations for the Fourier transform and the scattering transform as in (1.7). We compose the two symmetries into

$$z = \xi + i\lambda \in \mathbf{H} \tag{1.11}$$

and

$$S_z u(x) = e^{-ix(2\xi)} \lambda u(\lambda x).$$

b) Scattering parameters. These correspond to two symmetries, namely (i) the translation invariance, which at the level of the initial data yield

$$u(x) \to u(x-a),$$

and (ii) the phase invariance,

$$u(x) \to e^{-2i\theta} u(x).$$

We also assemble these together into a complex variable as

$$\kappa = a + i\theta, \tag{1.12}$$

and define

$$S_\kappa u(x) = e^{-2i\theta} u(x-a).$$

Then the set $\mathbf{M}_1$ of all single soliton data is given by

$$Q_{z,\kappa}(x) = S_z S_\kappa Q_0(x) = e^{-2i(\theta + x\xi)} \lambda Q_0(\lambda(x-x_0)), \qquad (z,\kappa) \in \mathbf{S}_1 := \mathbf{H} \times (\mathbb{C}/\ \pi i \mathbb{Z}),$$

where the soliton location x_0 is defined by

$$a = x_0 \lambda.$$

We can also represent $Q_{z,\kappa}$ using the first two flows,

$$Q_{z,\kappa} = S_z \Phi(\theta - a\ \cdot) Q_0 = \Phi(\theta - \xi x_0 - x_0\ \cdot) S_z Q_0.$$

Again we calculate

$$\mathcal{L}(Q_{z,\kappa}) \begin{pmatrix} \operatorname{sech}(2(\lambda(x-x_0))e^{-(i\theta + i\xi x + \lambda(x-x_0))} \\ -\operatorname{sech}(2(\lambda(x-x_0))e^{i\theta + i\xi x + \lambda(x-x_0)} \end{pmatrix} = z \begin{pmatrix} \operatorname{sech}(2\lambda(x-x_0))e^{-(i\theta + i\xi x + \lambda(x-x_0))} \\ -\operatorname{sech}(2(\lambda x - x_0))e^{i\theta + i\xi x + \lambda(x-x_0)} \end{pmatrix}, \tag{1.13}$$

which shows that z is an eigenvalue and, with ψ denoting the above eigenfunction to the eigenvalue z, we obtain

$$-e^{2\kappa} = \frac{\lim\limits_{x\to\infty} e^{-izx}\psi^2(x)}{\lim\limits_{x\to-\infty} e^{izx}\psi^1(x)}, \tag{1.14}$$

which gives the interpretation of $\kappa = i\theta + \lambda x_0$ as a scattering parameter.

It is interesting to describe the NLS and mKdV evolution in the single soliton space parametrized by $(z,\kappa) \in \mathbf{S}_1 = \mathbf{H} \times (\mathbb{C}/\ \pi i \mathbb{Z})$ as above. Clearly

$$e^{4it} Q_0(x - x_0)$$

is an NLS solution, and by scaling so is

$$e^{4i\lambda^2 t} \lambda Q_0(\lambda(x-x_0)).$$

The Galilean transform and phase shift give the general one soliton NLS solution

$$Q_{\theta,a,\xi,\lambda} = e^{-2i\theta_0 - 2i\xi x} e^{-i(4\xi^2 t - 4\lambda^2 t)} \lambda Q_0(\lambda(x + 4\xi t - x_0)). \tag{1.15}$$

Here the parameters λ and ξ stay fixed, while for a and θ we obtain

$$\dot a = -4\lambda\xi, \qquad \dot\theta = 2(\xi^2 - \lambda^2).$$

Thus for NLS (recall (1.11) and (1.12)) we have

$$\dot z = 0, \qquad 2\dot\kappa = 2(\dot a + i\dot\theta) = i(2z)^2, \tag{1.16}$$

Similarly, for mKdV

$$Q_0(x - 4t - x_0)$$

is a solution, scaling gives gives the rescaled solution

$$\lambda Q(\lambda(x - 4\lambda^2 t - x_0)),$$

and (1.9) specializes to the general complex soliton solution

$$\begin{aligned} &e^{2i\xi x}\Phi_3(t)\Phi_2(-6\xi_0 t)\Phi_1(12t\xi_0^2)\Phi_0(-8t\xi_0^3)\lambda Q_0(\lambda(x - x_0)) \\ &\quad = e^{i(2\xi x - t(8\xi^3 - 24\xi\lambda^2))} Q_0(\lambda(x - x_0 - 4\lambda^2 t + 12t\xi^2)) \end{aligned}$$

and

$$\dot{a} = 4\lambda^3 - 12\lambda\xi^2, \qquad \dot{\theta} = 4\xi^3 - 12\xi\lambda^2,$$

hence

$$\dot{z} = 0, \qquad 2\dot{\kappa} = i(2z)^3. \tag{1.17}$$

One can interpret these as Hamiltonian flows on

$$\mathbf{M}_1 = \{Q_{\theta,a,\xi,\lambda} : \theta \in \mathbb{R}/\pi\mathbb{Z}, a, \xi \in \mathbb{R}, \lambda > 0\},$$

where

$$Q_{\theta,a,\xi,\lambda}(x) = e^{-2i\theta - 2i\xi x}\lambda Q_0(\lambda x - a). \tag{1.18}$$

The trace formula (see Proposition 1.2) shows that the restriction of the Hamiltonians to the manifold $\mathbf{M}_1$ is

$$H_{NLS} = \frac{2}{3}\operatorname{Im}(2z)^3, \qquad H_{mKdV} = \frac{2}{4}\operatorname{Im}(2z)^4,$$

which could also be seen by a direct calculation. The restriction of the symplectic form

$$\sigma(u, v) = 2\operatorname{Im}\int u\bar{v}dx$$

defines a symplectic form on $\mathbf{M}_1$ which can be expressed in terms of the coordinates $(\lambda, \theta, a, \xi)$. We obtain the symplectic form by a direct calculation,

$$\omega = 8(d\lambda \wedge d\theta + da \wedge d\xi) = 2(d\kappa \wedge dz + d\bar{\kappa} \wedge d\bar{z}),$$

and the Hamiltonian equations for the NLS flow on $\mathbf{M}_1$ are

$$\dot{\lambda} = -\frac{1}{8}\frac{\partial H}{\partial\theta}, \quad \dot{\xi} = \frac{1}{8}\frac{\partial H}{\partial a}, \quad \dot{a} = -\frac{1}{8}\frac{\partial H}{\partial\partial\xi}, \quad \dot{\theta} = \frac{1}{8}\frac{\partial H}{\partial\lambda}.$$

which coincides with the dynamics above for NLS and mKdV in (1.16), (1.17).

A similar reasoning, also based on trace formulas, shows that the n-th Hamiltonian restricted to $\mathbf{M}_1$ is

$$H_n = \frac{2}{n+1}\operatorname{Im}(2z)^{n+1},$$

and the n-th evolution in $\mathbf{M}_1$ is

$$\dot{z} = 0, \qquad 2\dot{\kappa} = i(2z)^n. \tag{1.19}$$

1.4. **The Lax operator and the transmission coefficient.** The Lax operator associated to a state u is given by

$$\mathcal{L} = i \begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix},$$

and the associated spectral problem is

$$L\psi = z\psi.$$

The solutions to the above spectral problem will be called z-wave functions, or simply z-waves.

For z in the upper half-space one can consider two special z-waves ψ_l and ψ_r, called the Jost solutions, which have asymptotics

$$\psi_l(\xi, x, t) = \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix} + o(1)e^{\operatorname{Im} zx} \quad \text{as } x \to -\infty,$$

$$\psi_l(\xi, x, t) = \begin{pmatrix} T^{-1}(z)e^{-izx} \\ 0 \end{pmatrix} + o(1)e^{\operatorname{Im} zx} \quad \text{as } x \to \infty.$$

The function $T = T(z, u)$, called the *transmission coefficient*, is a meromorphic function in the upper half-space, and satisfies $|T| \geq 1$. As u evolves along any of the commuting flows of the NLS family, the transmission coefficient rests unchanged.

The L^2 size of a state u can be described using the transmission coefficient as

$$\|u\|_{L^2}^2 = \lim_{z \to i\infty} 2z \ln T(z, u).$$

Moving this relation to the real axis using the residue theorem yields the trace formula

$$\|u\|_{L^2}^2 = \frac{1}{\pi} \int_{\mathbb{R}} \ln |T(\xi/2)| d\xi + 2 \sum_j n_j \operatorname{Im}(2z_j), \tag{1.20}$$

where (z_j, n_j) are the poles of T with their multiplicity.

A function $u = Q_{z_0,\kappa}$ is a single soliton with spectral parameter z_0 if and only if its transmission coefficient has a pole exactly at z_0 and

$$T(z) = \frac{z - \bar{z}_0}{z - z_0}.$$

In particular z_0 is an eigenvalue of its Lax operator $\mathcal{L}$, and the corresponding eigenfunction ϕ_{z_0} is a multiple of both ψ_l and ψ_r. Then the scattering parameter κ can be read as the proportionality factor between the two Jost functions,

$$\psi_l = -e^{2\kappa}\psi_r. \tag{1.21}$$

1.5. **A full family of conservation laws.** In a prior article [23] the authors have extended the countable family of conservation laws for the 1-d cubic NLS and mKdV, associated to integer Sobolev indices, to a continuous family, associated to all real Sobolev exponents $s > -\frac{1}{2}$. Related conserved energies have been independently constructed by Killip-Visan-Zhang [22], for the range $-\frac{1}{2} < s \leq 1$.

Theorem 1.1 ([23])**.** *For each $s > -\frac{1}{2}$ there exist energy functionals E_s which are globally defined*

$$E_s : H^s \to \mathbb{R},$$

with the following properties:

(1) E_s is conserved along the NLS and mKdV flow.

(2) For all $u \in H^s$ the limit of[1] *$\mp \log |T|$ exists as a positive measure, and the trace formula* (2.10) *holds with absolute convergence in all sums and integrals.*

[1]The choice of signs $\mp$ corresponds to the defocusing/focusing case

(3) If $\|u\|_{l^2_1 DU^2} \leq 1$ then

$$\left|E_s(u) - \|u\|^2_{H^s}\right| \lesssim \|u\|^2_{l^2_1 DU^2} \|u\|^2_{H^s}.$$

(4) The map

$$H^\sigma \times (-\frac{1}{2}, \sigma] \ni (u,s) \to E_s(u)$$

is analytic provided $\frac{i}{2}$ is not an eigenvalue, and it is continuous in $u \in H^\sigma$ in general. It is also continuous in s, and analytic in s for $s < \sigma$.

Here the threshold $i/2$ is not important, and can be changed by scaling. Of course this would also change the energies E_s, though not in an essential way.

The Banach space $l^2 DU^2 = L^2 + DU^2$ is the inhomogeneous version of the DU^2 space, and contains all H^s spaces with $s > -\frac{1}{2}$. It is described in full detail in Appendix A of [23]. It can be viewed as a replacement for the unusable scaling critical space $H^{-\frac{1}{2}}$ and satisfies

$$\|u\|_{l^2 DU^2} \lesssim \|u\|_{B^{-\frac{1}{2}}_{2,1}} \lesssim \|u\|_{H^s}, \qquad s > -\frac{1}{2}. \tag{1.22}$$

The energies E_s in the theorem were defined in [23] in terms of the transmission coefficient T. A full description is provided in the next result, also from [23], which also doubles as a trace formula.

Proposition 1.2 (Trace formulas, [23]). *Let $N > [s]$ and $u \in \mathcal{S}$. In the upper half-space we define the function*

$$\Xi_s(z) = \operatorname{Im} \int_0^z (1+\zeta^2)^s d\zeta,$$

which does not depend on the path of integration. Then

$$\begin{aligned} E_s(u) = & \int (1+\xi^2)^s \operatorname{Re} \ln T(\xi/2) d\xi + 2\sum_k m_k \Xi(2z_k) \\ = & \; 4\sin(\pi s) \int_1^\infty (\tau^2-1)^s \Big[-\operatorname{Re} \ln T(i\tau/2) + \frac{1}{2\pi} \sum_{j=0}^N (-1)^j H_{2j} \tau^{-2j-1} \Big] d\tau + \sum_{j=0}^N \binom{s}{j} H_{2j}, \end{aligned} \tag{1.23}$$

where the k sum runs over all the poles z_k of T with multiplicity m_j.

If there are infinitely many poles for T in the upper half-space then the second expression above is always a convergent integral, whereas in the first expression we have a non-negative integral, plus a sum where all but finitely many terms are positive. This simultaneously allows us to interpret the trace of $\ln |T|$ on the real line as a non-negative measure, and to guarantee the convergence in the k summation.

We also remark on the contribution of the poles which are on the imaginary axis. Precisely the function Ξ_s is real analytic away from $z = i$. Thus the only nonsmooth dependence on u in E_s via the poles comes from the poles which are at i.

Here the choice of the function $(1+z^2)^s$ was somewhat arbitrary, all that matters is that it is holomorphic in the upper half-space minus $i[1,\infty)$ and has the appropriate behavior at infinity. In particular, the conserved Hamiltonians H_j correspond to the the functions z^j, and they can be expressed as

$$H_j(u) = \frac{1}{\pi} \int \xi^j \operatorname{Re} \ln T(\xi/2) d\xi + 2\sum_k m_k \dot{\Xi}_j(2z_k), \tag{1.24}$$

where

$$\dot{\Xi}_j(z) = \operatorname{Im} \int_0^{2z} \zeta^j d\zeta = \frac{1}{j+1} \operatorname{Im}(2z)^{j+1}.$$

For pure solitons the contribution of the first term vanishes, and we are left with

$$H_j(Q_{z,\kappa}) = \frac{2}{j+1} \operatorname{Im}(2z)^{j+1}, \tag{1.25}$$

as mentioned earlier in the paper.

To further clarify the assertions in the theorem, we note that the energy conservation result is established for regular initial data. By the local well-posedness theory, this extends to all H^s data above the (current) Sobolev local well-posedness threshold, which is $s \geq 0$ for NLS, respectively $s \geq \frac{1}{4}$ for mKdV. If s is below these thresholds, then the energy conservation property holds for all data at the threshold, i.e. for L^2 data for NLS, respectively $H^{\frac{1}{4}}$ data for mKdV. It is not known whether the two problems are well-posed below these thresholds and above the scaling; however, it is known that local uniformly continuous dependence fails, see [10]. Recently Harrop-Griffith, Killip and Visan [19] proved that the flow map extends to a continuous map on H^s, for $s > -\frac{1}{2}$, for both NLS and mKdV.

One key consequence of the above result is that, if the initial data is in H^s, then the solutions remain bounded in H^s globally in time in a uniform fashion:

Corollary 1.3 ([23])**.** *Let $s > -\frac{1}{2}$, $R > 0$ and u_0 be an initial data for either NLS or mKdV so that*

$$\|u_0\|_{H^s} \leq R$$

Then the corresponding solution u satisfies the global bound

$$\|u(t)\|_{H^s} \lesssim F(R,s) := \begin{cases} R + R^{1+2s} & s \geq 0 \\ R + R^{\frac{1+4s}{1+2s}} & s < 0 \end{cases}$$

This follows directly from the above theorem if $R \ll 1$. For larger R is still follows from the theorem, but only after applying the scaling (1.4). Here one needs to make the choice $\lambda = cR^{-2}$, with $c \ll 1$ if $s \geq 0$, respectively $\lambda = cR^{-\frac{2}{1+2s}}$ if $s < 0$.

1.6. **Multisolitons.** The main objective of this article is to study multisoliton solutions, both by investigating the geometry of the set of multisoliton states and by studying its stability under the family of commuting flows. The first step is to define multisoliton solutions.

A natural venue to define N-soliton solutions for NLS is to start with n single solitons $Q_1, \cdots, Q_N$, with soliton parameters $(z_1, \kappa_1), \cdots (z_n, \kappa_N)$. Assuming that the soliton speeds $\operatorname{Re} z_1, \cdots \operatorname{Re} z_N$ are distinct, these solitons separate at infinity, and one can actually prove (see [15]) the existence of a unique solution Q so that

$$Q - (Q_1 + \cdots + Q_N) \to 0 \quad \text{in } L^2 \qquad \text{as } t \to \infty. \tag{1.26}$$

However, the above venue does not readily extend to solitons with equal speeds, and also it involves the time evolution. Instead, we will take advantage of the complete integrability of the problem, and use the spectral picture for the Lax operator in order to define N-multisoliton states:

Definition 1.4. *A function Q is an N-multisoliton state with spectral parameters z_j, $1 \leq j \leq N$, if its transmission coefficient is*

$$T(z) = \prod_{j=1}^{N} \frac{z - \bar{z}_j}{z - z_j}. \tag{1.27}$$

The set of all N-multisolitons is denoted by $\mathbf{M}_N$.

The corresponding Lax operator has the values z_j as eigenvalues, with multiplicity corresponding to the multiplicity of the z_j's. If the spectral values z_j are all different then we define the associated scattering parameters using the corresponding eigenfunctions by (1.21). Alternatively one may define N solitons as stationary solutions to a linear combination of the first $2N$ flows [16, 27], an approach we do not pursue.

We remark that when the definition based on (1.26) applies, it yields the same spectral parameters z_j as in (1.27). On the other hand, the scattering parameters predicted by (1.21) and (1.26) are slightly different, as a shift in the effective scattering parameters occurs when two solitons interact. This shift was approximatively computed by Faddeev and Takhtajan [15], at least in the case of separated spectral parameters z_j. Precisely, if all the z_j are distinct then the effective soliton position $\hat{x}_j$ and phase $\hat{\theta}_j$ at infinity in (1.26) satisfy

$$\hat{x}_j - x_j \approx \frac{1}{2\operatorname{Im} z_j}\Big[\sum_{x_k<x_j} \ln\left|\frac{z_j-\bar{z}_k}{z_j-z_k}\right| - \sum_{x_j>x_k} \ln\left|\frac{z_j-\bar{z}_k}{z_j-z_k}\right|\Big]$$

and, modulo π,

$$\hat{\theta}_j - \theta_j \approx \sum_{x_k<x_j} \arg\frac{z_j-\bar{z}_k}{z_j-z_k} - \sum_{x_j>x_k} \arg\frac{z_j-\bar{z}_k}{z_j-z_k}$$

with errors that decay to zero as x_j's separate. On the other hand, these errors grow as the z_j's get closer to each other.

We note that the x_j's can always be separated by flowing far enough along the combined NLS-mKdV flow. Indeed, suppose that the spectral parameters $(z_j)_{j\leq N}$ are all simple and denote the scattering parameters by κ_j. Let s resp t be the times of the flow of NLS, respectively mKdV. Then

$$x_j(s,t)\operatorname{Im} z_j + i\theta_j(s,t) = \kappa_j(s,t) = \kappa_j + 2isz_j^2 + 4itz_j^3,$$

and we can choose (s_n, t_n) so that the distance between the x_j's tends to ∞.

It follows from calculations as in Faddeev and Takhatajan [15] that the subset of multisolitons with simple eigenvalues is smoothly parameterized by the

$$\{(z_j,\kappa_j)|z_j \in \{\operatorname{Im} z>0\}, \kappa_j \in \mathbb{R}/(\pi i\mathbb{Z}), z_j \neq z_k\}$$

On the other hand the singularity on the diagonal in the asymptotic formulas above reflects the fact, discussed in detail later on, that the parametrization of the set of multisoliton states via the soliton parameters is singular near the diagonals $z_j = z_k$. This leads us to a very interesting question:

Is the set $\mathbf{M}_N$ of all N multisoliton states a smooth manifold in some (any) reasonable Sobolev topology, or is it singular at the spectral values with higher multiplicity ?

The first aim of this article is to provide an answer to this question:

Theorem 1.5. *The set $\mathbf{M}_N$ is a uniformly smooth $4N$ dimensional symplectic submanifold of H^s for each $s > -\frac{1}{2}$. Here uniformity holds with respect to spectral parameters restricted to a compact subset of the upper half-space, but without any separation assumption.*

Further results concerning the structural properties of N-solitons are developed in Section 7.

The first challenge in the proof of the theorem will be to resolve the apparent singularity near the diagonal $z_j = z_k$. But a second, equally difficult challenge is to establish the uniformity in the theorem.

We now briefly outline the main steps in the construction of our smooth parametrization of the N-soliton manifold $\mathbf{M}_N$:

- To capture the symmetry of spectral parameters, we identify the set $\mathbf{z} = (z_1, \cdots z_N)$ of unordered spectral parameters with the set of symmetric polynomials $\mathbf{s}$ in $\mathbf{z}$,

$$s_j = \sum_{n=1}^{N} z_n^j$$

for $1 \leq j \leq N$, setting

$$W_N = \{\mathbf{s} \in \mathbb{C}^N : \operatorname{Im} z_n > 0\}.$$

Then we use the smooth topology of W_N as the topology on the space of unordered spectral parameters. We remark that, equivalently, instead of $\mathbf{s}$ above one could use the coefficients of the polynomial $\prod(z - z_j)$ appearing in the denominator of T in (1.27).
- Corresponding to scattering parameters $\kappa_j = 0$ we have a N-soliton state $Q_{\mathbf{z},0}$, which is shown to depend smoothly on $\mathbf{z}$ in the above smooth topology.
- We shift the scattering parameters away from 0 using the first $2N$ commuting flows, see (1.7), in the form

$$P(z) = \sum_{n=0}^{2n-1} \beta_n z^n$$

to define

$$Q_{\mathbf{z},\boldsymbol{\beta}} = \Phi(P)Q(\mathbf{z},0) = \prod_{n=0}^{2N-1} \Phi_n(2^{n-1}\beta_n)Q(\mathbf{z},0).$$

- The map

$$(\mathbf{z}, \boldsymbol{\beta}) \to Q_{\mathbf{z},\boldsymbol{\beta}}$$

provides locally a smooth parametrization of the N-soliton manifold $\mathbf{M}_N$.

We note that this parametrization is global, and corresponds to scattering parameters

$$\kappa_j = iP(z_j),$$

at least for distinct eigenvalues. It defines a diffeomorphism between the smooth manifold

$$\mathbf{S}^N = (W_N \times \mathbb{C}^N)/\{\kappa_j \in i\pi\mathbb{Z}\}$$

and $\mathbf{M}_N$. However, it is not a uniform parametrization.

1.7. **Soliton stability.** One can view the small data part of our earlier result in Theorem 1.1 as a stability statement in H^s for the zero solution of the NLS or mKdV equations. In the focusing case another interesting class of solutions are the pure N-soliton solutions, which belong to all Sobolev spaces H^s.

The second goal of the present article is to establish the H^s stability of these families of solutions. For convenience we state here a less precise form of the result, but which has the advantage that no further preliminaries are needed. A more precise form is provided in Section 8.

Theorem 1.6. *Let $s > -\frac{1}{2}$, and $u_0 \in \mathbf{M}_N$ be a pure N-soliton state. Then*

a) There exist $\varepsilon_0 > 0$ and $C > 0$ so that, for each initial data w_0 which satisfies

$$\|u_0 - w_0\|_{H^s} = \varepsilon < \varepsilon_0, \tag{1.28}$$

there exists another pure multisoliton data $v_0 \in \mathbf{M}_N$ so that the corresponding solutions for either NLS or mKdV satisfy

$$\sup_{t\in\mathbb{R}} \|w(t) - v(t)\|_{H^s} \leq C\varepsilon. \tag{1.29}$$

b) Furthermore, this result is uniform with respect to all N-soliton states u_0 with spectral parameters in a compact subset of the open upper half-plane, and holds for all commuting flows of the NLS hierarchy.

One can interpret part (b) above as saying that the constant C in (1.29) may be chosen to depend only on the spectral parameters, and staying bounded for spectral parameters in any compact subset of the open upper half-plane.

To place this result into context, we note that the stability of solitons has been an intensely studied area. During the last decade there have been several stability results proved using the integrable structure: Hoffman and Wayne [20] describe the use of the Bäcklund transform to obtain stability of (multi)solitons. Mizumachi and Pelinovsky [29] proved stability in L^2 and Cuccagna and Pelinovsky [12] proved asymptotic stability of single solitons for localized initial data.

Turning our attention to multisolitons, we remark that a much more restrictive form of this theorem has been previously proved in L^2 or H^1 under the additional assumption that the spectral parameters of the multi-soliton are distinct; Alejo-Muñoz [3] considered the dynamics of solitons for complex mKdV and their stability and the stability of breathers [4, 5]. Kapitula [21] and Contreras-Pelinovsky [11] studied the stability of NLS multisolitons.

Our result here drastically improves these prior results as follows:

- We allow multisolitons with multiple spectral parameters.
- We prove a result which is uniform near such multi-solitons.
- We prove stability in a full range of Sobolev spaces.

The proof of the above theorem is completed in Section 8. A key ingredient in the proof is provided by the trace formula in Proposition 1.2. This allows us to find energies for which the infimum on the set of potentials with given eigenvalues z_j with multiplicity is minimized for pure N-solitons with these spectral parameters. Moreover we may show that this energy is uniformly convex in the transverse direction in a uniform neighbourhood of this set of N-solitons. The stability then follows once we prove the uniform smoothness of the N-soliton manifold $\mathbf{M}_N$.

1.8. **Iterated Bäcklund transforms.** A key tool in proving the results in this paper is the Bäcklund transform, which adds or subtracts a soliton with given spectral and scattering parameter. On the level of the trace formula the action is very transparent: One adds or subtracts the contribution coming from this spectral value.

Its use in a setting without multiplicities is described by Hoffman and Wayne [20]. Similarly, the iterated Bäcklund transform adds or subtracts a multi-soliton [15, 28, 11, 21, 3, 9]; see also the variant in [7], which allows some extra flexibility. Unfortunately the naive iterated Bäcklund transform deteriorates as spectral parameters get close. We overcome this difficulty by

(1) Choosing an appropriate blow-up of the spectral and scattering coordinates near points with multiplicity, i.e. we prove that the parametrization described in Subsection 1.6 is smooth.
(2) Establishing a cancellation property for the iterated Bäcklund transform, and using it to verify the surjectivity of this parametrization.

The outcome of this analysis is

(i) A soliton addition map
$$B_+^N : H^s \times \mathbf{M}_N \to H^s,$$
which nonlinearly adds an N-soliton $Q_{\mathbf{z},\boldsymbol{\beta}}$ to an H^s state with no eigenvalues near $\mathbf{z}$.
(ii) A soliton removal map
$$B_-^N : H^s \to H^s \times \mathbf{M}_N,$$
which reverses the above process, nonlinearly splitting an H^s state near $\mathbf{M}_N$ into an N-soliton $Q_{\mathbf{z},\boldsymbol{\beta}}$ and an H^s state with no eigenvalues near $\mathbf{z}$.

For these maps we establish two smoothness and flow commutation properties:

- They are smooth, inverse maps.
- They commute with the NLS and mKdV flow, and all other commuting flows, which are viewed as acting separately on each of the inputs/outputs.

The arguments establishing this are local, and yield the smoothness of the N soliton set $\mathbf{M}_N$. To obtain the uniform smoothness of $\mathbf{M}_N$ we combine this with the uniform regularity of the energies and with relations derived from the trace formula.

Finally, we turn our attention to the full uniform smoothness of the soliton addition and removal maps. In full generality, this remains open:

Conjecture 1.7. *The maps B_+^N and B_-^N are uniformly smooth on bounded sets in H^s.*

A more precise version of this conjecture is stated in Section 6. A main difficulty in proving this is that our smooth $(\mathbf{z}, \boldsymbol{\beta})$ parametrization of the N-soliton manifold deteriorates as the solitons separate (which corresponds to $\boldsymbol{\beta} \to \infty$). Nevertheless, in the same section we prove some partial results in this direction for the soliton addition map;

- The map $(u, \mathbf{z}, \boldsymbol{\beta}) \to B_N^+(u, \mathbf{z}, \boldsymbol{\beta})$ is uniformly smooth on compact sets in $\boldsymbol{\beta}$.
- The same map is uniformly smooth in $(u, \boldsymbol{\beta})$

We also establish a uniform invertibility result for the soliton addition map with respect to its first argument, as follows:

Theorem 1.8. *The set of potentials $M_{\mathbf{z},\boldsymbol{\beta}}$ with spectral values $(\mathbf{z}, \boldsymbol{\beta})$ is a manifold of codimension $4N$. It is uniformly smooth in an ε neighborhood of $\mathbf{M}_N$, and uniformly transversal to $\mathbf{M}_N$.*

1.9. **An outline of the paper.** This paper aims to accomplish several goals, all having to do with multisoliton states and their perturbations in the completely integrable NLS and mKdV flows:

(1) Understand the structure of multisoliton states, with emphasis on close or multiple spectral parameters.
(2) Understand the geometry (smoothness and uniformity) of the multisoliton manifold.
(3) Study nearby states and their evolution using the (multi)soliton addition and removal maps.
(4) Prove uniform orbital stability of the multisoliton manifold.

Here we provide a road map for the reader:

1. The Lax operator $\mathcal{L}$ and the associated spectral problem. The goal of the next section is to provide an overview of this spectral problem, in particular the left and right Jost functions and more generally their linear combinations, which are called *wave functions*. These are in turn used to define the transmission coefficient $T(z)$ as a meromorphic function in the upper half-space. Finally, the transmission coefficient T has also played a key role in the construction of the continuous family of conservation laws; we provide an overview of these as well.

2. The Bäcklund transform. This allows one to add or remove a soliton with given spectral and scattering parameters (z, κ) to/from an existing state $u \in H^s$. It is, in turn, constructed using the wave functions for Lax wave operator. Section 3 is devoted to the Bäcklund transform, both in the standard form and in an extended form; the latter is needed in order to better understand its symmetry properties. A new notion we introduce here is that of analytic families of wave functions, and their associated Bäcklund transform.

3. The multisoliton addition and removal maps. These allow one to add/subtract an N-soliton to/from a given state, and are obtained by iterating Bäcklund transforms. They are classically defined relative to solitons with distinct spectral parameters, and are described in Section 4.

4. The smooth parametrization of the multisoliton addition and removal maps. Viewed as functions of the spectral and scattering parameters for N-solitons, the addition and removal maps are singular at the diagonal, near multiple spectral parameters. A key result of this paper is that this is a singularity of the parametrization, which can be removed by making a better choice for the parametrization of the sets of joint spectral and scattering parameters. This yields a smooth extension of the soliton addition and removal maps to solitons with higher multiplicity spectral parameters, and is naturally done using analytic families of wave functions. We introduce our extended spectral/scattering parameters, which in particular provide the desired reparametrization of multisoliton states in Section 6. The smoothness of both the multisoliton addition and removal maps is proved in the next section.

5. The uniformity question. Ideally, one would like both the multisoliton addition and removal maps to be uniformly smooth when restricted to spectral parameters in a compact subset of the upper half-space. One major difficulty we encounter is that our smooth parametrization of spectral/scattering parameters, although natural, is not uniform. Nevertheless, in Section 6 we are able to prove several partial uniformity results. These in particular lead us to the proof of one of our main results, namely the uniform regularity of the multisoliton manifold.

6. The structure of multisolitons. One can think of multisoliton states as a collection of bump functions, with exponential decay away from these bumps. In Section 7 we show that if sufficiently separated, these bump functions are exponentially close to lower rank multisolitons. This also leads to a good local uniform description of the multisoliton manifold $\mathbf{M}_N$ as an approximate sum of $\mathbf{M}_{N_j}$'s in the single bump regime, where our parametrization is uniform.

7. Uniform orbital stability of multisolitons. Section 8 contains the proof of the stability result. This relies heavily on our previous results on the uniformity properties for the soliton addition and removal maps, as well as on the conserved energies developed in our prior work.

8. 2*-solitons: a case study.* The aim of the last section of the paper is to provide a complete analysis for the case of 2-solitons. In particular we accurately describe both the 2-soliton manifold, as well as the NLS and mKdV flows on this manifold. This serves to both illustrate the concepts introduced in the rest of the paper, as well as to provide a full analysis of the interaction patterns of two solitons, both for the NLS and for the mKdV flows. Multiple pictures are also provided.

1.10. **Acknowledgements.** The first author was supported by the DFG through the SFB 611. The second author was supported by the NSF grant DMS-1800294 as well as by a Simons Investigator grant from the Simons Foundation.

2. An overview of the scattering transform

2.1. **Lax pair, Jost solutions and scattering transform.** Here we recall some basic facts about the inverse scattering transform for NLS and mKdV. Both the NLS evolution (1.1) and the mKdV evolution (1.2) are completely integrable, so we have at our disposal the inverse scattering transform conjugating the nonlinear flow to the corresponding linear flow. To describe their Lax pairs we consider the system

$$\begin{aligned} \psi_x &= \begin{pmatrix} -iz & u \\ -\bar{u} & iz \end{pmatrix} \psi \\ \psi_t &= i \begin{pmatrix} -[2z^2 - |u|^2] & -2izu + u_x \\ +2iz\bar{u} + \bar{u}_x & 2z^2 - |u|^2 \end{pmatrix} \psi, \end{aligned} \tag{2.1}$$

where z is a complex parameter. The focusing NLS equation arises as a compatibility condition for the system (2.1): For fixed z there exist two unique solutions ψ_1, ψ_2 to (2.1) with $\psi_1(0,0) = (1,0)$ and $\psi_2(0,0) = (0,1)$ if and only if u satisfies the nonlinear Schrödinger equation. The above is often referred to in the literature as the Lax pair for NLS.

If instead we want the canonical form $\mathcal{L}, \mathcal{P}$ with

$$\mathcal{L}_t = [\mathcal{P}, \mathcal{L}],$$

then we should view the first equation above as $\mathcal{L}\psi = z\psi$ where

$$\mathcal{L} = i \begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix}$$

and $\mathcal{P}$ is given by the second matrix in (2.1) where z has been eliminated using the relations $\mathcal{L}\psi = z\psi$,

$$\mathcal{P} = i \begin{pmatrix} 2\partial_x^2 + |u|^2 & -u\partial_x - \partial_x u \\ -\bar{u}\partial_x - \partial_x \bar{u} & -2\partial_x^2 - |u|^2 \end{pmatrix}.$$

This is equivalent to the pair of Kappeler and Grebert [17]. Much of this formalism can be found in the seminal paper by Ablowitz, Kaup, Newell and Segur [1]. The Lax operator $\mathcal{L}$ is the same for mKdV and for all the other commuting flows. It is only the operator $\mathcal{P}$ that will change.

The scattering transform associated to both the focusing NLS and the focusing mKdV is defined via the first equation of (2.1) which we write as linear system

$$\begin{cases} \dfrac{d\psi^1}{dx} = -iz\psi_1 + u\psi^2 \\[2ex] \dfrac{d\psi^2}{dx} = iz\psi^2 - \bar{u}\psi^1. \end{cases} \tag{2.2}$$

One part of the scattering data for this problem is obtained for $z = \xi$, real, by considering the relation between the asymptotics for ψ at $\pm\infty$. Precisely, one considers the Jost solutions ψ_l and ψ_r with asymptotics

$$\psi_l(\xi, x, t) = \begin{pmatrix} e^{-i\xi x} \\ 0 \end{pmatrix} + o(1) \quad \text{as } x \to -\infty, \quad \psi_l(\xi, x, t) = \begin{pmatrix} T^{-1}(\xi) e^{-i\xi x} \\ R(t,\xi) T^{-1}(\xi) e^{i\xi x} \end{pmatrix} + o(1) \quad \text{as } x \to \infty,$$

respectively

$$\psi_r(\xi, x, t) = \begin{pmatrix} L(t,\xi) T^{-1}(\xi) e^{-i\xi x} \\ T^{-1}(\xi) e^{i\xi x} \end{pmatrix} + o(1) \quad \text{as } x \to -\infty, \quad \psi_l(\xi, x, t) = \begin{pmatrix} 0 \\ e^{i\xi x} \end{pmatrix} + o(1) \quad \text{as } x \to \infty.$$

These are viewed as initial value problems with data at $-\infty$, respectively $+\infty$. We note that the T's in the two solutions ψ_l and ψ_r are the same since the Wronskian of the two solutions is constant:

$$\det(\psi_l, \psi_r) \to T^{-1}(\xi) \qquad \text{for } x \to \pm\infty.$$

The quantity $|\psi_1|^2 + |\psi_2|^2$ is also conserved, which shows that on the real line we have

$$|T| \geq 1, \qquad |T|^2 = 1 + |R|^2 = 1 + |L|^2.$$

Further, we have the symmetry $(\psi^1, \psi^2) \to (\bar{\psi}^2, -\bar{\psi}^1)$ which via the Wronskian leads to

$$L\bar{T} = \bar{R}T.$$

It is an immediate consequence of the existence of the Lax pair that as u evolves along the NLS flow (1.1), the functions L, R, T evolve according to

$$T_t = 0, \qquad L_t = -4i\xi^2 L, \qquad R_t = 4i\xi^2 R, \tag{2.3}$$

if u evolves according to the mKdV flow (1.2) then

$$T_t = 0, \qquad L_t = -8i\xi^3 L, \qquad R_t = 8i\xi^3 R, \tag{2.4}$$

and if z evolves according to the nth flow

$$T_t = 0, \qquad L_t = -i(2\xi)^n L, \qquad R_t = i(2\xi)^n R. \tag{2.5}$$

Thus one part of scattering map for u is given by

$$u \to R,$$

which maps the NLS flow (1.1) to the (Fourier transform of) the linear Schrödinger evolution, and simultaneously the mKdV flow to the linear Airy flow.

More generally, for any z in the closed upper half plane there exist the Jost solutions

$$\psi_l(\xi, x, t) = \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix} + o(1)e^{\operatorname{Im} zx} \quad \text{as } x \to -\infty,$$

$$\psi_l(\xi, x, t) = \begin{pmatrix} T^{-1}(z)e^{-izx} \\ 0 \end{pmatrix} + o(1)e^{\operatorname{Im} zx} \quad \text{as } x \to \infty,$$

This provides a holomorphic extension of T^{-1} to the upper half-space, and thus a meromorphic extension for T. Here T may have poles in the upper half-space, which correspond to non-real eigenvalues of $\mathcal{L}$. The poles of T must be isolated in the open upper half space, though they can accumulate on the real line.

For data u for which T is holomorphic in the upper half-space, the scattering data is fully described by the reflection coefficient R. If instead T is merely meromorphic, then the scattering data involves not only the function R on the real line, but also at least the singular part of the Laurent series of T at the poles. However, this still does not fully describe the problem, as by the results of Zhou [32], T may have poles in the upper half space accumulating at the real axis even for Schwartz functions u.

There is, however, one redeeming feature: All such poles are localized in a strip near the real axis if $u \in L^2$, and more generally in a polynomial neighbourhood $0 \le \operatorname{Im} z \lesssim_{\|u\|_{H^s}} (1 + |\operatorname{Re} z|)^{-2s}$ of the real line if $u \in H^s$ with $-1/2 < s < 0$. In the limiting case $s = -1/2$, smallness of u in $l^2 DU^2$ guarantees the localization of the poles in $0 \le \operatorname{Im} z \ll (1 + |\operatorname{Re} z|)$.

A key difference between real and nonreal z is that for real z, one essentially needs $u \in L^1$ in order to define the scattering data $L(\xi)$ and $T(\xi)$ in a pointwise fashion. This restricts the use of the inverse scattering transform to localized, rather than L^2 data. On the other hand, for z in the open upper half space it suffices to have some L^2 type bound on u in order to define $T(z)$.

Reconstructing u from the scattering data requires solving a Riemann-Hilbert problem, see [13] for this approach for the modified Korteweg-de Vries equation, as well as [8] for the focusing cubic NLS equation.

2.2. **Symmetries.** The main symmetries are multiplication by a phase, translations in x, modulations resp. translations in frequency, and scaling. We define them simultaneously on distributions by

$$f \to U_{\theta,\xi_0,x_0,\lambda} f = e^{i\theta + ix\xi} \lambda f(\lambda(x - x_0)).$$

This fixes a representation on a central extension of the Heisenberg group, a notion which we do not use in the sequel. On the Fourier side

$$\mathcal{F}\Big(e^{i\theta + ix\xi_0} \lambda f(\lambda(x - x_0))\Big)(\xi) = e^{i(\theta + x_0\xi_0)} e^{-ix_0\xi} \hat{f}(\lambda^{-1}(\xi - \xi_0)).$$

We compute the effect of the symmetries on the Lax operator and z waves:

$$i\begin{pmatrix} \partial_x & -e^{i\theta}u \\ \overline{-e^{i\theta}u} & -\partial_x \end{pmatrix} = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta t/2} \end{pmatrix} i\begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix}\begin{pmatrix} e^{-i\theta t/2} & 0 \\ 0 & e^{i\theta t/2} \end{pmatrix}.$$

Let $V(h)f = f(x-h)$. Then

$$i\begin{pmatrix} \partial_x & -V(h)u \\ \overline{-V(h)u} & -\partial_x \end{pmatrix} = V(h)i\begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix} V(-h).$$

Also,

$$i\begin{pmatrix} \partial_x & -e^{ix\xi}u \\ \overline{-e^{-ix\xi}u} & -\partial_x \end{pmatrix} = \begin{pmatrix} e^{ix\xi/2} & 0 \\ 0 & e^{-ix\xi/2} \end{pmatrix}\left[i\begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix} - \begin{pmatrix} \xi/2 & 0 \\ 0 & \xi/2 \end{pmatrix}\right]\begin{pmatrix} e^{-ix\xi/2} & 0 \\ 0 & e^{ix\xi/2} \end{pmatrix},$$

and with $R(\lambda)f(x) = f(\lambda x)$, the scaling by $\lambda > 0$ acts as follows:

$$i\begin{pmatrix} \partial_x & -\lambda u(\lambda x) \\ \overline{-\lambda u(\lambda x)} & -\partial_x \end{pmatrix} = i\lambda R_\lambda \begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix} R_{\lambda^{-1}}.$$

We obtain

Lemma 2.1. *We define*

$$\tilde{U}_{\theta,\xi,x_0,\lambda}\begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \lambda^{-1/2}\begin{pmatrix} e^{-i(\theta+\xi x_0)/2}\psi^1(\lambda^{-1}(.-x_0)) \\ e^{i(\theta+\xi x_0)/2}\psi^2(\lambda^1(.-x_0)) \end{pmatrix}.$$

Then

$$i\begin{pmatrix} \partial_x & -U_{\theta,\xi_0,x_0,\lambda}u \\ \overline{-\lambda U_{\theta,\xi_0,x_0,\lambda}u(\lambda x)} & -\partial_x \end{pmatrix}\Psi = i\begin{pmatrix} \partial_x & -u \\ -\bar{u} & -\partial_x \end{pmatrix}\tilde{U}_{\theta,\xi_0,x_0,\lambda}\Psi - \xi/2\tilde{U}_{\theta,\xi_0,x_0,\lambda}\Psi.$$

Moreover

$$T(U_{\theta,\xi_0,x_0,\lambda}u, z) = T(u, \lambda^{-1}(z-\xi_0/2)) \tag{2.6}$$

The last equation expresses the mismatch between transformations of Fourier variables and spectral variables.

2.3. **The transmission coefficient in the upper half-plane and conservation laws.** Our construction of fractional Sobolev conserved quantities in [23] relies essentially on the fact that the transmission coefficient T is preserved along both the NLS and mKdV flows. In principle this gives us immediate access to infinitely many conservation laws, but the question is whether one can relate (some of) them nicely to the standard scale of Sobolev spaces.

If u is a Schwartz function then $\ln|T|$ is a Schwartz function on the real line, and has a Taylor expansion

$$\ln T(z) \approx \frac{1}{2\pi i}\sum_{j=0}^{\infty} H_j(2z)^{-j-1}. \tag{2.7}$$

If T has no poles ith the upper half-space then by the residue theorem the conserved energies H_j can be expressed in terms of the values of T on the real axis,

$$H_k = \int \xi^k \ln|T(-\xi/2)|d\xi. \tag{2.8}$$

However, $\ln T$ may have poles in the upper half plane, and the right hand side in the formula (2.8) above has to be modified to account for the residues at the poles. Precisely, if the poles of T are located at z_j with multiplicities m_j then the counterpart of the relation (2.8) is

$$H_k = \int \xi^k \ln |T(-\xi/2)| d\xi + 2 \sum_j \frac{1}{k+1} m_j \operatorname{Im}(2z_j)^{k+1}. \tag{2.9}$$

This is clear if T has finitely many poles away from the real line, but can also be justified in general by interpreting the trace of $\ln |T|$ on the real line as a non-negative measure.

More generally, for any function $\eta : \mathbb{R} \to \mathbb{R}$ the expression

$$\int \eta(\xi) \ln |T(-\xi/2)| \, d\xi$$

is formally conserved.

Thus a natural candidate for a fractional Sobolev conservation law may be obtained by choosing any (real) function η so that

$$\eta(\xi) \approx (1 + \xi^2)^s.$$

However, there are two issues with such a general choice. First, it is quite difficult to get precise estimates for $\log |T|$ on the real line without assuming any integrability condition on u. Secondly, in the focusing case such a choice would still miss the poles of the transmission coefficient.

To remedy both of these issues, it is natural to use much more precise real weights which have a holomorphic extension at least in a strip around the real line. Our choice in [23] was to use the weights

$$\eta_s(\xi) = (1 + \xi^2)^s, \qquad s > -\frac{1}{2}.$$

which not only have the appropriate size on the real axis, but can also be extended as holomorphic functions to the subdomain $D = U \setminus i[1, \infty)$ of the upper half-space U.

In the absence of poles for T in the upper half-space one can formally define the conserved energies by

$$E_s(u) = \int (1 + \xi^2)^s \operatorname{Re} \ln T(\xi/2) d\xi.$$

By Cauchy's theorem the integral can be switched to the half-line $i[1, \infty)$ to give

$$E_s(u) = 4 \sin(\pi s) \int_1^\infty (\tau^2 - 1)^s \Big[-\operatorname{Re} \ln T(i\tau/2) + \frac{1}{2\pi} \sum_{j=0}^N (-1)^j H_{2j} \tau^{-2j-1} \Big] d\tau + \sum_{j=0}^N \binom{s}{j} H_{2j} \tag{2.10}$$

Here the conserved integer energies H_{2j} are used to remove the leading terms in the expansion at infinity, which is needed in order to insure the absolute convergence of the integral.

One key advantage to switching the integral into the upper halfspace is that the transmission coefficient is more robust there, depending only on Sobolev norms of u. For this reason, in [23] we adopt the formula (2.10) as the definition of the conserved energy E_s.

This works also in the case when the transmision coefficient T has poles in the upper half-space, Then T may have only finitely many poles on the half-line $i[1, \infty)$. We also note the role played by the smallness condition for u in $l^2 DU^2$, which is present in Theorem 1.1. This guarantees that T has a convergent multilinear expansion on the half-line $i[1, \infty)$, and in particular has no poles there.

3. The Bäcklund transform

The central object in this section is the intertwining operator, which in the form we consider is related to the one in the work of Cascaval, Gesztesy, Helge and Latushkin [9]. For related ideas and formulas we also refer the reader to more recent works [26], [2]. We also heavily exploit complex differentiability here, which brings in the tools of complex analysis. Unfortunately the dependence on the state u is not holomorphic, since the complex conjugate occurs in the Lax operator. To rectify this, it turns out to be useful to relax the relation between the off-diagonal entries of the Lax operator and to consider the generalized spectral problem

$$\psi_x = \begin{pmatrix} -iz & u_1 \\ -u_2 & iz \end{pmatrix} \psi. \tag{3.1}$$

3.1. **Regularity of Jost solutions.** We characterize the regularity of the Jost functions in the following summary of results of [23], where, for the left Jost function, we solve the system of integral equations

$$\begin{aligned} \phi^1(x) =& 1 + \int_{-\infty}^x u_1(y)\phi^2(y)dy, \\ \phi^2(x) =& \int_{-\infty}^x e^{2iz(x-y)} u_2(y)\phi^1 dy. \end{aligned}$$

for the renormalized functions

$$(\phi^1, \phi^2) = e^{izx}(\psi_l^1, \psi_l^2).$$

Lemma 3.1. *Let* $\mathbf{u} = (u_1, u_2) \in H^s$, $s > -\frac{1}{2}$, *and* $\operatorname{Im} z > 0$. *Then the left Jost function* ψ_l *satisfies*

$$(e^{izx}\psi_l)' \in H^s, \qquad e^{izx}\psi_l^2 \in H^{s+1},$$

$$\lim_{x\to-\infty} e^{izx}\psi_l = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\lim_{x\to\infty} e^{izx}\psi_l^1 = T^{-1}(z).$$

Moreover,

$$\|(e^{izx}\psi_l^1)'\|_{H^s} + \|e^{izx}\psi_l^2\|_{H^{s+1}} \le c\|(u_1, u_2)\|_{l^2DU^2}(\|u_1\|_{H^s} + \|u_2\|_{H^s}).$$

The map $(u_1, u_2, z) \to e^{izx}\psi_l$ *is holomorphic in* z, u_1 *and* u_2 *with all derivatives bounded by*

$$c(\|(u_1, u_2)\|_{l^2DU^2})(1 + \|u_1\|_{H^s} + \|u_2\|_{H^s})$$

for z *in a compact region in the upper half plane. The differential of* $e^{izx}\psi_l$ *at* $\mathbf{u} = 0$ *is given by*

$$\begin{pmatrix} 0 \\ \int_{-\infty}^x e^{2iz(x-y)} u_2(y)dy \end{pmatrix},$$

and the differential of $e^{-izx}\psi_r$ *by*

$$\begin{pmatrix} \int_x^\infty e^{-2iz(x-y)} u_1(y)dy \\ 0 \end{pmatrix}.$$

The map

$$l^2DU^2 \times l^2DU^2 \ni (u_1, u_2, z) \to 1/T(z) = W(\psi_l, \psi_r) \in \mathbb{C}$$

is holomorphic with derivatives bounded by $C(\|u_1\|_{l^2DU^2}, \|u_2\|_{l^2DU^2})$ *for* z *in a compact domain of the upper half plane. The expansion of* T^{-1} *at* $\mathbf{u} = 0$ *is given by*

$$T^{-1}(z) = 1 - \int_{x<y} e^{-2iz(x-y)} u_1(y)u_2(x)dydx + O(\|(e^{2i\operatorname{Re} zx}u_1, e^{-2i\operatorname{Re} zx}u_2)\|^4_{l^2_{\operatorname{Im} z}DU^2}). \tag{3.2}$$

3.2. **The spectrum of the Lax operator, wave functions and eigenfunctions.** We consider the scattering transform for the focusing nonlinear Schrödinger equation. The first equation of the zero curvature formulation is

$$\psi_x = \begin{pmatrix} -iz & u \\ -\bar{u} & iz \end{pmatrix} \psi. \tag{3.3}$$

Here we take $u \in l^2 DU^2$. Its transmission coefficient is a meromorphic function T in $\mathbb{C} \setminus \mathbb{R}$, with singularities (poles) at eigenvalues z of the Lax operator

$$\psi \to \mathcal{L}\psi = \mathcal{L}(u)\psi = i \begin{pmatrix} \partial & -u \\ -\bar{u} & -\partial \end{pmatrix} \psi, \tag{3.4}$$

or equivalently, if there is an L^2 function ψ which satisfies (3.1). The operator $\mathcal{L}$ is not selfadjoint, however it satisfies the conjugation relation

$$\mathcal{L}^* = M_0 \mathcal{L} M_0^{-1}, \qquad M_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.5}$$

Hence if z is an eigenvalue for L with eigenfunction ϕ, then it is also an eigenvalue for $\mathcal{L}^*$ with eigenfunction $M_0\phi$.

On the other hand, by conjugation it follows that $\bar{z}$ is an eigenvalue for both $\mathcal{L}$ and $\mathcal{L}^*$, with eigenfunctions $M\bar{\phi}$, respectively $M_0 M\bar{\phi}$, where

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.6}$$

We recall that if $\|u\|_{l^2 DU^2}$ is small then the eigenvalues all satisfy

$$\operatorname{Im} z \leq \varepsilon \langle \operatorname{Re} z \rangle,$$

see Corollary 5.11 of [23], or, by (3.2) and

$$\|e^{-i \operatorname{Re} z x} u\|_{l^2_{\operatorname{Im} z} DU^2} \lesssim \frac{\langle \operatorname{Re} z \rangle}{\operatorname{Im} z} \|u\|_{l^2_{\operatorname{Im} z} DU^2} \lesssim \frac{\langle \operatorname{Re} z \rangle}{\operatorname{Im} z} \langle 1/\operatorname{Im} z \rangle^{1/2} \|u\|_{l^2_1 DU^2}.$$

Here and later we use the standard notation $\langle x \rangle = \sqrt{1+x^2}$.

For z in the upper half-space we seek to describe solutions to $(\mathcal{L}(u) - z)\phi = 0$, which form a two dimensional vector space; these are called *wave functions*. For this it is convenient to use the left and right Jost functions ψ_l, ψ_r which decay exponentially at $-\infty$, respectively $+\infty$. The transmission coefficient is defined to be the meromorphic function in the upper half plane given by the inverse of their Wronskian,

$$T(z) = (\det(\psi_l, \psi_r))^{-1}.$$

Then we distinguish two scenarios:

- z is not an eigenvalue. Then ψ_l and ψ_r are linearly independent, and form a basis in the space of solutions.
- z is an eigenvalue. Then ψ_l and ψ_r are linearly dependent, and both are eigenfunctions.

In the first case we want to parametrize the wave functions which are unbounded as $x \to \pm\infty$, up to the multiplication by a complex number. We parametrize them by

$$\psi = T(z) \left(e^{-i(\beta_0 + \beta_1 z)} \psi_l + e^{i(\beta_0 + \beta_1 z)} \psi_r \right), \tag{3.7}$$

where $\beta_0 \in \mathbb{R}/\pi\mathbb{Z}$ and $\beta_1 \in \mathbb{R}$.

We can also relate this notation to our notation κ for scattering parameters for solitons, by setting

$$i(\beta_0 + \beta_1 z) = \kappa = \operatorname{Im} z x_0 + i\theta. \tag{3.8}$$

Here $x_0 \in \mathbb{R}$ and $\theta \in \mathbb{R}/\pi\mathbb{Z}$ can be thought of as the center point and the phase associated to ψ, and κ will be naturally interpreted later on as a scattering parameter in the context of the Bäcklund transform.

Moving u along the NLS flow corresponds to moving ψ along the $\mathcal{P}$ flow. It is not difficult to determine the dependence on time of the unbounded wave function parameters x_0 and θ when we evolve wave functions along the $\mathcal{P}$ flow. We recall that the leading part of $\mathcal{P}$ is $\begin{pmatrix} 2i\partial_x^2 & 0 \\ 0 & -2i\partial_x^2 \end{pmatrix}$ and hence, for the solution to

$$\psi_t = \mathcal{P}\psi,$$

the leading term near $-\infty$ is

$$e^{2itz^2+i(\beta_0+\beta_1 z)} \begin{pmatrix} 0 \\ e^{izx} \end{pmatrix} = e^{\operatorname{Im} z(x_0-4t\operatorname{Re} z)+i(\theta+2t(\operatorname{Re}^2 z-\operatorname{Im}^2 z))} \begin{pmatrix} 0 \\ e^{izx} \end{pmatrix},$$

and on the right it is

$$e^{-2iz^2t-i(\beta_0+\beta_1 z)} \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix} = e^{-\operatorname{Im} z(x_0-4t\operatorname{Re} z)-i(\theta+2t(\operatorname{Re}^2 z-\operatorname{Im}^2 z))} \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix}.$$

Thus

$$x_0(t) = x_0 - 4\operatorname{Re} zt, \qquad \theta(t) = \theta + 2(|\operatorname{Re} z|^2 - |\operatorname{Im} z|^2)t. \tag{3.9}$$

or with the complex notation, as for the pure soliton,

$$\kappa(t) = i(\beta_0 + \beta_1 z + 2tz^2). \tag{3.10}$$

A similar computation can be carried out for the mKdV flow, as well as for all of the other commuting flows.

Moreover, suppressing the time dependence for the rest of this section and setting $t = 0$, if ζ is neither real nor an eigenvalue then the inverse of $\mathcal{L}(u) - \zeta$ is given by

$$\begin{aligned}(\mathcal{L}(u)-\zeta)^{-1}f(x) =& T(\zeta)^{-1}\Big(\psi_r(x)\int_{-\infty}^x -\psi_l^2(y)f_1(y) + \psi_l^1(y)f_2(y)\,dy \\ &+ \psi_l(x)\int_x^\infty \psi_r^2(y)f_1(y) - \psi_r^1(y)f_2(y)\,dy\Big). \\ =& T(\zeta)^{-1}\left(\psi_l(x)\int_x^\infty M\psi_r\cdot f\,dy - \psi_r(x)\int_{-\infty}^x M\psi_l\cdot f\,dy\right).\end{aligned} \tag{3.11}$$

Similarly we normalize eigenfunctions ψ so that

$$\psi = -e^{-i(\beta_0+\beta_1 z)}\psi_l = e^{i(\beta_0+\beta_1 z)}\psi_r. \tag{3.12}$$

Together with (3.8), this allows one to understand κ, respectively x_0 and θ as scattering parameters, and we interpret heuristically "u contains a soliton with scale λ, modulation ξ, center x_0 and phase θ" as the statement that the Lax operator has an eigenvalue $z = -\xi/2 + i\lambda$ with scattering parameter κ given by x_0 and θ through (3.8). This will become more clear when we discuss the Bäcklund transform later on.

The multiplicity of eigenvalues is discussed next:

Lemma 3.2. *Suppose that ζ is an eigenvalue for $\mathcal{L}(u)$. Then the geometric multiplicity of ζ is* 1. *Let ψ be a ζ eigenfunction. Then the algebraic multiplicity of ζ is* 1 *if and only if*

$$\int \psi^1\psi^2\,dx \neq 0. \tag{3.13}$$

Further, we have

$$2i\int \psi^1\psi^2 dx = \frac{d}{dz}T^{-1}(\zeta). \tag{3.14}$$

Proof. If the geometric multiplicity were 2 then all solutions to (3.1) were bounded, and hence would decay exponentially at $\pm\infty$. This contradicts the fact that near ∞ there is one characteristic exponent with positive real part. The eigenvalue is simple if the equation

$$\mathcal{L}(u)\phi - \zeta\phi = \psi$$

is not solvable in L^2. Since, by the Fredholm alternative,

$$(\mathcal{L}^* - \bar\zeta)MM_0\bar\psi = 0,$$

the above equation is not solvable iff

$$\int \psi^1\psi^2 dx \neq 0.$$

To verify (3.14) we differentiate the system (3.1) with respect to the parameter z where $\psi = \psi_l$ is the left Jost function. Denoting $\tilde\psi = \frac{d}{dz}\psi_l$, it solves the system

$$\tilde\psi_x = \begin{pmatrix} -iz & u \\ -\bar u & iz \end{pmatrix}\tilde\psi - iM_0\psi_l,$$

with initial and terminal data

$$\tilde\psi(-\infty) = -ixe^{-izx}\begin{pmatrix}1\\0\end{pmatrix}, \qquad \tilde\psi(\infty) = e^{-izx}\begin{pmatrix}\partial_z T^{-1}(z)\\0\end{pmatrix},$$

since z is an eigenvalue. We recall that

$$T(z)^{-1} = W(\psi_l, \psi_r)$$

Then the relation (3.14) is obtained from

$$\lim_{x\to-\infty} W(\psi_l, \tilde\psi) = 0, \quad \lim_{x\to\infty} W(\psi_l, \tilde\psi) = \partial_z T^{-1}(z),$$

and

$$\partial_x W(\psi_l(x), \tilde\psi(x)) = 2i\psi_l^1\psi_l^2$$

by the fundamental theorem of calculus. □

Unbounded wave functions will play a crucial role also in the case when z is an eigenvalue. Suppose now that ϕ is an eigenfunction to the eigenvalue z of $L(u)$. If the wave function ψ to the same eigenvalue z is unbounded on one side then the same is true on the other side, and ϕ and ψ are a fundamental system. We may normalize ψ so that

$$\psi \sim e^{-\operatorname{Im} zx_0 - i\theta}e^{-izx}\begin{pmatrix}1\\0\end{pmatrix}$$

as $x\to\infty$ and

$$\psi \sim e^{\operatorname{Im} zx_0 + i\theta}e^{izx}\begin{pmatrix}0\\1\end{pmatrix}$$

as $x\to-\infty$. In contrast to the previous case (when z is not an eigenvalue), here x_0 and θ are uniquely determined, and we have the same normalization for all unbounded wave functions. With this convention

$$W(\phi, \psi) = 1, \tag{3.15}$$

and x_0 and θ are the same for both.

Lemma 3.3. *Suppose that z is an eigenvalue with ϕ an eigenfunction and ψ an unbounded wave function. Then the limit*

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} e^{-\varepsilon(x-x_0)^2}(\phi^1\psi^2+\phi^2\psi^1)\,dx$$

exists. The general unbounded wave function is given by

$$\psi+\zeta\phi.$$

If z is simple then there is a unique wave function so that the limit is 0. The limit defines a bijection between unbounded wave functions and $\mathbb{C}$. If the eigenvalue has higher multiplicity then it does not depend on the unbounded wave function.

As a consequence we obtain a natural parametrization of unbounded wave functions in the case of a simple eigenvalue.

Proof. By Lemma 3.1 the limits

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{x_0} e^{-\varepsilon(x-x_0)^2}\phi^2\psi^1 dx$$

and

$$\lim_{\varepsilon\to 0}\int_{x_0}^{\infty} e^{-\varepsilon(x-x_0)^2}\phi^1\psi^2 dx$$

exist. Since also $W(\phi,\psi)=\phi^1\psi^2-\phi^2\psi^1=1$ we obtain

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}} e^{-\varepsilon(x-x_0)^2}\phi^1\psi^2+\phi^2\psi^1 dx = 2\lim_{\varepsilon\to 0}\left[\int_{-\infty}^{x_0} e^{-\varepsilon(x-x_0)^2}\phi^2\psi^1 dx+\int_{x_0}^{\infty} e^{-\varepsilon(x-x_0)^2}\phi^1\psi^2 dx\right].$$

□

Finding a natural parametrization of unbounded z waves is important in the sequel. We will obtain implicitly a natural parametrization also for higher multiplicity.

3.3. **The intertwining operator.** The main tool in understanding the Bäcklund transform is the intertwining operator $D(u)$. Given $u\in H^s(\mathbb{R})$ with $s>-1/2$ we recall that $\mathcal{L}(u)$ is the associated Lax operator. Let z be a point in the upper half plane, and ψ an unbounded z wave. Then $\tilde\psi=\begin{pmatrix}\bar\psi_2\\-\bar\psi_1\end{pmatrix}$ is a $\bar z$ wave. The intertwining operator is the unique operator of the form

$$D(u)=\mathcal{L}(u)+A(x) \tag{3.16}$$

where $A:\mathbb{R}\to\mathbb{C}^{2\times 2}$ is chosen so that D annihilates both ψ and $\tilde\psi$. The matrix A is uniquely determined by this requirement. It turns out that there is a unique function $v\in H^s(\mathbb{R})$ so that the intertwining relation

$$\mathcal{L}(v)D(u)=D(u)\mathcal{L}(u)$$

holds. The map

$$(u,z,\psi)\to v$$

is called the *Bäcklund transform*. The construction is remarkable. It can be iterated, it gives useful formulas for the addition of multiple solitons, it works for multiple eigenvalues and it can be inverted by an intertwining operator based on eigenfunctions instead of unbounded z waves.

We want to trace the dependence of multiple Bäcklund transforms on the data. For that it turns out to be useful to relax the relation between u and $\bar u$, z and $\bar z$, and ψ and $\tilde\psi$: we consider a Lax operator of the form

$$i\begin{pmatrix}\partial & -u_1\\ -u_2 & -\partial\end{pmatrix},$$

two different values $z_1, z_2 \in \mathbb{C}\backslash\mathbb{R}$, and associated z_j waves ψ_1 and ψ_2. We define the intertwining operators - this time on intervals - by the requirement that the intertwining operator is of the form (3.16) and it has both ψ_j's in its null space.

The crucial benefit of this extension is that the iterated Bäcklund transform is easily seen to be invariant under exchanging any set of indices, which immediately implies a regular dependence of the iterated Bäcklund transform on the elementary symmetric polynomials of the z_j in the NLS/mKdV case.

We consider a pair of function $\mathbf{u} = (u_1, u_2)$ and the corresponding Lax operator

$$\psi \to \mathcal{L}\psi = \mathcal{L}(\mathbf{u})\psi = i \begin{pmatrix} \partial & -u_1 \\ -u_2 & -\partial \end{pmatrix} \psi. \tag{3.17}$$

We define z waves in the same fashion as for the Lax operator in the remaining part of this section.

Definition 3.4. *We denote the Wronskian by $W(.,.)$. Let $\zeta_1 \neq \zeta_2 \in \mathbb{C}\backslash\mathbb{R}$ and let ψ_j be ζ_j- wave functions associated to $\mathbf{u}$ and $I \subset \mathbb{R}$ an open set so that $W(\psi_1, \psi_2) \neq 0$. We define the intertwining operator on I by*

$$D\psi = D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})\psi = \big(\mathcal{L}(\mathbf{u}) - \zeta_2\big)\psi + (\zeta_2 - \zeta_1)\frac{W(\psi_2, \psi)}{W(\psi_2, \psi_1)}\psi_1. \tag{3.18}$$

It is not difficult to determine the kernel of this operator if I is an interval.

Lemma 3.5. *Let $\mathbf{u}$, $\boldsymbol{\zeta}$, $\boldsymbol{\psi}$ and I as above. Then*

$$D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})\psi_j = 0. \tag{3.19}$$

The operator $\mathcal{D}$ is symmetric under exchanging the indices,

$$D(\mathbf{u}, (\zeta_2, \zeta_1), (\psi_2, \psi_1)) = D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}). \tag{3.20}$$

Proof. It is easy to see that ψ_1 is in the null space. Assuming (3.20) we can argue in the same way for ψ_2. We turn to the proof of (3.20) and use the trilinear algebraic identity

$$W(\psi_1, \psi_2)\psi_3 + W(\psi_2, \psi_3)\psi_1 + W(\psi_3, \psi_1)\psi_2 = 0. \tag{3.21}$$

It implies

$$\zeta_2 W(\psi_1, \psi_2)\psi + (\zeta_2 - \zeta_1)W(\psi_2, \psi)\psi_1 = -\zeta_1 W(\psi_1, \psi_2)\psi + (\zeta_2 - \zeta_1)W(\psi_1, \psi)\psi_2.$$

We divide by $W(\psi_1, \psi_2)$ to obtain (3.20). □

The next construction is a crucial piece of the puzzle. Given $(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})$ as above, we search for a function $\mathbf{v}$ so that the following property holds

$$\mathcal{L}(\mathbf{v})D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}) = D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})\mathcal{L}(\mathbf{u}). \tag{3.22}$$

For such a $\mathbf{v}$ we will use the notation $\mathbf{v} = B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})$.

Both sides are second order operators with identical second order terms. We rewrite both sides of (3.22) as

$$(\mathcal{L}(\mathbf{u}))^2 + A_1\mathcal{L}(\mathbf{u}) + A_0 = (\mathcal{L}(\mathbf{u}))^2 + B_1\mathcal{L}(\mathbf{u}) + B_0.$$

where

$$\begin{aligned} A_1 - B_1 &= i\begin{pmatrix} 0 & u_1 - v_1 \\ u_2 - v_2 & 0 \end{pmatrix} + \frac{\zeta_2 - \zeta_1}{W(\psi_2, \psi_1)}\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -\psi_2^2\psi_1^1 & \psi_2^1\psi_1^1 \\ -\psi_2^2\psi_1^2 & \psi_2^1\psi_1^2 \end{pmatrix}\right]\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= i\begin{pmatrix} 0 & u_1 - v_1 - 2i\dfrac{\zeta_2 - \zeta_1}{W(\psi_2, \psi_1)}\psi_2^1\psi_1^1 \\ u_2 - v_2 + 2i\dfrac{\zeta_2 - \zeta_1}{W(\psi_2, \psi_1)}\psi_2^2\psi_1^2 & 0 \end{pmatrix} \end{aligned}$$

Thus $A_1 = B_1$ is equivalent to

$$\mathbf{v} := B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}) := \mathbf{u} + i\frac{2(\zeta_2 - \zeta_1)}{W(\psi_2, \psi_1)} \begin{pmatrix} \psi_1^1 \psi_2^1 \\ -\psi_1^2 \psi_2^2 \end{pmatrix}. \tag{3.23}$$

With this choice we see, using Lemma 3.5, that ψ_j, $j = 1, 2$ are in the null space of the both sides. Thus $A_0 = B_0$, and we have proved the intertwining relation (3.16). This computation motivates the following:

Definition 3.6. *We define the Bäcklund operator B by* (3.23)*.*

It also leads to the next result:

Theorem 3.7. *a) $D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})$ maps z waves of $\mathcal{L}(\mathbf{u})$ to z waves of $\mathcal{L}(B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}))$.*
b) We have

$$\mathcal{L}(B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})) \frac{1}{W(\psi_2, \psi_1)} \psi_1 = \zeta_2 \frac{1}{W(\psi_2, \psi_1)} \psi_1. \tag{3.24}$$

c) For all functions $\mathbf{u}$, pairwise disjoint ζ_j and ζ_j-waves ψ_j, $j = 1, 2, 3, 4$, the commutation relation

$$\begin{aligned} & D\Big(B(\mathbf{u}, (\zeta_1, \zeta_2), (\psi_1, \psi_2)), (\zeta_3, \zeta_4), D(\mathbf{u}, (\zeta_1, \zeta_2), (\psi_1, \psi_2))(\psi_3, \psi_4)\Big) D(\mathbf{u}, (\zeta_1, \zeta_2), (\psi_1, \psi_2)) \\ & = D\Big(B(\mathbf{u}, (\zeta_1, \zeta_4), (\psi_1, \psi_4)), (\zeta_3, \zeta_2), D(\mathbf{u}, (\zeta_1, \zeta_4), (\psi_1, \psi_4))(\psi_3, \psi_2)\Big) D(\mathbf{u}, (\zeta_1, \zeta_4), (\psi_1, \psi_4)) \end{aligned} \tag{3.25}$$

holds, and hence the iterated Bäcklund transform is symmetric in all indices.

Proof. Part a) is an immediate consequence of (3.22). To see Part b) let ϕ be a ζ_2 wave. Then by the definition of the intertwining operator (3.18)

$$D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})\phi = (\zeta_2 - \zeta_1) \frac{W(\psi_2, \phi)}{W(\psi_2, \psi_1)} \psi_1,$$

where $W(\psi_2, \phi)$ is constant and zero iff ϕ is a multiple of ψ_2. We choose ϕ linearly independent from ψ_2. Then the right hand side does not vanish. By the intertwining property,

$$\begin{aligned} \mathcal{L}(B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})) \frac{1}{W(\psi_2, \psi_1)} \psi_1 & = \frac{1}{(\zeta_2 - \zeta_1) W(\psi_2, \phi)} \mathcal{L}(B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})) D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi})\phi \\ & = \frac{1}{(\zeta_2 - \zeta_1) W(\psi_2, \phi)} D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}) \mathcal{L}(u)\phi \\ & = \zeta_1 \frac{1}{W(\psi_2, \psi_1)} \psi_1. \end{aligned}$$

To prove the commutation relation (3.25) we observe that both sides are second order operators with the same leading part. All the ψ_j are in the null space, and hence they are the same. ☐

We can now obtain the following inversion result by a simple direct computation:

Lemma 3.8. *Assume that*

$$\mathbf{v} = B(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}),$$

and let

$$\tilde{\psi}_2 = \frac{1}{W(\psi_2, \psi_1)} \psi_1, \quad \tilde{\psi}_1 = \frac{1}{W(\psi_1, \psi_2)} \psi_2.$$

Then

$$\mathbf{u} = B(\mathbf{v}, \boldsymbol{\zeta}, \tilde{\boldsymbol{\psi}}) \tag{3.26}$$

and

$$D(\mathbf{v}, \boldsymbol{\zeta}, \tilde{\boldsymbol{\psi}}) D(\mathbf{u}, \boldsymbol{\zeta}, \boldsymbol{\psi}) = (\mathcal{L}(\mathbf{u}) - \zeta_1)(\mathcal{L}(\mathbf{u}) - \zeta_2). \tag{3.27}$$

Moreover,

$$D(\mathbf{v},\boldsymbol{\zeta},\tilde{\boldsymbol{\psi}})\phi = (\mathcal{L}(\mathbf{u}) - \zeta_2)\phi - 2(\zeta_2-\zeta_1)\frac{W(\psi_2,\phi)}{W(\psi_2,\psi_1)}\psi_1. \tag{3.28}$$

Proof. The identity (3.26) is a consequence of (3.24) in Theorem 3.7 and of the definition of $D(\mathbf{u},\boldsymbol{\zeta},\boldsymbol{\psi})$. Both sides of (3.27) map z waves of $\mathcal{L}(\mathbf{u})$ to z waves of the same operator $\mathcal{L}(\mathbf{u})$. The kernel of the right hand side is spanned by the ζ_1 and ζ_2 waves of $\mathcal{L}(\mathbf{u})$. Since every ζ_j wave is mapped by $D(\mathbf{u},\boldsymbol{\zeta},\boldsymbol{\psi})$ into the null space of $D(\mathbf{v},\tilde{\boldsymbol{\zeta}})$ they also span the null space of the left hand side. This implies the formula (3.27). Finally (3.28) is a consequence of (3.21). ☐

We remark that interchanging the roles of $\mathbf{u}$ and $\mathbf{v}$ in (3.27) yields the symmetric relation

$$D(\mathbf{u},\boldsymbol{\zeta},\boldsymbol{\psi})D(\mathbf{v},\boldsymbol{\zeta},\tilde{\boldsymbol{\psi}}) = (\mathcal{L}(\mathbf{v})-\zeta_1)(\mathcal{L}(\mathbf{v})-\zeta_2).$$

This in turn allows one to also interpret the Bäcklund transform as an instance of the double commutation method as it is described for instance in Deift [14].

We may iterate the Bäcklund transform as follows. Let $u_1,u_2 \in H^s(\mathbb{R})$, $s > -\frac{1}{2}$, ζ_{j1},ζ_{j2}, $j=\overline{1,N}$, pairwise disjoint complex numbers, and associated wave functions ψ_{j1},ψ_{j2} for $\mathcal{L}(\mathbf{u})$. On a set where $W(\psi_{11},\psi_{12}) \neq 0$ we apply the corresponding Bäcklund transform for $\mathbf{u}$ via (3.23), as well as transform the other wave functions by

$$\psi^1_{j1} = D\psi_{j1}, \qquad \psi^1_{j2} = D\psi_{j2}$$

for $j \geq 2$. Then we repeat the process N times.

By Theorem 3.7, the iterated Bäcklund transforms are symmetric in all the indices - of course on a set where all the Wronskians are nonzero. So it is natural to seek a direct description for them. To achieve that we start with the $N \times N$ matrix M with complex entries

$$M_{jk} = \frac{iW(\psi_{2k},\psi_{2j-1})}{\zeta_{2k}-\zeta_{2j-1}}. \tag{3.29}$$

We define the map

$$Q(\phi_1,\phi_2)(\psi) = W(\phi_2,\psi)\phi_1.$$

We assume that M is invertible and denote $m = M^{-1}$. Then we have the following:

Theorem 3.9. *The following properties hold for the iterated Bäcklund transform:*

(a) *The operator $D^N = D^N(\mathbf{u},\boldsymbol{\zeta},\boldsymbol{\psi})$ is given by*

$$D^N = \left(I + \sum_{j,k=1}^N m_{kj} Q(\psi_{2j-1},\psi_{2k})(\mathcal{L}(\mathbf{u})-\zeta_{2k})^{-1}\right)\prod_{k=1}^N(\mathcal{L}(\mathbf{u})-\zeta_{2k}). \tag{3.30}$$

(b) *The output function $\mathbf{v} = \mathbf{B}^N_+(\mathbf{u},\boldsymbol{\zeta},\boldsymbol{\psi})$ is given by*

$$\mathbf{v} = \mathbf{u} + 2\sum_{kj} m_{kj}\begin{pmatrix}\psi^1_{2j-1}\psi^1_{2k}\\ -\psi^2_{2j-1}\psi^2_{2k}\end{pmatrix}. \tag{3.31}$$

(c) *In particular, the image of a z-wave ψ for $\mathbf{u}$ is a z-wave $D^N\psi$ for $\mathbf{v}$ where*

$$D^N\psi = \prod_{\ell=1}^N(z-\zeta_{2\ell})\left(\phi + \sum_{j,k=1}^N \frac{1}{z-\zeta_{2k}} m_{kj} W(\psi_{2k},\phi)\psi_{2j+1}\right), \tag{3.32}$$

provided z is not equal to one of the ζ_{2j} - otherwise we swap the odd and even indices.

(d) The functions

$$\phi_{2k} = \sum_{j=1}^{N} m_{jk}\psi_{2j-1} \tag{3.33}$$

are z_{2k} waves for $\mathcal{L}(\mathbf{v})$, and similarly with odd and even indices swapped. We obtain the concise formula for the iterated Bäcklund transform

$$\mathbf{v} = \mathbf{u} + \sum_{k=1}^{N} \begin{pmatrix} \phi^1_{2k}\psi^1_{2k} \\ -\phi^2_{2k-1}\psi^2_{2k-1} \end{pmatrix}. \tag{3.34}$$

Proof. (a) We begin with the product formula, where we remark that the operator D^N is an order N nondegenerate differential operator acting on 2 vectors, therefore it admits a system of $2N$ fundamental solutions, and is uniquely determined by such a system. The iterated Bäcklund transform is another N-th order operator with the same coefficient of the leading term. The null space of $D(\mathbf{u}, (\zeta_{2l-1}, \zeta_{2l}), \psi_{2l-1}, \psi_{2l})$ is spanned by ψ_{2l-1} and ψ_{2l}. Therefore by the iteration relation (3.25) it follows that the functions ψ_j , $1 \le j \le 2N$ form a fundamental system for D^N. Hence, it remains to show that the expression in (3.30) vanishes when applied to ψ_j.

Indeed, we have

$$\begin{aligned} D^N\psi_{2\ell-1} &= \prod_{k=1}^{N}(z_{2\ell-1} - z_{2k})\Big(\psi_{2\ell-1} + \sum_{k,j=1}^{N} \frac{1}{z_{2\ell-1} - z_{2k}} m_{kj} W(\psi_{2k}, \psi_{2\ell-1})\psi_{2j-1}\Big) \\ &= \prod_{k=1}^{N}(z_{2\ell-1} - z_{2k})\left(\psi_{2\ell-1} - \sum_{k,j=1}^{N} m_{kj} M_{\ell k}\psi_{2j-1}\right) = 0, \end{aligned}$$

and

$$\begin{aligned} D^N\psi_{2\ell} &= \prod_{k\neq\ell}(z_{2k} - z_{2\ell})\Big[(\mathcal{L}(\mathbf{u}) - z_{2\ell})\psi_{2\ell} + \sum_{j=1}^{N} m_{j\ell} Q(\psi_{2\ell}, \psi_{2j-1})\psi_{2\ell}\Big] \\ &= \prod_{k\neq\ell}(z_{2k} - z_{2\ell})\sum_{j=1}^{N} m_{j\ell} W(\psi_{2\ell}, \psi_{2\ell})\psi_{2j-1} = 0 \end{aligned} \tag{3.35}$$

since the Wronskian vanishes.

(b) Next we verify the formula (3.31). For this we use the intertwining relation

$$\mathcal{L}(\mathbf{v})D^N = D^N\mathcal{L}(\mathbf{u}),$$

where for D^N we use (3.30). The expression on the right admits an expansion in terms of powers of $\mathcal{L}(\mathbf{u})$,

$$D^N\mathcal{L}(\mathbf{u}) = \mathcal{L}(\mathbf{u})^{N+1} + \left(\sum_{k=1}^{N} -z_{2k} + \sum_{j,k=1}^{N} m_{kj} Q(\psi_{2k}, \psi_{2j-1})\right)\mathcal{L}(\mathbf{u})^N + \cdots$$

so we compute a similar expansion on the left,

$$\mathcal{L}(\mathbf{v})D^N = \mathcal{L}(\mathbf{u})D^N + \begin{pmatrix} 0 & u - v \\ \bar{u} - \bar{v} & 0 \end{pmatrix} D^N.$$

We use the expression for D^N, commute and identify the coefficients of $\mathcal{L}(\mathbf{u})^N$. This yields

$$\Big[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \sum_{j,k=1}^{N} m_{kj} \begin{pmatrix} -\psi_{2k}^2 \psi_{2j-1}^1 & \psi_{2k}^1 \psi_{2j-1}^1 \\ -\psi_{2k}^2 \psi_{2j-1}^2 & \psi_{2k}^1 \psi_{2j-1}^2 \end{pmatrix} \Big] = \begin{pmatrix} 0 & u_1 - v_1 \\ u_2 - v_2 & 0 \end{pmatrix},$$

which leads to the desired formula (3.31).

(c) The image of a z-wave ψ for $\mathbf{u}$ is a z-wave $D^N\psi$ for $\mathbf{v}$ by iterated application of Theorem 3.7 (a). The formula (3.32) is a direct consequence of (3.30).

(d) Here we consider the eigenfunction formula (3.33). Let ϕ be a $\zeta_{2\ell}$ wave. Then ϕ is mapped to a $\zeta_{2\ell}$ wave. As in (3.19) we get

$$D^N \phi = W(\psi_{2\ell}, \phi) \prod_{k \neq \ell} (z_{2k} - z_{2\ell}) \sum_{j=1}^{N} m_{j\ell} \psi_{2j-1}.$$

Formula (3.33) follows since the Wronskian is constant, and we may swap the odd and even indices. □

3.4. **The intertwining operator for NLS.** Our main interest is in the NLS equation, where

$$u_1 = u, \quad u_2 = \bar{u}, \quad z_1 = \zeta, \quad z_2 = \bar{\zeta}, \quad \psi_1 = \phi, \quad \psi_2 = M\bar{\phi} = \begin{pmatrix} -\bar{\phi}^2 \\ \bar{\phi}^1 \end{pmatrix},$$

and ϕ is a ζ-wave. Then the Wronskian is always nonzero,

$$W(\psi_1, \psi_2) = \det \begin{pmatrix} \phi^1 & -\bar{\phi}^2 \\ \phi^2 & \bar{\phi}^1 \end{pmatrix} = |\phi|^2,$$

so all the formulas in the previous subsection apply on the full real line. The intertwining operator becomes

$$D\psi = D(u, \zeta, \phi)\psi = \Big(\mathcal{L}(u) - \bar{\zeta} - 2i \operatorname{Im} \zeta \frac{\phi\phi^*}{|\phi|^2} \Big) \psi. \tag{3.36}$$

and the Bäcklund transform becomes

$$v = B(u, \zeta, \phi) := u + 4 \operatorname{Im} \zeta \frac{\phi^1 \bar{\phi}^2}{|\phi|^2}. \tag{3.37}$$

The wave function ϕ and $M\bar{\phi}$ are in the null space of $D(u, \zeta, \psi)$, and

$$D(u, \bar{\zeta}, M\bar{\phi}) = D(u, \zeta, \psi), \qquad B(u, \zeta, \phi) = B(u, \bar{\zeta}, \bar{\phi}),$$

which can be written out as

$$D(B(u, \zeta, \phi), \bar{\zeta}, \frac{\phi}{|\phi|^2}) = \mathcal{L}(u) - \bar{\zeta} - 2i \operatorname{Im} \zeta \frac{M\phi(M\phi)^*}{|\phi|^2}. \tag{3.38}$$

The intertwing relation (3.22) becomes

$$\mathcal{L}(B(u, \zeta, \phi)) D(u, \zeta, \phi) = D(u, \zeta, \phi) \mathcal{L}(u) \tag{3.39}$$

and we obtain

$$\mathcal{L}(B(u, \zeta, \phi)) \frac{\phi}{|\phi|^2} = \bar{\zeta} \frac{\phi}{|\phi|^2}. \tag{3.40}$$

If

$$v = B(u, \zeta, \phi),$$

then we have

$$u = B(v, \bar{\zeta}, \frac{1}{|\phi|^2}\phi) = B(v, \zeta, \frac{1}{|\phi|^2}M\bar{\phi}). \tag{3.41}$$

We will use the intertwining operator D in two cases:

- when ϕ is a wave function which is unbounded at both ends.
- when ϕ is an eigenfunction.

The remaining case when $\phi = \psi_l$ is not an eigenfunction is also of interest, but not relevant here.

We begin our discussion with the first case. Let ψ be a wave function for u, at the spectral parameter ζ, and which is unbounded at $\pm\infty$. We normalize it so that

$$\lim_{x\to-\infty} e^{-i\zeta x}\psi^2(x) = e^{\kappa},$$

$$\lim_{x\to\infty} e^{i\zeta x}\psi^1(x) = e^{-\kappa}.$$

for a unique choice

$$\kappa \in \mathbb{C}\backslash(\pi i\mathbb{Z}).$$

We recall that κ is uniquely determined if z is an eigenvalue, but can be chosen arbitrarily otherwise.

We know that the Lax operator for $v = B(u, \zeta, \psi)$ has an eigenvalue at ζ with associated eigenfunction

$$\phi = \frac{1}{|\psi|^2}\begin{pmatrix} -\overline{\psi^2} \\ \overline{\psi^1} \end{pmatrix}.$$

Then a brief calculation shows that

$$\phi = -e^{-\kappa}\psi_l = e^{\kappa}\psi_r \tag{3.42}$$

where ϕ_l and ϕ_r are the left resp. right Jost function for v.

Remark 3.10. *This property is what allows us to identify our use of κ as a notation for a scattering parameter, in the first section, to the current use of κ as a parameter for the unbounded eigenfunctions.*

It will often be convenient to use the alternative notation

$$e^{\kappa} = e^{i(\beta_0+\beta_1\zeta)},$$

with $\beta_0, \beta_1 \in \mathbb{R}$.

Lemma 3.11. *Let $\psi_l(u)$, $\psi_l(v)$, $\psi_r(u)$ and $\psi_r(v)$ are left resp. right Jost functions for u resp. v to the spectral parameter z and let ϕ be an unbounded ζ wave. Then*

$$D(u,\zeta,\phi)\psi_l(u) = (z-\bar{\zeta})\psi_l(v), \qquad D(u,\zeta,\phi)\psi_r(u) = (z-\bar{\zeta})\psi_r(v), \tag{3.43}$$

and, if the Jost functions at ζ are unbounded (resp. ζ is not a pole for T, equivalently ζ is not an eigenvalue), and

$$\phi = T(\zeta)^{-1}\Big(e^{-i(\beta_0+\beta_1\zeta)}\psi_l(\zeta) + e^{i(\beta_0+\beta_1\zeta)}\psi_r(\zeta)\Big),$$

$$\frac{1}{|\phi|^2}\begin{pmatrix} -\overline{\phi^1} \\ \overline{\phi^2} \end{pmatrix} = -e^{-i(\beta_0+\beta_1\zeta)}\psi_l(\zeta, v) = e^{i(\beta_0+\beta_1\zeta)}\psi_r(\zeta, v). \tag{3.44}$$

Moreover

$$D(B(u,\zeta,\phi),\bar{\zeta},\frac{\phi}{|\phi|^2})D(u,\zeta,\phi) = (\mathcal{L}(u)-\zeta)(\mathcal{L}(u)-\bar{\zeta}). \tag{3.45}$$

Proof. Since the operator $D(u,\zeta,\psi)$ becomes $\mathcal{L}(u)-\bar\zeta$ at infinity, we have

$$\lim_{x\to\infty} e^{izx}(D(u,\zeta,\phi)\psi_l(u))^1 = (z-\bar\zeta)T(z,u)^{-1},$$

which, together with the same calculation for the right Jost functions, implies (3.43). The second formula is an immediate consequence. For the last formula we use (3.38). □

If ζ is an eigenvalue, then ϕ_l and ϕ_r coincide up to a constant as above, see (3.12). Hence to characterize the normalized eigenfunction ϕ we can use the properties in the above Lemma for ψ_l for $x<x_0$, and for ψ_r for $x>x_0$. We combine Lemma 3.1 with the previous constructions.

Proposition 3.12. *(1) If ϕ is unbounded as $x\to\pm\infty$ then $D(u,\zeta,\phi): H^{s+1}\to H^s$ is injective and has closed range of codimension 2, with orthogonal complement spanned by $|\phi|^{-2}M_0\phi$ and $|\phi|^{-2}MM_0\bar\phi$ (which are the ζ, respectively $\bar\zeta$ eigenfunctions of $L(v)^*$). Further, $|\phi|^{-2}M\bar\phi$ is a ζ eigenfunction for $\mathcal{L}(v)$, and $|\phi|^{-2}\phi$ is a $\bar\zeta$ eigenfunction for $\mathcal{L}(v)$, and*

$$(z-\zeta)T(u,z) = (z-\bar\zeta)T(v,z).$$

(2) If ϕ is an eigenfunction then $D(u,\zeta,\phi): H^{s+1}\to H^s$ is surjective, with null space spanned by ϕ and $M\bar\phi$. Moreover,

$$(z-\bar\zeta)T(u,z) = (z-\zeta)T(v,z).$$

(3) If ζ is in the resolvent set of $\mathcal{L}(u)$ and ϕ is a ζ wave function for $\mathcal{L}(u)$ as in (3.7) then the maps

$$\zeta\times(x_0,\theta)\times u\to D(u,\zeta,\phi)\in L(H^{s+1},H^s)$$

and

$$H^s\ni u\to B(u,\zeta,\phi)-u\in H^{s+1}$$

are analytic and separately holomorphic as functions of ζ, $\bar\zeta$, u and $\bar u$, as discussed in the beginning of this section. They and their derivatives are uniformly bounded on the set

$$\{\delta\langle\operatorname{Re}\zeta\rangle < \operatorname{Im}\zeta\}\times\mathbb{R}\times(\mathbb{R}/\pi\mathbb{Z})\times(H^s\cap\{u:\|u\|_{l^2DU^2}<\delta/C\}),$$

for some $C>0$.

(4) If ζ is an simple eigenvalue of $\mathcal{L}(v)$ with eigenfunction ϕ then the maps

$$v\to\zeta(\mathcal{L}(v)),$$
$$v\to D(v,\zeta,\phi)\in L(H^{s+1},H^s),$$

and

$$H^s\ni v\to B(v,\zeta,\phi)-v\in H^{s+1}$$

are analytic and holomorphic as functions of ζ, $\bar\zeta$, u and $\bar u$.

Proof. To prove the claims we suppose that $f\in H^s$ and we study solutions to

$$D(u,\zeta,\phi)\psi = f. \tag{3.46}$$

Let ϕ be an unbounded wave function. Then

$$\lim_{x\to-\infty}\frac{\phi\bar\phi}{|\phi|^2} = \begin{pmatrix}0&0\\0&1\end{pmatrix},$$

and the equation at $-\infty$ becomes

$$\begin{pmatrix}\partial+iz & 0\\ 0 & -\partial+i\bar z\end{pmatrix}\psi = f.$$

Hence we obtain the unique solution (if it exists) by integration from $-\infty$. Similarly

$$\lim_{x\to+\infty}\frac{\phi\bar\phi}{|\phi|^2} = \begin{pmatrix}1&0\\0&0\end{pmatrix}$$

and we find the solution (if it exists) solving from ∞. Both solutions have to coincide at $x = 0$, which shows that we can solve (3.46) on a set of f of codimension 2.

By the previous lemma we know that $|\phi|^{-2}M\bar\phi$ is a ζ eigenfunction for $\mathcal{L}(v)$. Then by symmetries, $|\phi|^{-2}\phi$ is a $\bar\zeta$ eigenfunction for $L(v)$, and $|\phi|^{-2}MM_0\bar\phi$, respectively $|\phi|^{-2}M_0\phi$ are eigenfunctions for $\mathcal{L}(v)^*$ associated to the eigenvalues z, respectively $\bar z$.

To identify the co-kernel we compute the adjoint

$$D^* = \mathcal{L}^*(u) - \zeta + 2i \operatorname{Im} z \frac{\phi\bar\phi^T}{|\phi|^2} = \mathcal{L}^*(v) - \zeta + 2i \operatorname{Im} z \frac{M_0\phi(M_0\phi)^*}{|\phi|^2}.$$

Inserting the two eigenfunctions above for $\mathcal{L}(v)^*$ in this formula yields the desired basis for the kernel of D^*.

If ϕ is an eigenfunction then

$$\lim_{x\to-\infty} \frac{\phi\bar\phi}{|\phi|^2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\lim_{x\to\infty} \frac{\phi\bar\phi}{|\phi|^2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and every solution to the initial value problem (3.46) with prescribed initial data $\psi(0) = \psi_0$ is in H^{s+1}. We obtain the null space by choosing $f = 0$. It is an easy verification that ϕ and $M\phi$ span the kernel of $D(u, \zeta, \phi)$. □

Of course we can relax the connection and use $\mathbf{u}$ and two nonreal numbers z_1 and z_2. The crucial additional condition is that the Wronskian

$$\psi_1^1\psi_2^2 - \psi_1^2\psi_2^1$$

does not vanish. This is certainly true if

$$\|u_j\|_{l^2(DU^2)} + \|u_2\|_{l^2(DU^2)} < \delta, \quad |z_1 - \bar z_2| < \delta \text{ for some positive number } \delta.$$

3.5. **The Bäcklund transform associated to holomorphic families of wave functions.** Here we introduce a key generalization of the previous discussion of the Bäcklund transform, which will be critical later in the context of iterated Bäcklund transforms. Precisely, starting with an initial state u and some z with $\operatorname{Im} z > 0$, instead of single wave functions we consider a holomorphic family $\psi = \psi(x, \zeta)$ of ζ-wave functions, for ζ near z, or more generally for ζ in an open subset of the half-plane and x in an interval.

In the NLS case $u_2 = \bar u_1$ we consider holomorphic families of unbounded wave functions on $\mathbb{R}$. Away from the spectrum of $\mathcal{L}(u)$, the next lemma identifies such unbounded wave functions with a holomorphic function α which relate it to the left, respectively the right Jost functions:

Lemma 3.13. *Let $u \in H^s$, $s > -\frac{1}{2}$ and U an open set whose closure is compact in the upper half plane, without eigenvalues for $\mathcal{L}(u)$. Let α be a holomorphic function on U. Then there exists a unique holomorphic family of unbounded wave functions $\psi(u, z)$ with*

$$\lim_{x\to-\infty} e^{-izx}\psi^2(x, z) = e^{i\alpha(z)},$$

$$\lim_{x\to\infty} e^{izx}\psi^1(x, z) = e^{-i\alpha(z)}.$$

The same is true in the case $\mathbf{u} = (u_1, u_2)$ without assuming $u_2 = \bar u_1$.

Proof. We define

$$\psi(x, z) = T(z)(e^{-i\alpha(z)}\psi_l(x, z) + e^{i\alpha(z)}\psi_r(x, z)).$$

Uniqueness is easy to see. □

By contrast, at eigenvalues of $\mathcal{L}(u)$, holomorphic families of unbounded wave functions $\psi(u,z)$ the Taylor expansions are uniquely determined up to an order given by the multiplicity.

Lemma 3.14. *Let ζ be a zero of T^{-1} of multiplicity N. Then there exist α_j, $0 \le j < N$ so that for every holomorphic family of unbounded wave functions near $z = \zeta$, and α defined as above, we must have*

$$\alpha_0 - \alpha(\zeta) \in 2\pi\mathbb{Z},$$
$$\alpha_j = \alpha^{(j)}(\zeta), \qquad 1 \le j \le N-1,$$

for $1 \le j < N$. Conversely, given α_j there exists a unique polynomial α of degree at most $N-1$ satisfying the above conditions, along with an associated family of holomorphic unbounded wave functions.

Proof. Let $\psi_0 = \psi_l(\zeta)$ be the ζ eigenfunction for the eigenvalue ζ of multiplicity N. We choose $\psi(0)$ linearly independent of $\psi_0(0)$. There exists a unique solution to

$$\psi_x = \begin{pmatrix} -i\zeta & u \\ -\bar{u} & i\zeta \end{pmatrix} \psi$$

with these initial data. The Wronskian satisfies $W(\psi_0(0), \psi(0)) \neq 0$ and it is constant. Since ψ_0 decays exponentially as $x \to \pm\infty$, it follows that ψ is unbounded as $x \to \pm\infty$. We obtain a holomorphic family of unbounded wave functions near $z = \zeta$ by solving

$$\psi_x = \begin{pmatrix} -iz & u \\ -\bar{u} & iz \end{pmatrix} \psi$$

with the same initial data for $\psi(0,z)$. After multiplication by a holomorphic function we may assume that

$$\lim_{x\to\infty} e^{izx}\psi^1(x,z) =: e^{-i\alpha(z)}$$

and

$$\lim_{x\to-\infty} e^{-izx}\psi^2(x,z) = e^{i\alpha(z)}$$

for some holomorphic function α (we chose this normalization instead of the initial condition).

Let $\tilde{\psi}$ be another holomorphic family of wave functions near $z = \zeta$. Then we can represent it as

$$\tilde{\psi}(0,z) = \lambda(z)\psi(0,z) + \mu(z)\psi_l(0,z)$$

Both sides are solutions and hence this relation holds for all x. In particular, if $\tilde{\psi}(.,\zeta)$ is unbounded then $\lambda(\zeta) \neq 0$. Choosing a smaller neighborhood if necessary we divide by $\lambda(z)$ and, by an abuse of notation we obtain $\lambda(z) = 1$ and

$$\tilde{\psi}(x,z) = \psi(x,z) + \mu(z)\psi_l(x,z).$$

Clearly each choice of μ gives a holomorphic family of unbounded wave functions near ζ. Now

$$\lim_{x\to\infty} e^{izx}(\psi^1(x,z) + \mu(z)\psi_l^1(x,z)) = e^{-i\alpha(z)} + T^{-1}(z)\mu(z)$$

and

$$\lim_{x\to-\infty} e^{-izx}(\psi^2(x,z) + \mu(z)\psi_l^2(x,z)) = e^{i\alpha(z)}.$$

Hence the defining function $\tilde{\alpha}$ for $\tilde{\psi}$ is given by

$$\tilde{\alpha}(z) = \alpha(z) + \frac{1}{2}\ln(1 + e^{-\alpha(z)}T^{-1}(z)\mu(z)). \tag{3.47}$$

where the logarithm exists in a neighborhood of ζ since $T^{-1}(\zeta) = 0$ by an abuse of notation. Here T^{-1} vanishes of order N at ζ. Then $\alpha(\zeta) - \tilde{\alpha}(\zeta) \in 2\pi\mathbb{Z}$, and for $0 < j < N$ we must have

$$\tilde{\alpha}^{(j)}(\zeta) = \alpha^{(j)}(\zeta).$$

Conversely, let $\hat{\alpha}$ be a holomorphic function with $\hat{\alpha}^{(j)}(\zeta) = \alpha^{(j)}(\zeta)$ for $0 \le j < N$. Then

$$\hat{\mu}(z) = \exp(\hat{\alpha}(z) - \alpha(z))(1 + e^{-\alpha(z)}T^{-1}(z))^{-2}$$

yields $\hat{\alpha}$. □

We further remark that the above class of functions α express the order N matching between ψ_l and ψ_r at the pole. The Wronskian relation

$$W(\psi_l, \psi_r) = T^{-1}$$

shows that at ζ as in the lemma the vectors ψ_l and ψ_r agree exactly to order N up to a multiplicative factor. Away from the pole we must have

$$\psi(z) = T^{-1}(z)(e^{-i\alpha(z)}\psi_l(z) + e^{i\alpha(z)}\psi_r(z)).$$

So this multiplicative factor is exactly determined by α,

$$e^{-i\alpha(z)}\psi_l(z) = -e^{i\alpha(z)}\psi_r(z) + O((z-\zeta)^N).$$

The interesting feature of working with a holomorphic family of unbounded wave functions $\psi(u, z)$ is that we can propagate it across any associated Bäcklund transform in a way that carries full information. Again we consider the general case, but we also specialize to the NLS case.

Lemma 3.15. *Let I be an interval, $\zeta_1, \zeta_2 \in \mathbb{C}\backslash\mathbb{R}$ with $\zeta_1 \neq \zeta_2$, ψ_2 a ζ_2 wave and $\psi(x, z)$ a holomorphic family of wave functions for $x \in I$ and z in a neighborhood of ζ_1, $\psi_1 = \psi(., \zeta_1)$. Then*

$$\psi_v(x, z) = D(u, \zeta_1, \zeta_2, \psi_1, \psi_2) \begin{cases} \dfrac{\psi(., z) - \psi(., \zeta_1)}{z - \zeta_1} & z \neq \zeta_1 \\ \partial_z \psi(., z) & z = \zeta_1 \end{cases} \tag{3.48}$$

is a holomorphic family of wave functions for

$$v = B(u, \zeta_1, \zeta_2, \psi_1, \psi_2). \tag{3.49}$$

Let

$$\phi_1(x) = \frac{1}{W(\psi_1, \psi_2)}\psi_2 \qquad \phi_2(x) = \frac{1}{W(\psi_1, \psi_2)}\psi_1.$$

Then

$$\psi(., z) = \frac{1}{z - \zeta_2} D(v, \zeta_1, \zeta_2, \phi_1, \phi_2)\psi_v(., z), \qquad u = B(v, \zeta_1, \zeta_2, \phi_1, \phi_2). \tag{3.50}$$

Proof. It is obvious that $\psi_v(x, z)$ is holomorphic in z. By Theorem 3.7 $\psi_v(x, z)$ is an unbounded wave function for v if $z \neq \zeta_1$. By continuity the same is true for $z = \zeta_1$. A direct calculation shows that α does not change. The final assertion about inversion follows by (3.27). □

If $I = \mathbb{R}$, $u_2 = \bar{u}_1$, $\zeta_1 = \bar{\zeta}_1$, $\psi_2 = \begin{pmatrix} \bar{\psi}^2 \\ -\bar{\psi}^1 \end{pmatrix}$ and if ψ is an unbounded family parametrized by the same holomorphic function $\alpha(z)$ then ψ_v is also parametrized by $\alpha(z)$.

Conversely, we can start from v and ψ_v and recover u and ψ:

Lemma 3.16. *Let ψ_v be a holomorphic unbounded family of wave functions for v, and ζ an eigenvalue for $\mathcal{L}(v)$. Then with u and ψ defined by* (3.50)*, the relations* (3.48)*,* (3.49) *hold.*

4. The soliton addition and removal maps

In the previous section we have shown how to add one soliton to an existing state by applying a Bäcklund transform with respect to an unbounded wave function, and, in reverse, how to remove a soliton by applying a Bäcklund transform with respect to an eigenfunction.

The Bäcklund transforms can be iterated to add multiple solitons. Theorem 3.9 provides compact formulas provided the matrix M_{jk} is invertible. But this is always true for the focusing case, see Lemma 4.1 below.

Our aim in this section is to spell out the results of the last section for adding and subtracting mutiple solitons, and to provide algebraic proofs for the properties of the matrices m and M. Finally Lemma 4.4 will provide a sharp estimate of the uniform norm of multiple pure solitons.

To keep the analysis simple, for the computations in this section we only consider the case of distinct eigenvalues (spectral parameters). The soliton addition map $\mathbf{B}_+^N$ will add N prescribed solitons to a given state, and the soliton removal map $\mathbf{B}_-^N$ will remove n existing solitons from a given state.

Given an open subset U with compact closure of the complex upper half-space we define the nondegenerate phase space for N-solitons as

$$\mathbf{S}_U^{N,0} = \{\mathbf{s} = (\mathbf{z}, \boldsymbol{\kappa}) \in U^N \times (\mathbb{C}/i\pi\mathbb{Z})^N;\ z_i \neq z_j\},$$

where

$$\mathbf{z} = (z_1, \cdots, z_N), \qquad \boldsymbol{\kappa} = (\kappa_1, \cdots, \kappa_N).$$

4.1. The soliton addition map.

We will view the soliton addition map $\mathbf{B}_+^N$ as a map

$$\mathbf{B}_+^N : H^s \times \mathbf{S}_U^{N,0} \to H^s.$$

We denote the output by

$$H^s \ni u \to v = \mathbf{B}_+^N(u, (\mathbf{z}, \boldsymbol{\kappa})) \in H^s.$$

To define it we impose some natural restrictions, namely that the z_k are not poles for T_u. These will be satisfied for instance if the spectral parameters $\mathbf{z}$ are localized in a compact subset of the upper half-space and u is sufficiently small in H^s.

To describe it we start with the N distinct spectral data $\mathbf{z} = (z_1, \cdots, z_N)$ in U and corresponding scattering data $\boldsymbol{\kappa} = (\kappa_1, \cdots, \kappa_N)$. We denote $\mathbf{s} = (\mathbf{z}, \boldsymbol{\kappa})$ the corresponding element of $\mathbf{S}_U^{N,0}$. We consider the associated left and right z-waves $\psi_{k,l}$ and $\psi_{k,r}$, and use them to define the unbounded wave functions (z_j, ψ_j) for v which have spectral parameters κ_j (see (3.7) and (3.8)),

$$\psi_j = e^{-\kappa_j}\psi_{j,l} + e^{\kappa_j}\psi_{j,r}.$$

We inductively apply n Bäcklund transforms as follows,

$$\psi_j^{(k+1)} = D(u^{(k)}, \psi_k^{(k)}, z_k)\psi_j^{(k)}, \qquad u^{(k+1)} = B(u^{(k)}, \psi_k^{(k)}, z_k)$$

where we initialize

$$\psi_k^{(1)} = \psi_k, \qquad u^{(1)} = v$$

Then we define the soliton addition map as

$$\mathbf{B}_+^N(v, \mathbf{s}) := u^{(N+1)}$$

We will also denote the iterated Bäcklund transform as

$$B^N(u, \mathbf{z}, \boldsymbol{\kappa}) = \prod_{k=1}^{N} D(u^{(k)}, \psi_k^{(k)}, z_k). \tag{4.1}$$

We specialize the formulas of Theorem 3.9. We start with the symmetric matrix M with complex entries

$$M_{jk} = \frac{i\psi_k^* \psi_j}{\bar{z}_k - z_j}$$

Lemma 4.1. *Suppose that the imaginary part of z_j is positive, that the z_j are pairwise disjoint and that the $\psi_j \in \mathbb{C}^2$ are nonzero. Then M_{jk} is positive definite.*

Proof. We define the nonzero functions

$$\Psi_j(t) = e^{iz_j t}\psi_j \in L^2((0,\infty);\mathbb{C}^2).$$

which are linearly independent since z_j are distinct. Then M is their Gramian matrix. □

We denote by m the inverse matrix $m = M^{-1}$. Then the formulas in Theorem 3.9 take the following form:

(1) The iterated intertwining operator is given by

$$D^N(\mathbf{z}, \boldsymbol{\kappa}) = \left(I + \sum_{j,k=1}^{N} m_{kj}\psi_j\psi_k^*(\mathcal{L}_v - \bar{z}_k)^{-1} \right) \prod_{l=1}^{N}(\mathcal{L}_v - \bar{z}_l). \tag{4.2}$$

The image of a z-wave ψ for v is a z-wave $D^N\psi$ for u where

$$D^N = \prod_{l=1}^{N}(z - \bar{z}_l) \left(I + \sum_{j,k=1}^{N} \frac{1}{z - \bar{z}_k} m_{kj}\psi_j\psi_k^* \right). \tag{4.3}$$

(2) The output function $v = \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\kappa})$ is given by

$$v = u + 2m_{kj}\psi_j^1\bar{\psi}_k^2. \tag{4.4}$$

(3) The functions

$$\phi_j = m_{jk}\psi_k \tag{4.5}$$

are $\bar{z}_j$ eigenfunctions for $\mathcal{L}_v$ with scattering parameters κ_j. Moreover,

$$v = u + \sum_{j=1}^{N} \phi_j^1\bar{\psi}_j^2.$$

We can represent the iterated Bäcklund transform as follows:

$$\begin{aligned} D^N &= \sum_{j=1}^{N} \left(1 - \sum_{k=1}^{N} \phi_k\psi_k^*(\mathcal{L}_u - \bar{z}_k)^{-1}\right) \prod_{\ell=1}^{n}(\mathcal{L}_u - \bar{z}_\ell) \\ &= \prod_{l=1}^{N}(\mathcal{L}_v - \bar{z}_\ell)\Big(1 - \sum_{k=1}^{N}(\mathcal{L}_v - \bar{z}_k)^{-1}\phi_k\psi_k^*\Big) \\ &= \sum_{j=1}^{N} \left(1 - \sum_{k=1}^{N} M\bar{\phi}_k M\psi_k^t(\mathcal{L}_u - z_k)^{-1}\right) \prod_{\ell=1}^{n}(\mathcal{L}_u - z_\ell). \end{aligned} \tag{4.6}$$

4.2. **The soliton removal map.** We will view the soliton removal map $\mathbf{B}_-^N$ as a map

$$\mathbf{B}_-^N : \mathbf{V}_U^{N,0} \to H^s \times \mathbf{S}_U^{N,0}.$$

We denote the output by

$$H^s \ni v \to \mathbf{B}_-^N(v) = (u, \mathbf{z}, \boldsymbol{\kappa}) \in \mathbf{S}_U^{N,0} \times H^s.$$

To define it we again impose some natural restrictions on $v \in H^s$, namely we select an open subset U inside the upper half-space with compact closure, and assume that the transmission coefficient $T_v(z)$ has exactly N simple poles z_j in U.

Now the spectral parameters $\mathbf{z}$ are defined as the poles of T_v within K. These will be simple eigenvalues of $\mathcal{L}_v$; then their conjugates $\bar{\mathbf{z}}$ will also be simple eigenvalues of $\mathcal{L}_v$.

We denote by $(\phi_1, \cdots, \phi_N)$ a corresponding set of eigenfunctions for $\mathbf{z}$. The scattering parameters $\boldsymbol{\kappa}$ will be determined by the relations

$$\phi_j = -e^{-\kappa_j} \psi_{j,l} = e^{\kappa_j} \psi_{j,r},$$

comparing the left and right wave functions to the eigenfunction.

Then we define the rest of the soliton removal map $\mathbf{B}_-^N$ exactly as we have previously defined the soliton addition map $\mathbf{B}_+^N$, but starting from v and the $\mathbf{z}$ eigenfunctions $(\phi_1, \cdots, \phi_N)$. To describe the soliton removal map we start with the symmetric matrix m with entries

$$m_{jk} = \frac{i\phi_k^* \phi_j}{\bar{z}_j - z_k}.$$

We denote $M = m^{-1}$. Then we have the following formulas for the removal map

(1) The operator D^N is given by

$$D^N = \left(I + \sum_{j,k=1}^N M_{kj} \phi_j \phi_k^* (\mathcal{L}_v - \bar{z}_k)^{-1} \right) \prod_{\ell=1}^N (\mathcal{L}_v - \bar{z}_\ell). \tag{4.7}$$

In particular the image of a z-wave ψ for v is a z-wave $D^n \psi$ for u where

$$D^n = \prod_{\ell=1}^n (z - z_\ell) \Big(I - \sum_{j,k=1}^n \frac{1}{z - \bar{z}_k} M_{kj} \phi_j \phi_k^* \Big). \tag{4.8}$$

(2) The output function $u = \mathbf{B}_-^N(v, \mathbf{s}, \boldsymbol{\kappa})$ is given by

$$u = v - M_{jk} \phi_k^1 \bar{\phi}_j^2. \tag{4.9}$$

(3) The functions

$$\psi_j = M_{jk} \phi_k \tag{4.10}$$

are z_j wave functions for $\mathcal{L}_u$ with scattering parameters κ_j.

4.3. **Connecting the two maps.** Here we briefly discuss the relation between the soliton addition and removal maps in the context of isolated eigenvalues. It follows from the corresponding result for single simple eigenvalues that the maps are inverses.

Theorem 4.2. *We have*

$$\mathbf{B}_-^N \circ \mathbf{B}_+^N = Id \tag{4.11}$$

for a finite number N of simple eigenvalues.

The soliton addition and removal operations are symmetric with the roles of ϕ and ψ essentially reversed. This is a consequence of the construction by iterative Bäcklund transforms, but it is also a consequence of a purely algebraic relation.

Lemma 4.3. *For nonzero $\psi_j \in \mathbb{C}^2$ nonzero define*

$$M_{jk} = \frac{i\psi_k^* \psi_j}{\bar{z}_k - z_j}, \qquad m = M^{-1}, \qquad \phi_j = m_{jk}\psi_k.$$

Then we have the converse relation

$$m_{jk} = \frac{i\phi_k^* \phi_j}{z_k - \bar{z}_j}.$$

Proof. Let z be a complex number, different from the z_k and $\bar{z}_k$ and considering the 2×2 matrices

$$\chi = 1 + \sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - \bar{z}_k},$$

$$\chi^+ = 1 - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k},$$

we compute

$$\begin{aligned}
\chi\chi^+ &= I + \sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - \bar{z}_k} - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k} + \sum_{k,j=1}^{n} \frac{\phi_k \psi_k^*}{z - \bar{z}_k} \frac{\psi_j \phi_j^*}{z - z_j} \\
&= I + \sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - z_k} - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k} + \sum_{j,k=1}^{n} \phi_k \psi_k^* \psi_j \phi_j^* \frac{1}{\bar{z}_k - z_j} \left(\frac{1}{z - \bar{z}_k} - \frac{1}{z - z_j} \right) \\
&= I + \sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - z_k} - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k} + \sum_{j,k=1}^{n} M_{jk} \phi_k \phi_j^* \left(\frac{1}{z - z_j} - \frac{1}{z - \bar{z}_k} \right) \\
&= I + \sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - z_k} - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k} + \sum_{j=1}^{n} \frac{1}{z - z_j} \psi_j \phi_j^* - \sum_{k=1}^{n} \frac{1}{z - z_k} \phi_k \psi_k^* \\
&= I.
\end{aligned}$$

This implies that $\chi^+\chi = 1$ which yields

$$\sum_{k=1}^{n} \frac{i\phi_k \psi_k^*}{z - \bar{z}_k} - \sum_{k=1}^{n} \frac{i\psi_k \phi_k^*}{z - z_k} = - \sum_{j,k=1}^{n} \frac{\psi_j \phi_j^*}{z - z_j} \frac{\phi_k \psi_k^*}{z - \bar{z}_k} = - \sum_{j,k=1}^{n} \psi_j \phi_j^* \phi_k \psi_k^* \frac{1}{\bar{z}_k - z_j} \left(\frac{1}{z - \bar{z}_k} - \frac{1}{z - z_j} \right).$$

Identifying the residues we obtain the relations

$$\phi_k \psi_k^* = i \sum_{j=1}^{n} \frac{\phi_j^* \phi_k}{\bar{z}_k - z_j} \psi_j \psi_k^*,$$

or equivalently

$$\phi_k = i \sum_{j=1}^{n} \frac{\phi_j^* \phi_k}{\bar{z}_k - z_j} \psi_j,$$

which leads to the desired conclusion. ∎

For later use we include here another algebraic relation related to the soliton addition/removal transforms. Precisely, consider the Hermitian 2×2 positive definite matrix

$$A = \sum_{j,k=1}^{n} m_{kj} \psi_j \psi_k^*$$

where we note that the difference between u and v is one of the off-diagonal entries of this matrix. Rather than trying to bound that particular entry, we produce a bound for the entire matrix, via its trace.

Lemma 4.4. *We have*

$$TrA = 2\sum_{j=1}^{n} \operatorname{Im} z_j. \tag{4.12}$$

As a consequence, we obtain a uniform bound for $|u - v|$.

Proof. This lemma seems to have little to do with the context of our problem. Writing the trace of A in the form

$$\operatorname{Tr} A = \sum_{j,k=1}^{n} m_{kj}\psi_k^*\psi_j,$$

this becomes a statement which only involves the (complex) dot products of ψ_j and ψ_k. We consider first the case when ψ_j take values in $\mathbb{C}^n$ assuming that they are linearly independent. The statement of the lemma follows then by continuity.

To prove the above trace property we represent the Gram matrix as

$$(\psi_k^*\psi_j) = \operatorname{diag}(z_j)M - M\operatorname{diag}(\bar{z}_k).$$

Thus our trace becomes

$$\begin{aligned}\operatorname{Tr} A =& \sum_{j,k=1}^{n} m_{kj}\psi_k^*\psi_j \\ =& \operatorname{Tr}_{\mathbb{C}^n}((\operatorname{diag}(z_j)M - M\operatorname{diag}(\bar{z}_k))M^{-1}) = \operatorname{Tr}_{\mathbb{C}^n}(\operatorname{diag}(z_j) - \operatorname{diag}(\bar{z}_k)) \\ =& 2\sum_{j=1}^{n} \operatorname{Im} z_j.\end{aligned}$$

We remark that here the ψ's can be in an arbitrary Hilbert space, therefore the Gram matrix $\psi_j\bar{\psi}_k^t$ can be any arbitrary symmetric non-negative matrix. □

5. The extended soliton addition and removal maps

So far, we have only considered the iterated Bäcklund transform corresponding to isolated eigenvalues, which can be viewed as a smooth map

$$\mathbf{B}_+^N : H^s \times \mathbf{S}_U^{N,0} \to H^s, \qquad u \times \mathbf{s} \to v.$$

restricted to states u with no eigenvalues at $\mathbf{z}$. The problem with this setting is that when we endow $\mathbf{S}_U^{N,0}$ with the obvious smooth structure derived from $(\mathbf{S}_1)^N$, the soliton addition map does not admit a smooth extension to the diagonal with multiple eigenvalues.

Our contention here is that this does not reflect an inherent lack of smoothness for the soliton addition map at multiple eigenvalues, but rather the fact that we are using the wrong smooth structure on $\mathbf{S}_U^{N,0}$. To rectify that, our first step is to consider the iterated Bäcklund transform associated to holomorphic families of unbounded wave functions. Precisely, we start with

- A compact set U in the upper half-space,
- A state $u \in H^s$ with no eigenvalues in U
- A holomorphic family of unbounded wave functions $\psi(z)$ in a neighbourhood of U, associated to u, also with associated α as in Lemma 3.13.

Let $z_j \in U$ be pairwise disjoint for $1 \leq j \leq J$. Let $n_j \geq 1$ for $1 \leq j \leq J$ and $N = \sum n_j$. By an iterated Bäcklund transform corresponding to the holomorphic unbounded wave function ψ, we add N solitons at z_j with corresponding multiplicities n_j. By an iterated application of Lemma 3.15, the result only depends on the z_j and

$$(\partial^k \alpha(z_j))_{1\leq j\leq J,\ 0\leq k<n_j}. \tag{5.1}$$

We obtain an associated soliton addition map

$$\mathbf{B}^N_{\mathbf{z},\psi} : H^s \times U^N \to H^s.$$

A-priori this map, as a function of spectral parameters $\mathbf{z}$, is smooth, indeed real analytic, as well as symmetric. However the eigenvalues $\mathbf{z}$ of the Lax operator are only determined up to permutations, and they do not depend smoothly on the Lax operator in the case of multiplicities. Instead the elementary symmetric polynomials

$$s_j = \sum_{k=1}^{N} z_k^j, \qquad 1 \leq j \leq N$$

depend smoothly on the potential, and we will see that the soliton addition map is invariant under such a permutation and depends smoothly on the elementary symmetric polynomials. Secondly, we will parametrize the holomorphic wave function α recorded at $\mathbf{z}$ via $2N$ real variables $\beta_0, \cdots, \beta_{2N-1}$, which turn out to be naturally associated to the first $2N$ commuting flows. These enhancements are the topic of this section.

We seek to study further its regularity properties as well as its parametrization. For this we consider separately the spectral parameters $\mathbf{z}$ and the unbounded wave functions.

5.1. **The spectral data and the characteristic polynomial.** The spectral data $\mathbf{z}$ for the soliton addition map can be encoded the characteristic polynomial

$$P_{\mathbf{z}}(z) = \prod_{j=1}^{n}(z - z_j) = z^N + \sum_{k=1}^{N}(-1)^k s_k z^{N-k}$$

where $\mathbf{s} = \{s_k : 0 \leq k \leq N\}$ are the elementary symmetric polynomials in $\mathbf{z}$ with $s_0 = 1$.

Since the soliton addition map is symmetric as a function of $\mathbf{z}$, it is natural to seek to view it as a smooth function of s_k's, rather than separately in each individual eigenvalues. Because of this, on the space of spectral data $\mathbf{z}$ we will not use the product topology.

Instead, we will denote the space of spectral data by $\mathbb{C}^N_{sym}$, and interpret it as the space of unordered N-uples of complex numbers with the smooth topology defined by the elementary symmetric polynomials s_k.

The correspondence between the two topologies is continuous but not smooth.

Lemma 5.1. *A) Let $\mathbf{z} \in \mathbb{C}^N$. Then*

$$|\mathbf{s}(\mathbf{z})| \leq (1 + |\mathbf{z}|)^N$$

and

$$|\mathbf{z}| \leq \sqrt{2}|\mathbf{s}(\mathbf{z})|.$$

For all $\mathbf{s}$ there exists $\mathbf{z}$ with $\mathbf{s} = \mathbf{s}(\mathbf{z})$.

B) Let $\mathbf{z}, \mathbf{w} \in \mathbb{C}^N$. Then

$$|\mathbf{s}(\mathbf{z}) - \mathbf{s}(\mathbf{w})| \leq c_N(1 + |\mathbf{z}| + \mathbf{w}|)^{N-1}|\mathbf{z} - \mathbf{w}|.$$

C) Let $\mathbf{s}, \sigma \in \mathbb{C}^N$. Then there exist $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^N$ with $\mathbf{s} = \mathbf{s}(\mathbf{z})$, $\sigma = \mathbf{s}(\mathbf{w})$ and

$$|\mathbf{z} - \mathbf{w}| \leq C(|\mathbf{s}|, |\sigma|)|\mathbf{s} - \sigma|^{\frac{1}{N}}.$$

Proof. The first inequality in A) is immediate with the l^1 norm instead of the L^2 norm, which implies the bound in the l^2 norm. The components of $\mathbf{z}$ are the roots of

$$\sum_{n=0}^{N} s_{N-n} z^n = 0.$$

which are contained in the open disc with the given radius by the theorem of Gerschgorin. Part B is an immediate calculation. For Part C we study the dependence of roots on the coefficients of a polynomial. ☐

5.2. **The scattering data and holomorphic families of unbounded wave functions.** We have seen that a holomorphic family of unbounded wave functions ψ can be uniquely described (up to a multiplicative constant) via a holomorphic function $\alpha(z)$; because of this, we will identify the notations $\mathbf{B}^N_{\mathbf{z},\psi}$ and $\mathbf{B}^N_{\mathbf{z},\alpha}$. In the case of distinct eigenvalues $\mathbf{z}$, the associated soliton addition map depends only of $\kappa_j = i\alpha(z_j)$. Suppose now that we have multiple eigenvalues z_k with multiplicity n_k. In view of Lemma 3.15, the associated soliton addition map may depend on

$$\partial^j \alpha(z_j), \qquad j = 0, n_k - 1.$$

Thus we can use the equivalence relation

Definition 5.2. *For two holomorphic functions α and $\tilde{\alpha}$ we say that*

$$\alpha = \tilde{\alpha} \pmod{P_{\mathbf{z}}}$$

if there exists a holomorphic function q so that

$$\alpha(z) - \tilde{\alpha}(z) = q(z) P_{\mathbf{z}}(z).$$

Then we can rephrase the above discussion as

Lemma 5.3. *Assume that $\alpha = \tilde{\alpha} \pmod{P_{\mathbf{z}}}$. Then $\mathbf{B}^N_{\mathbf{z},\alpha} = \mathbf{B}^N_{\mathbf{z},\tilde{\alpha}}$.*

This equivalence relation will allow us to replace the holomorphic function α by an equivalent polynomial with degree at most $N - 1$.

Lemma 5.4. *Let $P_{\mathbf{z}}$ be the characteristic polynomial and α a holomorphic function in a neighbourhood of $\mathbf{z}$. Then there exists an unique polynomial (remainder)*

$$\tilde{\alpha}(z) = \sum_{k=0}^{N-1} \alpha_j z^j$$

so that

$$\tilde{\alpha} = \alpha \pmod{P_{\mathbf{z}}}.$$

Furthermore, $\tilde{\alpha}$ depends holomorphically on the symmetric polynomials $\mathbf{s}$.

Proof. We consider a contour γ around the zeroes $\mathbf{z}$ of $P_{\mathbf{z}}$. We must have

$$\int_\gamma z^j \frac{\alpha(z) - \tilde{\alpha}(z)}{P(z)}\, dz = 0, \qquad j \geq 0,$$

but only the first N such relations are independent. This yields

$$\int_\gamma z^j \frac{\alpha(z)}{P(z)}\, dz = \int_\gamma z^j \frac{\tilde{\alpha}(z)}{P(z)}\, dz, \qquad j = 0, N-1.$$

The left hand side is determined by α, and depends holomorphically on the symmetric polynomials in $\mathbf{z}$ (which are the coefficients of $P_{\mathbf{z}}$). This in turn uniquely determines the first N coefficients in the Taylor series for $\dfrac{\tilde{\alpha}(z)}{P(z)}$ at infinity, which in turn uniquely determines $\tilde{\alpha}(z)$. ☐

The $N-1$ degree polynomial $\tilde{\alpha}$ can be viewed as our generalized scattering parameter, and is identified via its (complex) coefficients $\boldsymbol{\alpha} = (\alpha_0, \cdots, \alpha^{N-1})$,

$$\alpha(z) = \sum_{j=0}^{N-1} \alpha_j z^j.$$

This is endowed with the smooth topology of $\mathbb{C}^N$.

However, there is also an equivalent alternative choice, which we will give preference to in this paper. Precisely, instead of working with complex polynomials of degree $N-1$, it is sometimes more convenient to work with real polynomials of degree $2N-1$.

If two real polynomials are equal modulo $P_{\mathbf{z}}$ then they are also equal modulo $P_{\bar{\mathbf{z}}}$ so they must[2] be equal modulo $P_{\mathbf{z}} P_{\bar{\mathbf{z}}}$. Thus the above lemma concerning α is replaced by

Lemma 5.5. *Let $P_{\mathbf{z}}$ be the characteristic polynomial and α a holomorphic function in a neighbourhood of $\mathbf{z}$. Then there exists an unique real polynomial*

$$\beta(z) = \sum_{k=0}^{2N-1} \beta_j z^j$$

so that

$$\beta = \alpha \pmod{P_{\mathbf{z}}}$$

Furthermore, $\boldsymbol{\beta} := (\beta_0, \cdots, \beta_{2N-1})$ depends analytically on the symmetric polynomials $\mathbf{s}$.

There is a (real) linear one-to one connection between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which is analytic in the symmetric polynomials $\mathbf{s}$. Thus the topologies determined by the $\boldsymbol{\alpha}$, respectively the $\boldsymbol{\beta}$ representations of the scattering parameters are equivalent.

5.3. **The smooth soliton parameters.** Based on the previous discussion, it is natural to define the phase space $\mathbf{S}_U^N$ associated to an open set U with compact closure in the open upper half-space as

$$\mathbf{S}_U^N = \{(\mathbf{z}, \boldsymbol{\beta}); \mathbf{z} \in \mathbb{C}^N_{sym}, \ \mathbf{z} \subset U, \quad \boldsymbol{\beta} \in \mathbb{R}^{2N}\}$$

with the smooth topology given by the symmetric polynomials $\mathbf{s}$ for $\mathbf{z}$ and the smooth topology in $\mathbb{R}^{2N}$ for $\boldsymbol{\beta}$.

It is easily seen that we have a smooth embedding of the N soliton set with pairwise different eigenvalues

$$\mathbf{S}_U^{N,0} \subset \mathbf{S}_U^N,$$

which is provided by the matching

$$\kappa_j = i\boldsymbol{\beta}(z_j).$$

Thus, one can view $\mathbf{S}_U^N$ as the completion of $\mathbf{S}_U^{N,0}$ with respect to the above topology. Our contention is that this is the correct smooth parametrization for extending the soliton addition and removal maps as smooth inverse maps to spectral parameters $\mathbf{z}$ with multiplicity.

Another symmetry which is readily seen at the level of κ_j is that κ_j are only uniquely determined modulo πi. At the level of $\boldsymbol{\beta}$, this yields the equivalence relation, denoted by A, defined

$$\boldsymbol{\beta}_1 \equiv \boldsymbol{\beta}_2 \quad \text{iff} \quad \boldsymbol{\beta}_1(z_j) = \boldsymbol{\beta}_2(z_j) \pmod{\pi i}.$$

This relations is now $\mathbf{z}$ dependent. There are two interesting observations to make:

- This is a discrete relation, uniformly in $\mathbf{z} \in U$. Thus the local smooth topologies on $\mathbf{S}_U^N$ and $\mathbf{S}_U^N/A$ coincide.

[2]Here we recall that all the z_j's in $\mathbf{z}$ are in the upper halfspace, so $P_{\mathbf{z}}$ and $P_{\bar{\mathbf{z}}}$ have no common roots.

- The dimension of the symmetry lattice depends on the multiplicities in $\mathbf{z}$. This corresponds to some periods approaching infinity as eigenvalues collapse.

In this section we carry out the first step of the analysis, and show that

Proposition 5.6. *The soliton addition map*

$$(u, (\mathbf{z}, \boldsymbol{\beta})) \to v = \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}) := \mathbf{B}_{\mathbf{z},\boldsymbol{\beta}}^N(u)$$

is one to one in a suitable setting, and commutes with every flow of the NLS hierarchy whenever this flow is a continuous extension of the flow on Schwartz functions. Moreover we have the energy relation (trace formula)

$$E_s(v) = E_s(u) + 2\sum_{k=1}^{N} m_k \Xi_s(2z_k) \tag{5.2}$$

and in particular

$$\|v\|_{L^2}^2 = \|u\|_{L^2}^2 + 2\sum_{k=1}^{N} \operatorname{Im} z_k. \tag{5.3}$$

Here for the flow of $\boldsymbol{\beta}$ we use the induced linear maps determined by the relations (1.19) where for κ we use the $\boldsymbol{\beta}$ representation $\kappa = i\sum \beta_n z^n$. This is also explicitely spelled out later in (6.2).

Proof. The parameters $\boldsymbol{\beta}$ yield a unique well-defined holomorphic family of unbounded wave functions $\psi_u = \psi_u(z, \boldsymbol{\beta})$ associated to u. Then by iteratively applying N Bäcklund transforms to the pair (u, ψ_u) corresponding to the eigenvalues $\mathbf{z}$, we obtain the pair (v, ψ_v) where ψ_v is another unbounded wave function with the same parameter $\boldsymbol{\beta}$. By Lemma 3.14, it follows that $\boldsymbol{\alpha}$ and thus $\boldsymbol{\beta}$ is uniquely determined by v modulo $\pi\mathbb{Z}$. This proves injectivity.

Conversely, let v be a potential with eigenvalues $\mathbf{z}$, possibly with multiplicities. Then by Lemma 3.14 applied to each eigenvalue there exists a unique polynomial α of degree at most $N-1$ which generates a family of holomorphic wave functions ψ_v associated to v. Following Lemma 3.16 we successively remove poles while propagating back the unbounded family of wave functions, until after N steps we obtain a pair (u, ψ_u) without any eigenvalues for $\mathcal{L}(u)$ in U. Then, by Lemmas 3.15, 3.16, (v, ψ_v) is the image of (u, ψ_u) through the iterated Bäcklund transformation. The assertion on the norms is an immediate consequence of the trace formula (1.23). ☐

6. The regularity of soliton addition and removal

With the setting of the previous section in place, we return to the question of the regularity of the soliton addition map with multiplicities. As a consequence we obtain a precise description of the pure soliton manifolds and the structure of the phase space and its dynamics.

6.1. The results.

Our goal here is to study the regularity properties of the soliton addition and the soliton removal maps. We begin by describing our setup, which requires the following elements:

- An open subset U of the upper half-space with compact closure in the open upper half plane.
- The set of N tuples $\mathbf{z}$ of complex numbers in U, up to permutations. We consider it as an analytic manifold with the analytic structure given by the elementary symmetric polynomials $\mathbf{s}$ in N variables, and use the notation $\mathbb{C}_{sym}^N$. Each $\mathbf{z}$ can also be identified with its characteristic polynomial

$$P_{\mathbf{z}}(z) = \prod_{k=1}^{N}(z - z_k).$$

and also, equivalently, with the real polynomial $P_{\mathbf{z}}P_{\bar{\mathbf{z}}}$.

- The associated soliton phase space $\mathbf{S}_U^N$ defined by

$$\mathbf{S}_U^N = \{(\mathbf{z}, \boldsymbol{\beta}) \subset \mathbb{C}^N_{sym} \times \mathbb{R}^{2N},\, \mathbf{z} \subset U\}/A.$$

We identify $\boldsymbol{\beta}$ with the polynomial

$$\boldsymbol{\beta}(z) = \sum_{j=0}^{2N-1} \beta_j z^j.$$

The set A is a discrete equivalence relation: We identify $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ if for $z_j \in \mathbf{z}$ we have

$$\boldsymbol{\beta}_1(z_j) - \boldsymbol{\beta}_2(z_j) \in \pi\mathbb{Z} \text{ and } \boldsymbol{\beta}_2^{(m)}(z_j) = \boldsymbol{\beta}_1^{(m)}(z_j) \text{ for } 1 \le m < m_j = \text{ the multiplicity of } z_j. \tag{6.1}$$

The set $\mathbf{S}_U^N$ carries the natural analytic structure defined by the analytic structure of the $\mathbf{z}$ and the Euclidean structure of the $\boldsymbol{\beta}$. The n-th flow acts on $\mathbf{S}_U^N$ by

$$\dot{\mathbf{z}} = 0, \quad \dot{\boldsymbol{\beta}}(z) = 2^{n-1} z^n \ (\mod P_{\mathbf{z}}P_{\bar{\mathbf{z}}}). \tag{6.2}$$

This can be viewed as Hamiltonian flows on the phase space endowed with the symplectic form

$$\omega = \sum_{k=0}^{2N-1} \beta_k \wedge d \sum_{j=1}^{N} \operatorname{Im} z_j^{k+1},$$

generated by the Hamiltonians

$$H_n(\mathbf{z}, \boldsymbol{\beta}) = \frac{2^n}{n+1} \operatorname{Im} \sum_{j=1}^{N} z_j^{n+1}.$$

- The N pure soliton set $\mathbf{M}_U^N$ of pure N solitons with spectral parameters in U. There is a natural map

$$\mathbf{M}_U^N \ni u \to \mathbf{z},$$

whose fibers are denoted by $M_{\mathbf{z}}$.

- The space $\mathbf{V}_U^0 \subset H^s$ of states with no spectral parameters in U,

$$\mathbf{V}_U^0 = \{u \in H^s : \sigma(\mathcal{L}(u)) \cap U = \emptyset\}.$$

- The space $\mathbf{V}_U^N \subset H^s$ of N soliton states with spectral parameters in U,

$$\mathbf{V}_U^N = \{u \in H^s : \#\sigma(\mathcal{L}(u)) \cap U) = N\}.$$

There is the obvious natural map

$$\mathbf{V}_N \ni u \to \mathbf{z} \in \mathbb{C}^N_{sym},$$

which is easily seen to be analytic.

- The m-th Hamiltonian of the hierarchy is a sum

$$H_m(u) = \sum_{j=1}^{m/2} H_{m,j}(u),$$

with

$$H_{2m,2}(u) = \int |u^{(m)}|^2 dx,$$

$$H_{2m+1,2}(u) = \frac{1}{i} \int u^{(m)} \partial_x \overline{u^{(m)}} dx.$$

The second index is half of the homogeneity in u. In $H_{m,j}$ there are $m+2-2j$ derivatives distributed over the $2j$ terms. The m-th Hamiltonian defines a flow on Schwartz space, and in particular on pure solitons $\mathbf{M}_U^N$.

H_2 is the Hamiltonian of the Schrödinger equation and H_3 is the Hamiltonian of mKdV.

In this context we consider the soliton addition map

$$\mathbf{B}_+^N : \mathbf{V}_U^0 \times \mathbf{S}_U^N \to \mathbf{V}_U^N,$$

and the soliton removal map

$$\mathbf{B}_-^N : V_U^N \to V_U^0 \times \mathbf{S}_U^N.$$

In the previous section we have seen that these maps are inverse maps. Here we study their regularity. We begin with the soliton addition map.

Theorem 6.1. *a) Let $s > -1/2$. The soliton addition map*

$$\mathbf{V}_U^0 \times \mathbf{S}_U^N \ni (u, \mathbf{z}, \boldsymbol{\beta}) \to v = \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}) \in \mathbf{V}_U^N$$

of adding N solitons is smooth, uniformly on compact sets in $\boldsymbol{\beta}$, for $\mathbf{z} \in K^N$ for a compact subset $K \subset U$.

b) The soliton addition map $\mathbf{B}_+^N$ is uniformly smooth globally in u and $\boldsymbol{\beta}$, for u restricted to a bounded set in H^s and $\mathbf{z}$ restricted as above.

This result provides the proper context to study the pure N-soliton set $\mathbf{M}_U^N$, which, by Proposition 5.6, can be described as

$$\mathbf{M}_U^N = \mathbf{B}_+^N(0, \mathbf{S}_U^N). \tag{6.3}$$

For this set we will prove Theorem 1.5, which we restate for convenience in the following:

Theorem 6.2. *Let $s > -\frac{1}{2}$, U and N as above. Then the pure N soliton set $\mathbf{M}_U^N$ is a uniformly smooth $4N$ dimensional Riemannian submanifold of H^s.*

The pure N solitons belong to all H^s spaces, and using a different s will yield an equivalent Riemannian structure in the metric sense. By construction there exists a smooth natural diffeomorphism $\mathbf{S}_U^N \to \mathbf{M}_U^N$, which commutes with the first $2N$ flows. This gives $\mathbf{S}_U^N$ a smooth Riemannian structure.

It is also natural to consider the foliation of $\mathbf{M}_U^N$ relative to the spectral parameter $\mathbf{z}$. It is not difficult to see that this is a smooth foliation, as we show later that $\mathbf{z}$ is a smooth nondegenerate function on the entire space $\mathbf{V}_U^N$. We conjecture the following:

Conjecture 6.3. *The fibers $\mathbf{M}_\mathbf{z}$ provide a uniformly smooth foliation of $\mathbf{M}_U^N$.*

The global structure of $\mathbf{M}_\mathbf{z}$ depends on the multiplicities of the spectrum. If all eigenvalues are simple then it is diffeomorphic to $(\mathbb{R} \times \mathbf{S}^1)^N$. Moreover the induced diffeomorphisms obtained by choosing a function in $\mathbf{M}_\mathbf{z}$ and flowing it with the first $2N$ flows are uniformly smooth.

The topology of the fiber is different if there are multiplicities: Some of the S^1 components became real lines as the spectral parameters approach multiplicities. As a consequence the smooth structure defined by the flow maps cannot give a uniformly smooth parametrization as we approach multiple eigenvalues.

All the complications above occur already for the case of two solitons, which we will study the two soliton case in depth in Section 9. The two soliton manifold with a double eigenvalue z_0 is diffeomorphic to

$$\mathbb{R}^3 \times \mathbb{S}^1.$$

For instance if $n = 2$, $z_0 = i$ and $|\beta_2| + |\beta_3| \gg 1$ then the soliton distance is about

$$R + i\theta \approx \log(-2(\beta_2 + i\beta_3)).$$

Here R denotes the distance between bump locations, and θ is the phase shift between the two bumps. Hence (β_2, β_3) with the Euclidean topology cannot uniformly describe the soliton distance.

We continue with the properties of soliton removal map:

Theorem 6.4. *a) The soliton removal map*

$$\mathbf{V}_U^N \ni v \to \mathbf{B}_-^N(v) = (u, \mathbf{z}, \boldsymbol{\beta}) \in \mathbf{V}_U^0 \times \mathbf{S}_U^N$$

of removing N solitons is smooth, uniformly on compact sets in $\boldsymbol{\beta}$ and $\mathbf{z}$ in a compact subset of U^N. The map commutes with the flows in the same sense as for the soliton addition map.

As a corollary of these two results, we have:

Theorem 6.5. *The soliton addition map*

$$\mathbf{B}_+^N : \mathbf{V}_U^0 \times \mathbf{S}_U^N \to \mathbf{V}_U^N$$

is a local diffeomorphism with respect to the smooth structure of H^s for all $s > -\frac{1}{2}$.

A natural question to ask here is whether the soliton addition and removal maps are uniformly smooth globally, for u restricted to a bounded set. For this to be meaningful, one has to use the phase space $\mathbf{S}_U^N$ endowed with the Riemannian metric induced from the pure N-soliton manifold $\mathbf{M}_U^N$. We conjecture the following:

Conjecture 6.6. *Identifying $\mathbf{S}_U^N$ and $\mathbf{M}_U^N$, the soliton addition and removal maps are uniformly smooth globally,*

$$\mathbf{B}_+^N : \mathbf{V}_U^0 \times \mathbf{M}_U^N \to \mathbf{V}_U^N, \qquad \mathbf{B}_-^N : \mathbf{V}_U^N \to \mathbf{V}_U^0 \times \mathbf{M}_U^N$$

for u restricted to a bounded set in H^s and the spectral parameters $\mathbf{z}$ restricted to a compact subset of U.

The remainder of this section contains proofs of these results, after a preliminary discussion of symmetric functions.

6.2. **Symmetric functions and elementary symmetric polynomials.**

Lemma 6.7. *Let $U \subset \mathbb{C}$ and $V \subset X$ be open, where X is a Banach space. Let $f : U^N \times V \to \mathbb{C}$ be a continuous (C^k, C^∞, analytic) function, so that for every $x \in V$*

(1) $U^N \ni \mathbf{z} \to f(\mathbf{z}, x)$ is invariant under permutations.
(2) $U^N \ni \mathbf{z} \to p(\mathbf{z}, x)$ is holomorphic.

Let $\mathbf{s}(U^N)$ be the (open) range under the map to the first N elementary symmetric functions. Then there exists a continuous (C^k, C^∞, analytic) function $\tilde{f} : \mathbf{s}(U^N) \times V \to \mathbb{C}$ so that for every $x \in X$

(1) $\mathbf{s}(U^N) \ni \mathbf{s} \to \tilde{f}(\mathbf{s}, x)$ is holomorphic.
(2) $\tilde{f}(\mathbf{s}(\mathbf{z})) = f(\mathbf{z}, x)$.

Proof. Let $U^{N,0} \subset U^N$ be the set with pairwise disjoint complex numbers. The map

$$U^0 \ni \mathbf{z} \to \mathbf{s}$$

is locally biholomorphic, in which case the claim is trivial.

Let $\mathbf{z} \in U^N$ be a point where all the variables are identical. The Taylor series for f at the special point converges uniformly in a neighborhood. The partial sums up to degree M are symmetric polynomials f_M which can be written as

$$f_M = q_M(\mathbf{s}).$$

The q_M's converge uniformly in a neighborhood of $\mathbf{s}(\mathbf{z})$ since the same is true for the f_M. The limit is a holomorphic function $\tilde{f}$ in a neighborhood of $\mathbf{s}(\mathbf{z})$.

If the z's are grouped into separated clusters, we can do the same with the elementary symmetric functions for the clusters: Let $U_j \subset \mathbb{C}$ be open sets with compact pairwise disjoint closures. Let $\{z_n\}_{n\le N}$ be N points in the union of the U_j and $(s_j)_{1\le j\le N}$ be the elementary symmetric polynomials. Let N_j be the number of z's in U_j. We claim that the map

$$(s_n) \to (s^j_m)_{1\le j\le J, 1\le m\le N_j}$$

is holomorphic. To see that we note that an integration over a suitable contour yields

$$\lambda^l_m := \sum_{1\le m\le N_j} (z^j_m)^l = \frac{1}{2\pi i}\int_\gamma z^l \frac{f'(z)}{f(z)} dz$$

where $\{s^l_n\}$ and $\{\lambda^l_n\}$ are algebraically diffeomorphic.

The inverse is given by

$$\sum_{n=0}^N (-1)^n s_n z^{N-n} = \prod_{m=1}^M \sum_{n=0}^{N_m} (-1)^n s^n_m z^{N_m-n}.$$

These definitions immediately carry over to the setting with X.

□

6.3. **A first proof of Theorem 6.1(a).** Let u be as in the theorem, $u_1 = u$ and $u_2 = \bar{u}$ and $\mathbf{z}_0 \in U$.

By Lemma 3.1 there exists a neighborhood $V \in H^s(\mathbb{R};\mathbb{C}^2)$ of $\mathbf{u}$ and a neighborhood

$$V_{\mathbf{z}} = \{z : |z - z_{j,0})| < \varepsilon\} \subset U$$

such that the map

$$V \times V_{\mathbf{z}} \ni (v_1, v_2, \mathbf{z}) \to (e^{-\operatorname{Re} zx}\psi_l, e^{\operatorname{Re} zx}\psi_r) \in (L^\infty \cap DH^s \times H^{s+1}) \times (H^{s+1} \times L^\infty \cap DH^s)$$

is analytic with uniformly bounded derivatives. It is an immediate consequence that the iterated Bäcklund transform is analytic in the spectral parameters $\mathbf{z}$, $\boldsymbol{\beta}$ and holomorphic in u, uniformly for bounded $\mathbf{u}$, $\mathbf{z}$ and $\boldsymbol{\beta}$ in compact sets.

This is weaker than the statement of Theorem 6.1 since we claim smoothness in the elementary symmetric polynomials. To see this let

$$\psi_1(x, z_1) = e^{-i\alpha(z_1)}\psi_l(x, z_1) + e^{i\alpha(z_1)}\psi_r(x, z_1)$$

and for $z \in \bar{V}_{\mathbf{z}}$

$$\psi_2(x, z_2) = \begin{pmatrix} \overline{\psi^2_1(x, \bar{z})} \\ -\overline{\psi^1_1(x, \bar{z})} \end{pmatrix}.$$

The iterated Bäcklund transform is now analytic in (u_1, u_2), $\mathbf{z}_1 = \mathbf{z}$ and $\mathbf{z}_2 = \bar{\mathbf{z}}$. Fix (u_1, u_2), $\mathbf{z}_1$ and $\mathbf{z}_2$ pairwise disjoint. The set where the Wronskian vanishes,

$$\{x : W(\psi_1(x, z_{1,j}), \psi_2(x, z_{2,k})) \neq 0\},$$

is the complement of a finite set. By Theorem 3.9 the iterated Bäcklund transform is symmetric under symmetric permutations of the $\mathbf{z}_1$ and $\mathbf{z}_2$ separately for these values of x, and by continuity for all x.

By Lemma 6.7 the iterated Bäcklund transform is a holomorphic function of the elementary symmetric polynomials separately in $\mathbf{z}_1$ and $\mathbf{z}_2$ - more precisely we have to remove a small neighborhood of the set where some Wronskian vanish. The derivatives with respect to the elementary symmetric polynomials are bounded by the derivatives with respect to $\mathbf{z}_1$ reps. $\mathbf{z}_2$, hence we obtain a smooth dependence on the elementary symmetric polynomials.

6.4. **The key regularity lemma.** Here we turn our attention to a second proof of Theorem 6.1, where we aim to provide a more algebraic argument. This proof depends on a result on holomorphic functions, which is virtually independent from the problem at hand. Let $U \subset \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ be an open subset with compact closure in the open upper half plane and $W \subset X$ open. We consider functions $\psi : U \times W \to \mathbb{C}^N$, holomorphic in both arguments and such that

$$|\psi(z_1, w_1)| \le 2|\psi(z_2, w_2)|$$

for $|z_1 - z_2| \ll 1$ and $\|w_1 - w_2\|_X \ll 1$. Using the Cauchy integral we deduce that

$$|\partial^\gamma_{z,w}\psi(z,w)| \le C_{|\gamma|}(1 + d^{-|\gamma|})|\psi(z,w)| \tag{6.4}$$

where d is the distance from (z,w) to the complement of $U \times W$. Let $\delta > 0$, $\phi^1, \phi^2 : U \times W \to \mathbb{C}$ holomorphic with

$$|\phi^1(z,w)| \le \delta|\psi(z,w)| \qquad |\phi^2(z,w)| \le \delta|\psi(z,w)|. \tag{6.5}$$

Let $U_0 \subset U^N$ the subset of pairwise disjoint tuples and $V_0 = \mathbf{s}(U_0)$ resp $V = \mathbf{s}(U)$. We define the Hermitian matrix M by

$$M_{jk} = \frac{i\psi^*(z_k,w)\psi(z_j,w)}{\bar z_k - z_j},$$

and denote by m its inverse. Then we define the function $g : U_0 \times W \to \mathbb{C}$ by

$$g(\mathbf{s}, w) = \sum_{j,k=1}^N \phi^{1*}(z_j,w) m_{jk}\phi^2(z_k,w), \tag{6.6}$$

where $\mathbf{s}$ denotes the elementary symmetric functions.

Lemma 6.8. *The function g has a unique analytic extension to $U^N \times W$. Moreover*

$$|\partial^\gamma_{\mathbf{s},w} g(\mathbf{s},w)| \le c_{N,|\gamma|}\delta(1 + d^{-N|\gamma|}).$$

Here the power of d is controlled via the Cauchy integral and Lemma 5.1.

Proof. At a given point $(\mathbf{z}_0, w_0)$ we divide both ϕ_j and ψ by $|\psi(z_0,w_0)|$ and we may assume that $|\psi(z_0,w_0)| = 1$ in the sequel.

We substitute $\tilde z_j$ for $\bar z_j$ and $\tilde w$ for $\bar w$ and define

$$M_{jk} = i\frac{(\overline{\psi(\overline{\tilde z_k}, \overline{\tilde w})} \cdot \psi(z_j,w)}{\tilde z_k - z_j},$$

which is a holomorphic function of $\mathbf{z}, \widetilde{\mathbf{z}}, w$ and $\tilde w$. Similarly we extend the function g. Let C be the cofactor matrix of M. Then we can write

$$m = \frac{1}{\det M} C^T$$

and

$$g(\mathbf{s}, \tilde{\mathbf{s}}, w, \tilde w) = \overline{\phi^2(\overline{\tilde z_k})} m_{kj}\phi^1(z_j) = \frac{\overline{\phi^2(\overline{\tilde z_k}, \tilde w)} C^T_{kj}\phi^1(z_j,w)}{\det M}.$$

Now we consider the symmetry properties of both the numerator and the denominator. Permuting two values z_j and z_k exchanges to rows in M and permuting two values $\tilde z_j$ and $\tilde z_k$ exchanges two columns. Under either operation $\det(M)$ changes sign. Since

$$MC^T = C^T M = \det(M)1$$

we see that interchanging z_j and z_k exchanges two rows in C and the whole sign of C. As a consequence both $\det M$ and

$$\overline{\phi^2(\overline{\tilde z_k})} C^T_{kj}\phi^1(z_j),$$

considered as holomorphic functions of $\widetilde{\mathbf{z}}$ and $\mathbf{z}$, are antisymmetric in the z_j. Then we can smoothly factor

$$\overline{\phi^2(\overline{\tilde{z}_k}, \overline{\tilde{w}})} C^T_{kj} \phi^1(z_j, w) = G(\mathbf{z}, \tilde{\mathbf{z}}, w, \tilde{w}) \prod_{j \neq k} (z_j - z_k) \prod_{j \neq k} (\tilde{z}_j - \tilde{z}_k).$$

The same applies for $\det M$,

$$\det M = H(\mathbf{z}, \widetilde{\mathbf{z}}, w, \tilde{w}) \prod_{j \neq k} (z_j - z_k) \prod_{j \neq k} (\tilde{z}_j - \tilde{z}_k).$$

Here the functions G and H are holomorphic functions in $\mathbf{z}$ and $\tilde{\mathbf{z}}$, and separately symmetric in both $\mathbf{z}$ and $\tilde{\mathbf{z}}$ (this is the reason we separated the variables $\mathbf{z}$ and $\bar{\mathbf{z}}$ in the first place). By Lemma 6.7, every symmetric holomorphic function is a holomorphic function in the elementary symmetric polynomials. Hence by a slight abuse of notations we will write

$$G(\mathbf{z}, \tilde{\mathbf{z}}, w, \tilde{w}) = G(\mathbf{s}, \tilde{\mathbf{s}}, w, \tilde{w}), \qquad H(\mathbf{z}, \tilde{\mathbf{z}}, w, \tilde{w}) = H(\mathbf{s}, \tilde{\mathbf{s}}, w, \tilde{w}).$$

We still need to divide these two functions G and H. To do this we return to the diagonal $\tilde{\mathbf{z}} = \bar{\mathbf{z}}$, $\tilde{w} = \bar{w}$ and claim that there we have

$$H(\mathbf{z}, \overline{\mathbf{z}}, w, \bar{w}) \gtrsim 1. \tag{6.7}$$

But there we can take advantage of the positivity of M. We can write M at $w = \bar{z}$ as the sum of M positive matrices which correspond to the components of the ψ's,

$$M_{jk} = \sum_{m=1}^{M} \frac{i\psi^m(z_j)\overline{\psi^m(z_k)}}{\bar{z}_k - z_j} =: \sum_{m=1}^{M} M^m_{jk}.$$

Here, to insure nondegeneracy, we first rotate the coordinates so that all components $\psi^m(z_j)$ are of comparable size, which we can do if $m \geq 2$.

Then

$$M^m = \begin{pmatrix} \psi^m(z_1) & 0 & \dots & 0 \\ 0 & \psi^m(z_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi^m(z_M) \end{pmatrix} \left(\frac{i}{\bar{z}_k - z_j} \right)_{jk} \begin{pmatrix} \overline{\psi^m(z_1)} & 0 & \dots & 0 \\ 0 & \overline{\psi^m(z_2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \overline{\psi^m(z_M)} \end{pmatrix}$$

and

$$\det M^m = \prod_{j=1}^{n} |\psi^m(z_j)|^2 \frac{\prod_{j<k}(z_j - z_k)(\bar{z}_j - \bar{z}_k)}{\prod_{j,k}(\bar{z}_j - z_k)}$$

Since we have bound one of these these determinants from below, we can also bound from below the determinant of the sum, thereby proving our claim (6.7).

Now we write, again on the diagonal,

$$g(\mathbf{z}, w) = \frac{G(\mathbf{s}, \bar{\mathbf{s}}, w, \bar{w})}{H(\mathbf{s}, \bar{\mathbf{s}}, w, \bar{w})}$$

where both the numerator and denominator are smooth and the denominator is bounded from below. The conclusion of the Lemma follows.

□

6.5. **The pointwise regularity of soliton addition.** This is the first part of the second proof of both part (a) and (b) of Theorem 6.1, where we consider the pointwise regularity of the soliton addition map. We consider $u \in \mathbf{V}_U^0$ and $(\mathbf{z}, \boldsymbol{\beta}) \in \mathbf{S}_U^N$, and let $v = \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})$. For arbitrary fixed $x \in \mathbb{R}$ we study the dependence of $v(x)$ on $u, \mathbf{z}, \boldsymbol{\beta}$. Here $v(x)$ is obtained via the following steps:

(1) Since the transmission coefficient $T(u)$ has no poles in U, it follows that the left and right Jost functions $\psi_l(u, z)$ and $\psi_r(u, z)$ are uniformly independent (the Wronskian is the inverse of $T(z)$), analytic in u and uniformly holomorphic in z.

(2) Given $\boldsymbol{\beta} \in \mathbb{R}^{2N}$, we produce a holomorphic family of wave functions by setting
$$\psi(u, z, \boldsymbol{\beta}) = e^{-i\boldsymbol{\beta}(z)}\psi_l(u, z) + e^{i\boldsymbol{\beta}(z)}\psi_r(u, z).$$

(3) Based on our interpretation of the soliton addition map in terms of the holomorphic wave function, it follows that
$$\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}) = \mathbf{B}_{\mathbf{z},\psi}^N(u).$$
The latter expression can be viewed as the outcome of N Bäcklund transforms, so it is analytic in the z_j's, in u and $\bar{u}$, and in α (via ψ).

(4) If the z's are distinct, then we can use the results in Section 4 to obtain an expression for $v - u$ as
$$v - u = \sum_{j,k=1}^{N} \bar{\psi}^1(z_j, u) m_{jk} \psi^2(z_k, u),$$
where m is the inverse of the matrix M given by
$$M_{jk} = \frac{i\psi^*(z_k, u)\psi(z_j, u)}{\bar{z}_k - z_j}.$$

(5) Now we use Lemma 6.8 to conclude that this expression has an uniformly smooth extension to the diagonal, as a function of $\mathbf{s}$, $\boldsymbol{\beta}$ and u.

By the above considerations we have defined a soliton addition map $\mathbf{B}_+^N$ on $\mathbf{V}_U^0 \times \mathbf{S}_U^N$ which is smooth for fixed x. At this point we are still lacking uniformity both with respect to x, u, $\mathbf{z}$ and $\boldsymbol{\beta}$.

We now consider the question of uniformity. To start with, we need to describe more accurately the Jost functions ψ_l, ψ_r. Consider ψ_l for instance. Taking out the exponentials, we set $\tilde{\psi}_l = e^{izx}\psi_l$, for which we have the ode's

$$\begin{cases} \dot{\tilde{\psi}}_l^1 = & v\tilde{\psi}_l^2 \\ \dot{\tilde{\psi}}_l^2 = & 2iz\tilde{\psi}_l^2 + \bar{v}\tilde{\psi}_l^1 \end{cases}. \tag{6.8}$$

The map
$$H^s \ni u \to \tilde{\psi}_l^1 \in C_b(\mathbb{R})$$
is uniformly smooth in u (see Lemma 3.1), and also uniformly holomorphic in z. In a symmetric way we set $\tilde{\psi}_r = e^{-izx}\psi_r$ with similar properties.

With these notations, our holomorphic family of wave functions becomes
$$\psi(u, z, \boldsymbol{\beta}) = e^{-i(\boldsymbol{\beta}(z)+zx)}\tilde{\psi}_l(u, z) + e^{i(\boldsymbol{\beta}(z)+zx)}\psi_r(u, z).$$

To gain uniformity, we first assume that $\boldsymbol{\beta}$ is in a compact set. Then the second term above is leading if $x < 0$ while the first term is dominant when $x > 0$. Suppose for instance that $x > 0$. Then we take out the second exponential factor, and redefine the holomorphic family of wave functions as
$$\psi(u, z, \boldsymbol{\beta}) = \tilde{\psi}_l(u, z) + e^{2i(\boldsymbol{\beta}(z)+2zx)}\psi_r(u, z).$$
Our choice of x now insures that for this family we have uniform regularity at x with respect to all parameters. Then, by Lemma 6.8, we obtain the corresponding uniform regularity for $v(x)$, as stated in part (a) of the theorem.

Next we move to part (b) of the theorem. We fix $\mathbf{z}$, and consider the question of uniform regularity in u and $\boldsymbol{\beta}$. Differentiating $v-u$ we obtain a representation

$$\partial_u(v-u)(x) = \partial_u \psi^{1*} m \psi^2 + \psi^{1*} m \partial_u \psi^2 - \psi^{1*} m \partial_u M m \psi^2,$$

and similarly for the $\boldsymbol{\beta}$ derivatives. For higher derivatives with respect to $\boldsymbol{\beta}$ and u we obtain a similar but longer expansion but with more instances of m separated by differentiated M. For each differentiated $\partial^k M$ (in either u or $\boldsymbol{\beta}$) we can separate the variables $\mathbf{z}$ and $\bar{\mathbf{z}}$ (e.g. using the exponential representation) and represent them as rapidly convergent sums (integrals) of terms of the form

$$\partial^{k_1}\psi_j g(z_j) \otimes \partial^{k_2}\psi_j^* g(\bar{z}_j).$$

By Cauchy-Schwartz, it remains to obtain a uniform bound for expressions of the form

$$\partial^k \psi_j^* \overline{g(z_j)}\, m_{jn}\, \partial^k \psi_n g(z_n).$$

For the two components of ψ we have the regularity

$$\partial^k \left[e^{\boldsymbol{\beta}(z)} \psi_l(u,z) \right] = e^{\boldsymbol{\beta}(z)+izx} f^k(z,u),$$

where f^k is uniformly holomorphic in z.

So we need to bound uniformly an expression of the form

$$\overline{e^{\boldsymbol{\beta}(z_j)+iz_jx} f(z_j)}\, m_{jn}\, e^{\boldsymbol{\beta}(z_n)+iz_nx} f(z_n),$$

where f is holomorphic.

If we could simply discard the ψ_r component of ψ, then we would factor out the phase $e^{\boldsymbol{\beta}(z)+izx}$ and then just apply Lemma 6.8. As it is, we can still use each of the two components of ψ to define its own non-negative matrix M^1, respectively M^2 so that $M = M^1 + M^2$. correspondingly we get $m \leq m^1$ and $m \leq m^2$. Applying Lemma 6.8 to each of these components, it follows that we could bound m on vectors of the form

$$b_1(z)\psi_1, \qquad b_2(z)\psi_2,$$

with b_1, b_2 holomorphic. It remains to see that we can obtain a representation

$$e^{\boldsymbol{\beta}(z)+izx} f(z) = b_1(z)\psi^1 + b_2(z)\psi^2.$$

Cancelling phases this is equivalent to

$$f(z) = b_1(z)(\tilde{\psi}_l^1 + e^{-2\boldsymbol{\beta}(z)-2izx}\tilde{\psi}_r^1) + b_2(z)(\tilde{\psi}_l^2 + e^{-2\boldsymbol{\beta}(z)-2izx}\tilde{\psi}_r^2).$$

Here we do not want b_1 or b_2 to depend on the exponentials, for that would likely make them unbounded. So we strengthen the above relation to a system

$$\begin{aligned} b_1(z)\tilde{\psi}_l^1 + b_2(z)\tilde{\psi}_l^2 &= f, \\ b_1(z)\tilde{\psi}_r^1 + b_2(z)\tilde{\psi}_r^2 &= 0, \end{aligned}$$

This is uniformly solvable since the Wronskian of $\tilde{\psi}_l$ and $\tilde{\psi}_r$ is constant and of size $O(1)$ for $z \in U$, because $u \in \mathbf{V}_U^N$.

6.6. **The H^s regularity of soliton addition.** Theorem 6.1 claims uniform smoothness for the soliton addition map as a map to H^s in two contexts, corresponding to part (a) and part (b). At this point we know that, in both contexts, for each x the map

$$\mathbf{V}_U^0 \times \mathbf{S}_U^N \ni (u, \mathbf{s}, \boldsymbol{\beta}) \to v(x) - u(x) \in \mathbb{C}$$

is smooth, with appropriate uniformity statements. The next step is to prove similar H^s bounds for $u - v$ and its linearization. To achieve this, we divide and conquer. We split the real axis into unit intervals, and seek to understand the H^s regularity within each interval. For a reference point x, we study the H^s regularity of $u - v$ in the interval $I = (x-1, x)$.

We follow the analysis in the previous subsection, but working on unit intervals instead of at a fixed point x. One can think of the construction as having two stages:

(i) From the data u to the renormalized Jost functions $\tilde{\psi}_l$, $\tilde{\psi}_r$.

(ii) From the renormalized Jost functions to $v - u$.

As long as $u \in \mathbf{V}_U^0$ and $z \in U$, the map

$$H^s \ni u \to \tilde{\psi}_l, \tilde{\psi}_r \in H^{s+1}(I)$$

is holomorphic in z and analytic in u. Then the same argument as in the previous subsection shows that the soliton addition map

$$\mathbf{V}_U^0 \times \mathbf{S}_U^N \ni (u, \mathbf{s}, \boldsymbol{\beta}) \to v - u \in H^{s+1}(I)$$

is analytic, with the same uniformity statements as before.

The new difficulty we face here is in the transition from the local H^s regularity to the global H^s regularity. For this we need to gain the ℓ^2 summation with respect to unit intervals. We remark that this gain is not straightforward, i.e. it does not happen at the level of $\tilde{\psi}_l, \tilde{\psi}_r$. Instead, the best we can say is that we have the localized bounds

$$\|\tilde{\psi}_l^2\|_{H^{s+1}(x-1,x)} \leq c\|e^{\operatorname{Im} z(y-x)} u(y)\|_{H^s(-\infty,x)}$$

and

$$\|\tilde{\psi}_l^1(x) - \tilde{\psi}_l^1(.)\|_{H^{s+1}(x-1,x)} \leq c\|e^{\operatorname{Im} z(y-x)} u(y)\|_{H^s(-\infty,x)},$$

with an implicit constant depending only on $\|u\|_{H^s}$. Similar bounds will hold for the linearizations. Excluding finitely many intervals where u might concentrate, we can assume that we also have smallness,

$$\|\tilde{\psi}_l^2\|_{H^{s+1}(x-1,x)} + \|\tilde{\psi}_l^1(x) - \tilde{\psi}_l^1(.)\|_{H^{s+1}(x-1,x)} \ll 1$$

This in turn gives pointwise smallness, and thus a bound from below

$$|\tilde{\psi}_l^1(x)| \gtrsim 1.$$

Similarly, we will have

$$|\tilde{\psi}_r^2(x)| \gtrsim 1.$$

Hence, on the interval $(x-1, x)$ it is natural to compare the renormalized Jost functions $\tilde{\psi}_l$ and $\tilde{\psi}_r$ with $\tilde{\psi}_l^1(x)e_1$, respectively $\tilde{\psi}_r^2(x)e_2$. Thus, within the interval I we arrive at a reference configuration which corresponds to a pure soliton. However, this is not the soliton with parameters $(\mathbf{z}, \boldsymbol{\beta})$; that would correspond to having $\tilde{\psi}_l^1 = 1$ and $\tilde{\psi}_r^1 = 1$. Instead $\boldsymbol{\beta}$ is readjusted to

$$\tilde{\boldsymbol{\beta}}(x) = \boldsymbol{\beta} + \frac{1}{2}\log(\tilde{\psi}_l^1(x)/\tilde{\psi}_r^2(x)).$$

This also should be seen as a function of z and u.

We denote by $Q_{\mathbf{z},\boldsymbol{\beta}}$ the pure soliton with parameters $(\mathbf{z}, \boldsymbol{\beta})$. Then our analysis above allows us to conclude that we have the localized bound

$$\|v - u - Q_{\mathbf{z},\tilde{\boldsymbol{\beta}}(x)}\|_{H^{s+1}(I)} \lesssim \|\operatorname{sech}[\delta(y-x)]u(y)\|_{H^s_y}, \qquad 0 < \delta < \min\{\operatorname{Im} z_j\}.$$

Similar bounds will also hold for the linearization.

To conclude, we need to show that the square summability in I survives as we vary the soliton parameter $\tilde{\boldsymbol{\beta}}$. The key property here is that $\tilde{\boldsymbol{\beta}}$ does not vary much, $|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}| \lesssim 1$. Hence it suffices to verify the property

$$\sum_I \sup_{|\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}|\lesssim 1} \|Q_{\mathbf{z},\tilde{\boldsymbol{\beta}}}\|^2_{H^{s+1}(I)} \lesssim 1$$

Here the value of s is not important. But this is easy to see, as the pure N-solitons with spectral parameters in U are uniformly bounded in all H^s spaces, while the change in β corresponds to the flow along the first $2N$ commuting flows, so the β derivative of $Q_{\mathbf{z},\boldsymbol{\beta}}$ is also uniformly bounded in all H^s norms. In effect in the next section we prove that the N-solitons are exponentially decaying away from at most N bumps, so the same applies to the localized norms in the above formula.

6.7. **The multisoliton manifold.** This subsection is devoted to the proof of Theorem 1.5 resp. 6.2. We begin with the case $s = 0$, where the notations are simpler. The argument for other H^s spaces with $s > -\frac{1}{2}$ is similar, and is outlined at the end of the section.

We recall that the family $\mathbf{M}_U^N$ of pure N-solitons can be described using the soliton addition map $\mathbf{B}_+^N$,

$$\mathbf{M}_U^N = \{\mathbf{B}_+^N(0, \mathbf{s}, \boldsymbol{\beta});\ (\mathbf{s}, \boldsymbol{\beta}) \in \mathbf{S}_U^N\}.$$

as a subset of the set of N-soliton states $\mathbf{V}_U^N$

$$\mathbf{V}_U^N = \{v = \mathbf{B}_+^N(u, \mathbf{s}, \boldsymbol{\beta});\ u \in \mathbf{V}_U^0,\ (\mathbf{s}, \boldsymbol{\beta}) \in \mathbf{S}_U^N\}.$$

On $\mathbf{V}_U^N$ we define the real valued map

$$F : \mathbf{V}_U^N \ni v \to F(v) = E_0(v) - 2\sum_k \operatorname{Im} z_k = \|v\|_{L^2}^2 - \sum_k 2\operatorname{Im} z_k. \tag{6.9}$$

which gives the soliton free L^2 energy of v. In view of the trace formula (1.2), the map F is obviously uniformly smooth, non-negative and it vanishes on pure solitons. Thus also its derivative vanishes at pure N solitons. We claim that the following three properties hold:

(1) the N-soliton manifold can be described as

$$\mathbf{M}_U^N = \{v \in \mathbf{V}_U^N : DF(v) = 0\}, \tag{6.10}$$

(2) the Hessian of F evaluated at $v = \mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta}) \in \mathbf{M}_U^N$ is nondegenerate,

$$D^2F(v)[w, w] \geq C^{-1}\|w\|_{L^2}^2, \qquad w \in \operatorname{Range}(D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})), \tag{6.11}$$

(3) the u differential of $\mathbf{B}_+^N$ at $u = 0$ is nondegenerate,

$$\|w\|_{L^2} \leq C\|D_u D^N(0, \mathbf{z}, \boldsymbol{\beta})w\|_{L^2}, \qquad w \in L^2. \tag{6.12}$$

We proceed to prove these three claims. Since F vanishes quadratically at pure solitons, we must have $DF(v) = 0$ whenever v is a pure soliton. Now suppose that $DF(v) = 0$ with $v = \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})$. The trace identities imply that

$$F(\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})) = \|u\|_{L^2}^2, \tag{6.13}$$

hence, differentiating in the w direction,

$$2\operatorname{Re}\int uw\, dx = D_u(F \circ \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}))|_{u=0}(w) = DF(\mathbf{B}_+^N(u, \mathbf{z}, \beta))D_u\mathbf{B}_+^N(u, \mathbf{z}, \beta)(w),$$

which vanishes for $w = u$ only if $u = 0$, or, equivalently, if $D^N(u, \mathbf{z}, \boldsymbol{\beta})$ is a pure N soliton. This implies the claim (6.10).

Moreover, we can also calculate the Hessian in (6.13) as

$$\begin{aligned} 2\|w\|^2 &= D_u^2(F \circ \mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}))[w, w] \\ &= D_v^2 F(\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}))][D_u\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})w, D_u\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})w] \\ &\quad + D_v F(\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta}))D_u^2\mathbf{B}_+^N(u, \mathbf{z}, \boldsymbol{\beta})[w, w]. \end{aligned}$$

We evaluate this formula at pure N solitons, using the fact that DF vanishes there:

$$\begin{aligned} 2\|w\|^2 &= D_v^2 F \circ \mathbf{B}_+^N(0, \mathbf{s}, \boldsymbol{\beta})[D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})w, D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})w] \\ &= \|D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})w\|^2 - 2D^2 \operatorname{Im} s_1(D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})w, D_u\mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta})w). \end{aligned}$$

The map to the elementary symmetric functions is given by a nondegenerate contour integral of the transmission coefficient, and hence it is uniformly smooth. Thus we obtain (6.12), By Theorem 6.1 the two norms in (6.12) must be equivalent, so (6.11) also follows.

This gives important information on the uniformly smooth maps $v \to F(v)$ and $v \to DF(v)$. Let $R(\mathbf{z}, \boldsymbol{\beta})$ be the range of $D_u D^N(0, \mathbf{z}, \boldsymbol{\beta})$, which by (6.12) is a closed subspace of L^2, with codimension $4N$. By (6.11) D^2F is positive definite on this subspace. Thus D^2F defines a linear map from L^2 to L^2 with a $4N$ dimensional null space. The restriction to $R(\mathbf{z}, \boldsymbol{\beta})$ defines a uniformly invertible operator. Then by the implicit function theorem the set $\{v \in L^2;\ DF(v) = 0\}$ is a uniformly smooth $4N$-dimensional manifold, which concludes the proof of the theorem.

6.8. **Regularity of soliton removal.** Here we give the proof of Theorem 6.4.

6.8.1. *The spectrum.* Let $v \in \mathbf{V}_U^N$ be as above, or equivalently, assume that $\mathcal{L}(v)$ has exactly N eigenvalues (counting with multiplicity) in U. These are denoted by $\mathbf{z} = \{z_j\}$ and can be described as the poles of T, not necessarily distinct. We call

$$P(z) = \prod_{j=1}^{N} (z - z_j) = \sum_{n=0}^{N} (-1)^n s_n z^{N-n}$$

the characteristic polynomial. As seen in Lemma 5.1, the relation between the symmetric polynomials and the roots is Hölder continuous but not smooth. The next lemma shows that these polynomials can be smoothly recovered from T, which in turn depends smoothly on v away from the poles.

Lemma 6.9. *Let $U \subset \mathbb{C}$ be open, X be a complex Banach space, $W \subset X$ open and*

$$f : U \times W \to \mathbb{C}$$

holomorphic. Let $K \subset U$ be compact, $w_0 \in W$ such that $f(z, w_0) \neq 0$ for $z \in \partial K$. Then there exists $\varepsilon > 0$ so that $f(z, w)$ does not vanish for $z \in \partial K$, $|w - w_0| < \varepsilon$. The number of zeroes in K of $f(., w)$ is independent of w. Let $(s_n(w))_{n \leq N}$ be the elementary symmetric polynomials of the roots. Then

$$B_\varepsilon(w_0) \ni w \to s_n(w) \in \mathbb{C}$$

is holomorphic.

Proof. We may assume that ∂K is a union of closed nonintersecting positively oriented C^1 Jordan curves γ. Then

$$\lambda_k := \sum_{n=1}^{N} z_n^k = \frac{1}{2\pi i} \int_\gamma \zeta^k \frac{\partial_z f(\zeta, w)}{f(\zeta, w)} d\zeta.$$

Then each s_n can be written as a polynomial in λ_k and vice versa. □

6.8.2. *Regularity of soliton removal: Finding the scattering data.* Next we consider the question of recovering $\boldsymbol{\beta}$. For this we use the left and right Jost functions, and recall that $\boldsymbol{\beta}(z_j)$ is the proportionality constant between them at the poles, and should be accurate to the order of the pole,

$$\psi_l(x, z) + e^{2i\boldsymbol{\beta}(z)} \psi_r(x, z) = O(z - z_j)^{m_j} \qquad \text{for } z \text{ near } z_j \text{ and } x \text{ in a compact set.}$$

To define $\boldsymbol{\beta}$ we fix some $x \in \mathbb{R}$ and compare $\psi_l(z)$ and $\psi_r(z)$ at x. Each of these two values depends analytically on z and also on v for v near v_0.

Hence, near each z_j^0 we find a ball B_j where either we have $|\Psi_l^1((z, x)| \geq \frac{1}{2} |\Psi_l((z, x)|$ or $|\Psi_l^2((z, x)| \geq \frac{1}{2} |\Psi_l((z, x)|$. To fix the notations assume the former.

Here the size r of each B_j depends on the Lipschitz constant for $|\psi_l|^{-1}\psi_l$ in z at the point x. These balls can overlap, and we identify them if the centers are much closer, i.e.

$$B_j = B_k \qquad \text{if } \ |z_j - z_k| \ll r$$

and

$$2B_j \cap 2B_k = \emptyset \qquad \text{if } \ |z_j - z_k| \gg r$$

We choose r so that these are the only alternatives. The same will hold not only for v_0, but also for v in a neighbourhood. Then we locally define the function $\boldsymbol{\beta}_0$ as

$$e^{2i\boldsymbol{\beta}_0(z)} = -\frac{\psi_l^1(z)}{\psi_r^1(z)}.$$

(or using the second component, or a linear combination, whichever works for v near some given state v_0.) This is holomorphic near z_j, with a smooth local dependence on v. By the Chinese remainder theorem (see Lemmas 5.4, 5.5) there exists a unique real polynomial $\boldsymbol{\beta}(z)$ of degree at most $2N-1$, so that

$$\boldsymbol{\beta} = \boldsymbol{\beta}_0 \qquad (\mod P_{\mathbf{z}} P_{\bar{\mathbf{z}}}).$$

This will define the scattering parameters $\boldsymbol{\beta}$ for v, in a manner that depends smoothly on $v \in H^s$.

We go one step further, and also define a corresponding holomorphic family of unbounded wave functions by setting

$$\psi(z) = T_v(z)(e^{-i\boldsymbol{\beta}(z)}\psi_l(x,z) + e^{i\boldsymbol{\beta}(z)}\psi_r(x,z)).$$

This will also depend smoothly in H^{s+1}_{loc} on $v \in H^s$. This suffices for the local regularity, but we also need to investigate more carefully what happens near $\pm\infty$. Consider or instance a neighborhood of ∞. We can localize v there to $\tilde{v}$, which is now small in H^s. Then we can write the wave function ψ as a wave function for $\tilde{v}$, with scattering parameter $\tilde{\beta}$ which depends smoothly on $v \in H^s$.

6.8.3. *Smooth soliton removal: Finding the background.* Once the spectral and scattering parameters are smoothly recovered, we can recover also v in terms of u following the removal transformation. To see that we can start by density with the case of distinct eigenvalues. In this case the iterated soliton removal maps are smooth with respect to u and the result is independent of the order.

The case of multiple eigenvalues is obtained as a limit, using Lemma 6.8, since the unbounded wave functions ψ obtained above have a smooth dependence on u and are holomorphic in z. This yields pointwise bounds. The H^s bounds can be dealt with exactly as in the case of the soliton addition map. This splits into two parts: (i) locally, which is exactly the same as before, and (ii) near infinity, where, as discussed above, this is identical to the corresponding argument for the soliton addition map.

6.8.4. *Soliton removal: Uniformity for $\boldsymbol{\beta}$ in a compact set.* What changes here is that the left and right Jost functions have a single bump which depends nicely on v. The location of the bump depends only on $\boldsymbol{\beta}_1$. Then we choose x_0 near this peak, which insures that $\boldsymbol{\beta}$ depends uniformly smoothly on v, and also that ψ has a similar dependence on v away from the bump. The rest is similar to the soliton addition map.

7. The structure of solitons

Single solitons can be seen as bump functions, with uniform exponential decay away from the center of the bump. Here we investigate the similar question for multisolitons. Precisely, we will show that each N-soliton can be viewed as a collection of at most N unit sized bumps, with exponential decay in between and at infinity. Forthermore, each of these bumps has to be exponentially close to a lower dimensional soliton.

Our main result concerning the structure of N multisolitons is as follows:

Theorem 7.1. *a) The N multisoliton solutions are functions with exactly N bumps (possibly overlapping), and exponential decay away from these bumps.*

b) If bumps separate into k groups at distance at least R, then the multisoliton can be approximately viewed as the sum of k multisoliton solutions, with an accuracy of $O(e^{-cR})$.

Proof. a) Let $Q = Q_{\mathbf{z},\boldsymbol{\beta}}$ be an N-soliton and $P_{\mathbf{z}}$ its characteristic polynomial. Denote by $R = P_{\mathbf{z}}P_{\bar{\mathbf{z}}}$, which has real coefficients. Consider the action of the n-th flow on Q, or more precisely on $\boldsymbol{\beta}$. This gives

$$\dot{\boldsymbol{\beta}} = i2^{n-1}z^n \qquad (\mod R).$$

It follows that the first $2N$ flows are linearly dependent when acting on Q.

Precisely, to any real polynomial

$$R(z) = \sum_{j=0}^{2N} r_j (2z)^j$$

we can associate the Hamiltonian

$$H_R = \sum_{j=0}^{2N} r_j H_j.$$

Then for $R = P_{\mathbf{z}}P_{\bar{\mathbf{z}}}$ defined above, we know that Q is a steady state for the H_R flow. This is equivalent to

$$DH_R(Q) = 0$$

which shows that Q solves a semilinear ODE of order $2N$. The linear part of this ODE is given by the operator $R(D_x) = R(\frac{1}{i}\partial_x)$. So we can rewrite this ODE in the form

$$R(D_x)Q = N(Q^{(\leq 2N-1)}) \tag{7.1}$$

Equivalently, we can rewrite this as a first order system for the variables

$$y = \{D^j Q; j = 0, 2N-1\},$$

namely

$$\partial_x y = iAy + N(y), \tag{7.2}$$

where the matrix A has R as a characteristic polynomial and N is polynomial and contains quadratic and higher order terms.

The state $y = 0$ is a fixed point for the system (7.2), and the eigenvalues for the linearization around $y = 0$ are $\pm i z_k$, neither of which is on the imaginary axis. Hence 0 is a hyperbolic fixed point for this dynamical system. Hence, by the Hartman-Grobman theorem, the dynamics around $y = 0$ are well described by the corresponding linearized flow, up to a local Hölder continous homeomophism with Hölder continuous inverse.

Now we are able to complete the qualitative description of the solitons. We consider the localized mass of Q in unit intervals $I_j = [j, j+1]$,

$$M_j = \int_{I_j} |Q|^2 \, dx,$$

In intervals where M_j is small, all of the Cauchy data of Q must be small so the Hartman-Grobman theorem applies. But the total mass is finite, so there can be only finitely many intervals where M_j is large. Outside this finite number of intervals, the soliton Q must follow the linearized dynamics and decay exponentially. This completes the proof of part (a) of the theorem, modulo the counting of the bumps; we still need to show that, if ϵ is small enough, then there are at most N regions where $|Q| > \epsilon$. This will follow as a corollary of the proof in (b).

b) Denote by $R \gg 1$ the smallest gap between two bumps. Then in between each two bumps, we will find a smallest value for y,

$$|y(x_k)| \lesssim e^{-cR}, \qquad k = 1, K.$$

Away from x_k, y will grow exponentially. We use the x_k as sharp cut points for Q, splitting it on the intervals $I_k = (x_k, x_{k+1})$ where $x_0 = -\infty$ and $x_{K+1} = +\infty$,

$$Q = Q_1 + \cdots + Q_{K+1}, \qquad Q_k = 1_{I_k} Q.$$

On one hand, we have the obvious energy relation

$$\|Q\|_{L^2}^2 = \sum \|Q_k\|_{L^2}^2.$$

On the other hand, we investigate the relation between the transmission coefficients of Q and those of Q_j. We work with z away from the spectral parameters $\mathbf{z}$ of Q, say on a contour γ around $\mathbf{z}$. For such z, the renormalized Jost function $\tilde{\psi}_l$ associated to Q satisfies

$$|\tilde{\psi}_l| \gtrsim 1, \qquad \lim_{x \to \infty} \tilde{\psi}_l = T_Q^{-1}(z).$$

Furthermore, around the points x_k the coupling between the two components of the $\tilde{\psi}_l$ equation (6.8) is exponentially small, therefore we also obtain the exponential smallness

$$|\tilde{\psi}_l^2(x_k)| \lesssim e^{-cR}, \qquad |\tilde{\psi}_r^1(x_k)| \gtrsim 1.$$

Next, for the same $z \in \gamma$ we consider the corresponding renormalized Jost function $\tilde{\psi}_{j,l}$ for Q_j and the associated transmission coefficients $T_{Q_j}^{-1}(z)$. There the effective evolution is in I_j, with initial data

$$\tilde{\psi}_{j,l}(x_j) = e_1,$$

and terminal data

$$\tilde{\psi}_{j,l}(x_{j+1}) = T_{Q_j}^{-1}(z) e_1 + c e_2.$$

Now on the interval I_j we compare $\tilde{\psi}_l$ and $\tilde{\psi}_{j,l}$, which solve the same equation and have nearly collinear data. Their e_1 components must be nearly proportional, so it immediately follows that we must have the relation

$$T_{Q_j}^{-1}(z) = \frac{\tilde{\psi}_l^1(x_{j+1})}{\tilde{\psi}_l^1(x_j)} + O(e^{-cR}).$$

Multiplying these relations, it follows that for z on our curve γ we have

$$T_Q^{-1}(z) = \prod_{j=0}^{K} T_{Q_j}^{-1}(z) + O(e^{-cR}).$$

This implies that the product on the right must have the same number of zeroes as the left hand side, call them $\tilde{\mathbf{z}}$, and further that the zeros of the left hand side $\mathbf{z}$ and $\tilde{\mathbf{z}}$ must be close,

$$d(\mathbf{z}, \tilde{\mathbf{z}}) \lesssim e^{-cR},$$

for some new uniform constant c.

Applying the soliton removal map to Q_j within the contour γ we get

$$\mathbf{B}_-^{N_j} Q_j = (u_j, \tilde{\mathbf{z}}_j, \boldsymbol{\beta}_j),$$

where $\tilde{\mathbf{z}}_j$ are the poles of T_{Q_j}, which represent a subset of $\tilde{\mathbf{z}}$. By the trace formula (1.20) for Q_j we get

$$\|Q_j\|_{L^2}^2 = \|u_j\|_{L^2}^2 + 4 \operatorname{Im} \tilde{\mathbf{z}}_j,$$

while by the trace formula for Q,

$$\|Q\|_{L^2}^2 = 4 \operatorname{Im} \mathbf{z}.$$

Summing up in the first relation and comparing with the second, we obtain

$$\sum_{j=0}^{K} \|u_j\|_{L^2}^2 \lesssim e^{-2cR}.$$

A corollary of this is that each Q_j must have at least an eigenvalue within γ, or else it would have to have a very small L^2 norm. This implies that there can be at most N such Q_j, which completes the proof in part (a).

Finally, we define the multi-solitons

$$\tilde{Q}_j = \mathbf{B}_+^{\mathbf{N_j}}(\mathbf{0}, \tilde{\mathbf{z}}_\mathbf{j}, \boldsymbol{\beta}_\mathbf{j}) \tag{7.3}$$

By the uniform regularity of the soliton addition map, we have

$$\|Q_j - \tilde{Q}_j\|_{L^2} \lesssim e^{-cR},$$

so that

$$Q = \sum \tilde{Q}_j + O_{L^2}(e^{-cR}),$$

as desired. We note that the L^2 bound in the error can easily be upgraded to any higher Sobolev norm by interpolation. This concludes the proof of the theorem. □

An interesting question which emerges from the proof of the above theorem is whether one can lift the above correspondence to the level of the soliton manifolds. Above we have defined a map

$$\mathbf{M}^N \ni Q \to \Gamma(Q) := \{\tilde{Q}_j\} \in \prod \mathbf{M}^{N_j} \tag{7.4}$$

with the property that

$$\|Q - \sum \tilde{Q}_j\|_{H^s} \lesssim e^{-cR}.$$

One could also argue in reverse fashion, namely start with the solitons $\tilde{Q}_j$ and sum them,

$$v = \sum \tilde{Q}_j,$$

Then the same argument as in the proof of the theorem shows that v is a near soliton, in the sense that its residual energy is small. Precisely, if

$$\mathbf{B}_-^{\mathbf{N}}\mathbf{v} = (\mathbf{u}, \mathbf{z}, \boldsymbol{\beta}).$$

then we have

$$\|u\|_{H^s} \lesssim e^{-cR}, \qquad s > -\frac{1}{2}.$$

Hence by the mapping properties of the soliton addition, it follows that the map

$$\times \mathbf{M}^{N_j} \ni \{\tilde{Q}_j\} \to \tilde{\Gamma}(\{\tilde{Q}_j\}) = Q := \mathbf{B}_+^{\mathbf{N}}(\mathbf{0}, \mathbf{z}, \boldsymbol{\beta}) \in \mathbf{M}^{\mathbf{N}}, \tag{7.5}$$

is a near addition in the uniform norm,

$$\|Q - \sum \tilde{Q}_j\|_{H^s} \lesssim e^{-cR}.$$

If follows that the two manifolds $\mathbf{M}^N$ and $\sum \mathbf{M}^{N_j}$ are locally $O(e^{-cR})$ close. Here the product can be interpreted as a smooth manifold via the addition map, since the manifolds $\mathbf{M}^{N_j}$ are locally uniformly transversal; this is because the elements in their tangent space are exponentially localized near the $O(R)$ separated points x_j.

But the two manifolds are also locally uniformly smooth; it follows that they must also be close in any smooth topology:

Theorem 7.2. *Let Q be an N soliton with R separated bumps, and let $\tilde{Q}_j$ be as* (7.3)*. Then locally, near Q, respectively $\sum \tilde{Q}_j$, the manifolds $\mathbf{M}^N$ and $\sum \mathbf{M}^{N_j}$ are $O(e^{-cR})$ close as smooth manifolds.*

We note that this does not inply that either of the maps Γ, respectively $\tilde{\Gamma}$, are uniformly smooth near identity maps between the two manifolds. This would require a uniform regularity statement for the soliton removal map, which we wo not have. Nevertheless, we conjecture that such a result should be true.

8. The stability result

Here we prove the stability result using the regularity of the soliton addition map. We first restate the result in a more accurate form:

Theorem 8.1. *Let $s > -\frac{1}{2}$, and U a compact subset of the open upper half-plane. There exist $\varepsilon_0 > 0$ and $C > 0$ so that the following is true. Let v be a pure N-soliton solution for either NLS or mKdV with initial data $v_0 \in \mathbf{M}_U^N$. If the initial data w_0 for another solution w satisfies*

$$\|v_0 - w_0\|_{H^s} = \varepsilon \leq \varepsilon_0, \tag{8.1}$$

then there exists another pure N-soliton solution $\tilde{v}$ so that

$$\sup_{t \in \mathbb{R}} \|w(t) - \tilde{v}(t)\|_{H^s} \leq C\epsilon. \tag{8.2}$$

b) Furthermore, this result is uniform with respect to all N-soliton solutions with spectral parameters in a compact subset of the open upper half-plane.

Here we remark that the spectral data for $\tilde{v}$ are not necessarily in U, however they must be in a small neighbourhood of U. The uniformity assertion corresponds to the fact that constant C in (8.2) depends only on the compact set U.

Proof. We denote by $\mathbf{z}_0, \boldsymbol{\beta}_0$ the spectral, respectively scattering parameters for v_0. The transmission coefficient for v_0 is then given by

$$T_{v_0}(z) = \prod_{k=1}^{N} \frac{z - \bar{z}_{k0}}{z - z_{k0}},$$

and has poles at $\mathbf{z}_0$.

Away from the poles, the transmission coefficient depends smoothly on the input function. Hence, if ε_0 is small enough (depending only on U), it follows that the transmission coefficent of w_0 has exactly N poles $\mathbf{z}$ in a small neighbourhood of U, and that $\mathbf{z}$ is close to $\mathbf{z}_0$,

$$d(\mathbf{z}, \mathbf{z}_0) \lesssim \varepsilon$$

where the distance is measured using the symmetric polynomials.

We now apply the soliton removal map to w_0, denoting

$$\mathbf{B}_-^N(w_0) = (u_0, \mathbf{z}, \boldsymbol{\beta}),$$

and define the initial data

$$\tilde{v}_0 = \mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta}).$$

By the trace formula (5.2) in the trace theorem, the H^s energy of w_0 splits into

$$E_s(w_0) = E_s(u_0) + \sum_{k=1}^{N} \Xi_s(z_k) =: F_s(w_0) + \sum_{k=1}^{N} \Xi_s(z_k)$$

where $F_s(w_0)$ denotes the "no soliton energy" of w_0. This is uniformly smooth in the H^s topology, see [23] and also Theorem 1.1. It is also nonnegative and vanishes on the N-soliton manifold $\mathbf{M}_U^N$, so it vanishes of second order on the N-soliton manifold $\mathbf{M}_U^N$. It follows that

$$F_s(w_0) \lesssim \epsilon^2$$

which reinterpreted in terms of u_0 shows that

$$E_s(u_0) \lesssim \epsilon^2.$$

Since E_s is positive definite for small data, it follows that

$$\|u_0\|_{H^s} \lesssim \epsilon,$$

and by the uniform regularity of the soliton addition map,

$$\|w_0 - \tilde{v}_0\|_{H^s} \lesssim \epsilon.$$

The N-soliton solution $\tilde{v}$ with initial data $\tilde{v}_0$ to either NLS or mKdV is given by

$$\tilde{v}(t) = \mathbf{B}_+^N(0, \mathbf{z}, \boldsymbol{\beta}(t)),$$

where the parameter $\beta(t)$ depends on whether we consider the NLS or mKdV flow.

Denote by u the solution to NLS or mKdV with initial data u_0. Since E_s is conserved, this remains small,

$$\|u(t)\|_{H^s} \lesssim \varepsilon.$$

On the other hand the soliton addition map commutes with the flows, so we must have

$$w(t) = \mathbf{B}_+^N(u(t), \mathbf{z}, \beta(t)).$$

Using our result on the uniform regularity of the soliton addition map in Theorem 6.1, it follows that

$$\|w(t) - \tilde{v}(t)\|_{H^s} \lesssim \|u(t)\|_{H^s} \lesssim \varepsilon,$$

which concludes the proof of our theorem. □

9. Double eigenvalues

In this section w undertake a case study of double eigenvalues to gain some additional intuition and to provide some examples of multisoliton dynamics. For some early computations in this direction we refer the reader to [30].

9.1. **The asymptotic shift due to interaction.** We begin with the case of two different eigenvalues $z_1 \neq z_2$ and the z_j waves for the Lax operator with trivial potential,

$$\psi_1 = \begin{pmatrix} e^{\gamma_1 - iz_1 x} \\ e^{-\gamma_1 + iz_1 x} \end{pmatrix}$$

with $|\operatorname{Re}\gamma_1| \lesssim 1$. We assume that $\operatorname{Re}\gamma_2$ is large and choose

$$\psi_2 = \begin{pmatrix} 1 \\ e^{-2\gamma_2 + iz_2 x} \end{pmatrix}.$$

In this regime it is convenient to apply the iterated Bäcklund transform. The second intertwining operator (the one with respect to the index 2) is

$$D_2 = \begin{pmatrix} i\partial - \bar{z}_2 & 0 \\ 0 & -i\partial - \bar{z}_2 \end{pmatrix} - 2i \frac{\operatorname{Im} z_2}{1 + e^{-2\operatorname{Re}(\gamma_2 - iz_2 x)}} \begin{pmatrix} 1 & e^{-\overline{(\gamma_2 - iz_2 x)}} \\ e^{-(\gamma_2 - iz_2 x)} & e^{-2(\operatorname{Re}\gamma_2 + \operatorname{Im} z_2 x)} \end{pmatrix}$$

We apply the second intertwining operator to ψ_1,

(9.1)

$$D_2\psi_1 = \begin{pmatrix} \left[(z_1 - z_2) + 2i\operatorname{Im} z_2 \dfrac{e^{-2\operatorname{Re}\gamma_2 - 2\operatorname{Im} z_2 x}}{1 + e^{-2\operatorname{Re}\gamma_2 - 2\operatorname{Im} z_2 x}}\right] e^{\gamma_1 - iz_1 x} - 2i\operatorname{Im} z_2 \dfrac{e^{-\overline{\gamma_2 - iz_2 x} - \gamma_1 + iz_1 x}}{1 + e^{-2\operatorname{Re}\gamma_2 - 2\operatorname{Im} z_2 x}} \\ \left[(z_1 - \bar{z}_2) - 2i\operatorname{Im} z_2 \dfrac{e^{-2(\operatorname{Re}\gamma_2 + \operatorname{Im} z_2 x)}}{1 + e^{-2(\operatorname{Re}\gamma_2 + \operatorname{Im} z_2 x)}}\right] e^{-\gamma_1 + iz_1 x} - 2i\operatorname{Im} z_2 \dfrac{e^{-(\gamma_2 - iz_2 x)} e^{\gamma_1 - iz_1 x}}{1 + e^{-2(\operatorname{Re}\gamma_2 + \operatorname{Im} z_2 x)}} \end{pmatrix}$$

Without interaction the positions of the solitons would be the point x_j where both components of ψ_j have the same size,

$$x_j = -\frac{\operatorname{Re}\gamma_j}{\operatorname{Im} z_j}.$$

We assume without loss of generality $x_2 \le x_1$. We are interested in the case that the two soliton function will consist of two separated bumps. We define their position as the point where the amplitude has a local maximum, or, equivalently, where both components of $D_2\psi_1$ have the same size. The second soliton is far to the left of the first soliton if

$$\operatorname{Im} z_2 e^{-\operatorname{Re}\gamma_2 - \operatorname{Im} z_2 x_1} \ll |z_1 - z_2|.$$

Then

$$D_2\Psi(x_1) = \Big(1 + O\Big(\frac{\operatorname{Im} z_2}{|z_1 - z_2|} e^{-\operatorname{Re}\gamma_2 - \operatorname{Im} z_2 x_1}\Big)\Big) \begin{pmatrix} z_1 - z_2 \\ z_1 - \bar z_2 \end{pmatrix}$$

and due to the exponential factor $e^{\pm(\gamma_1 - i z_1 x)}$

$$y_1 = \frac{1}{2\operatorname{Im} z_1} \ln \frac{|z_1 - \bar z_2|}{|z_1 - z_2|} + O\Big(\frac{\operatorname{Im} z_2}{|z_1 - z_2|} e^{-\operatorname{Re}\gamma_2 - \operatorname{Im} z_2 x_1}\Big). \tag{9.2}$$

This gives the asymptotic shift due to the interaction when the solitons are well separated. it is not difficult to work out the shape of the solitons in this case.

9.2. **An algebraic computation.** In the sequel we seek for a more detailed understanding when the solitons are well separated with a separation independent of the distance between the eigenvalues, a much more involved task. We consider again two states ψ_1 and ψ_2 associated to eigenvalues z_1, z_2 and consider the corresponding matrix M

$$M = i \begin{pmatrix} \dfrac{\psi_1^*\psi_1}{\bar z_1 - z_1} & \dfrac{\psi_2^*\psi_1}{\bar z_2 - z_1} \\ \dfrac{\psi_1^*\psi_2}{\bar z_1 - z_2} & \dfrac{\psi_2^*\psi_2}{\bar z_2 - z_2} \end{pmatrix}.$$

Let m be the inverse of M. We evaluate the expression

$$w = 2\bar\psi_j^2 m_{jk} \psi_k^1, \tag{9.3}$$

which arises in the definition of the Bäcklund transform in (4.5). Our first task is to compute the determinant of M,

$$\begin{aligned}\det M &= \prod_{i,j=1}^{2} \frac{-1}{\bar z_i - z_j} (|z_1 - \bar z_2|^2 |\psi_1|^2 |\psi_2|^2 - 4 \operatorname{Im} z_1 \operatorname{Im} z_2 |\psi_1^*\psi_2|^2) \\ &= \prod_{i,j=1}^{2} \frac{-1}{\bar z_i - z_j} \Big(|z_1 - z_2|^2 |\psi_1|^2 |\psi_2|^2 + 4 \operatorname{Im} z_1 \operatorname{Im} z_2 (|\psi_1|^2|\psi_2|^2 - |\psi_1^*\psi_2|^2\Big),\end{aligned}$$

where we have used

$$|z_1 - \bar z_2|^2 = |z_1 - z_2|^2 + 4 \operatorname{Im} z_1 \operatorname{Im} z_2. \tag{9.4}$$

We recall that we can choose

$$\psi_j = \begin{pmatrix} e^{\gamma_j} \\ e^{-\gamma_j} \end{pmatrix}, \qquad \gamma_j = -i \sum_{k=0}^{3} \beta_k z_j^k - i z_j x.$$

Then we can rewrite the expression

$$D = -\prod_{i,j=1}^{2} (\bar{z}_i - z_j) \det M$$

as

$$\begin{aligned} D &= 4|z_1 - z_2|^2 \cosh(2\operatorname{Re}\gamma_1)\cosh(2\operatorname{Re}\gamma_2) + 16\operatorname{Im} z_1 \operatorname{Im} z_2 \,|\sinh(\gamma_1-\gamma_2)|^2 \\ &= 2|z_1-z_2|^2[|\cosh(\gamma_1+\gamma_2)|^2 + |\sinh(\gamma_1+\gamma_2)|^2 + |\cosh(\gamma_1-\gamma_2)|^2] \\ &\quad + 2[|z_1+z_2|^2 - 2(z_1 z_2 + \bar{z}_1\bar{z}_2)]\,|\sinh(\gamma_1-\gamma_2)|^2 . \end{aligned} \tag{9.5}$$

We can read off important parts of the structure. At the right hand side of the first equality we see a sum of two terms, the first containing a factor $|z_1 - z_2|^2$, and the second a factor $|\gamma_1 - \gamma_2|^2$. This vanishes quadratically exactly when $z_1 = z_2$ and $\gamma_1 = \gamma_2$ modulo $i\pi$, or, equivalently, if ψ_1 and ψ_2 are collinear.

On the right hand side of the second equality we consider z_j, γ_j and $\bar{z}_j$ resp. $\bar{\gamma}_j$ as separated variables and see that exchanging (z_1, γ_1) and (z_2, γ_2) while keeping the complex conjugates changes the sign.

In the complement of the set where $\det M$ vanishes we can write

$$w = -\frac{2A}{D}$$

where A is given by 4 times

$$(e^{-\bar{\gamma}_1}, e^{-\bar{\gamma}_2})\begin{pmatrix} \frac{1}{2}|z_2-\bar{z}_1|^2 \operatorname{Im} z_1(e^{\gamma_2+\bar{\gamma}_2}+e^{-\gamma_2-\bar{\gamma}_2}) & -i\operatorname{Im} z_1 \operatorname{Im} z_2(\bar{z}_1-z_2)(e^{\gamma_1+\bar{\gamma}_2}+e^{-\gamma_1-\bar{\gamma}_2}) \\ -i\operatorname{Im} z_1 \operatorname{Im} z_2(\bar{z}_2-z_1)(e^{\gamma_2+\bar{\gamma}_1}+e^{-\gamma_2-\bar{\gamma}_1}) & \frac{1}{2}|z_2-\bar{z}_1|^2 \operatorname{Im} z_2(e^{\gamma_1+\bar{\gamma}_1}+e^{-\gamma_1-\bar{\gamma}_1}) \end{pmatrix}\begin{pmatrix} e^{\gamma_1} \\ e^{\gamma_2} \end{pmatrix}.$$

We rewrite A as follows:

$$\begin{aligned} A &= 4|z_2-\bar{z}_1|^2 \{\operatorname{Im} z_1(\exp(2i\operatorname{Im}\gamma_1)\cosh(\gamma_2+\bar{\gamma}_2) + \operatorname{Im} z_2(\exp(2i\operatorname{Im} z_2)\cosh(\gamma_1+\bar{\gamma}_1)\} \\ &\quad - 4i\operatorname{Im} z_1 \operatorname{Im} z_2\Big\{(\bar{z}_1-z_2)(\exp(\gamma_1-\bar{\gamma}_1+\gamma_2+\bar{\gamma}_2) + \exp(-\gamma_1-\bar{\gamma}_1+\gamma_2-\bar{\gamma}_2)) \\ &\qquad\qquad + (\bar{z}_2-z_1)(\exp(\gamma_1+\bar{\gamma}_1+\gamma_2-\bar{\gamma}_2) + \exp(\gamma_1-\bar{\gamma}_1-\gamma_2-\bar{\gamma}_2)\Big\} \\ &= 4|z_1-z_2|^2(\operatorname{Im} z_1 e^{2i\operatorname{Im}\gamma_1}\cosh(2\operatorname{Re}\gamma_2) + \operatorname{Im} z_2 e^{2i\operatorname{Im}\gamma_2}\cosh(2\operatorname{Re}\gamma_1)) \\ &\quad + 8i\operatorname{Im} z_1 \operatorname{Im} z_2\Big\{(z_1-z_2)e^{\gamma_1+\gamma_2}\sinh(\bar{\gamma}_1-\bar{\gamma}_2) + (\bar{z}_1-\bar{z}_2)e^{-\bar{\gamma}_1-\bar{\gamma}_2}\sinh(\gamma_1-\gamma_2)\Big\} \\ &= |z_1-z_2|^2\Big(2\operatorname{Im}(z_1+z_2)(e^{2i\operatorname{Im}\gamma_1}\cosh(2\operatorname{Re}\gamma_2) + e^{2i\operatorname{Im}\gamma_2}\cosh(2\operatorname{Re}\gamma_1)) \\ &\qquad + ((z_1-z_2)+(\bar{z}_1-\bar{z}_2))(e^{\gamma_1-\bar{\gamma}_1}\cosh(\gamma_2+\bar{\gamma}_2) - e^{\gamma_2-\bar{\gamma}_2}\cosh(\gamma_1+\bar{\gamma}_1))\Big) \\ &\quad + 8i\operatorname{Im} z_1 \operatorname{Im} z_2\Big\{(z_1-z_2)e^{\gamma_1+\gamma_2}\sinh(\bar{\gamma}_1-\bar{\gamma}_2) + (\bar{z}_1-\bar{z}_2)e^{-\bar{\gamma}_1-\bar{\gamma}_2}\sinh(\gamma_1-\gamma_2)\Big\} \\ &= |z_1-z_2|^2\Big\{2\operatorname{Im}(z_1+z_2)\left(e^{\gamma_1+\gamma_2}\cosh(\bar{\gamma}_1-\bar{\gamma}_2) + e^{-\bar{\gamma}_1-\bar{\gamma}_2}\cosh(\gamma_1-\gamma_2)\right) \\ &\qquad - i(z_1-z_2)e^{-\bar{\gamma}_1-\bar{\gamma}_2}\sinh(\gamma_1-\gamma_2) - i(\bar{z}_1-\bar{z}_2)e^{\gamma_1+\gamma_2}\sinh(\bar{\gamma}_1-\bar{\gamma}_2)\Big\} \\ &\quad + i[|z_1+z_2|^2 - 2(z_1z_2+\bar{z}_1\bar{z}_2)] \\ &\qquad \times\Big\{(z_1-z_2)e^{\gamma_1+\gamma_2}\sinh(\bar{\gamma}_1-\bar{\gamma}_2) + (\bar{z}_1-\bar{z}_2)e^{-\bar{\gamma}_1-\bar{\gamma}_2}\sinh(\gamma_1-\gamma_2)\Big\} \end{aligned} \tag{9.6}$$

Both A and D are smooth. It is an easy consequence that $w = -2AD^{-1}$ is smooth in the set $\{z_1 \neq z_2\} \cap \{\gamma_1 - \gamma_2 \notin i\pi\mathbb{Z}\}$ and that it vanishes if $\gamma_1 - \gamma_2 \notin i\pi\mathbb{Z}$ but $z_1 = z_2$.

In order to resolve the apparent singularity at the zeroes of the denominator, we view z_j and $\bar{z}_j$ as separate variables, and similarly for γ_j and $\bar{\gamma}_j$. We first observe that both A and D are odd with respect to the separate symmetries

$$(z_1, \gamma_1) \leftrightarrow (z_2, \gamma_2),$$

respectively

$$(\bar{z}_1, \bar{\gamma}_1) \leftrightarrow (\bar{z}_2, \bar{\gamma}_2).$$

Then their ratio is invariant under both separate exchanges. To capture the cancellation allowed by this symmetry we introduce the auxiliary variables γ, α

$$2\gamma = \gamma_1 + \gamma_2 \qquad \alpha = \frac{\sinh(\gamma_1 - \gamma_2)}{z_1 - z_2}, \tag{9.7}$$

and cancel a $|z_1 - z_2|^2$ factor. We obtain

Lemma 9.1. *With the notations in* (9.7)*, the expression w in* (9.3) *can be represented in the nondegenerate form*

$$w = -\frac{2A_0}{D_0}$$

where

$$\begin{aligned} A_0 = {} & 2\,\mathrm{Im}(z_1 + z_2)\left(e^{2\gamma}\cosh(\bar{\gamma}_1 - \bar{\gamma}_2) + e^{-2\bar{\gamma}}\cosh(\gamma_1 - \gamma_2)\right) \\ & - i\alpha(z_1 - z_2)^2 e^{-2\bar{\gamma}} - i\bar{\alpha}(\bar{z}_1 - \bar{z}_2)^2 e^{2\gamma} + i[|z_1 + z_2|^2 - 2(z_1 z_2 + \bar{z}_1\bar{z}_2)]\left\{\bar{\alpha}e^{2\gamma} + \alpha e^{-2\bar{\gamma}}\right\} \end{aligned} \tag{9.8}$$

and

$$D_0 = 2\Big[|\cosh(2\gamma)|^2 + |\sinh(2\gamma)|^2 + |\cosh(\gamma_1 - \gamma_2)|^2\Big] + 2\Big[|z_1 + z_2|^2 - 2(z_1 z_2 + \bar{z}_1\bar{z}_2)\Big]|\alpha|^2 \tag{9.9}$$

Assuming that z_1, z_2 are confined to a (small) compact subset of the upper half-plane, $s_1 = z_1 + z_2$, $s_2 = z_1^2 + z_2^2$ we interpret this expression as a zero homogeneous form

$$w = w(\mu, s_1, s_2)$$

in the complex variables

$$\mu = (\mu_1, \mu_2, \mu_3, \mu_4) = (\cosh(2\gamma), \sinh(2\gamma), \cosh(\gamma_1 - \gamma_2), \alpha),$$

with smooth coefficients which are symmetric functions separately in (z_1, z_2) and $(\bar{z}_1, \bar{z}_2)$, resp. smooth coefficients in s_1 and s_2.

For the function w we note the pointwise bound:

$$|w| \lesssim \frac{(|\mu_1| + |\mu_2|)(|\mu_3| + |\mu_4|)}{|\mu|^2} := w_0, \tag{9.10}$$

which in particular shows that for unbalanced μ's w must be small:

$$|w| \approx 1 \quad \Longrightarrow \quad |\mu_1| + |\mu_2| \approx |\mu_3| + |\mu_4|.$$

We also have similar bounds for the derivatives of w with respect to μ,

$$|\mu|^{|\alpha|}\left|\partial_\mu^\alpha w\right| \lesssim w_0. \tag{9.11}$$

One might be tempted to parametrize w as a function of z_1, z_2, γ and α, but $\cosh(\gamma_1 - \gamma_2)$ can only be viewed locally as a smooth function of α for $\gamma_1 - \gamma_2$ away from $(\frac{1}{2} + \mathbb{Z})\pi i$. Thus it is better to think of these variables, together with $z_1 + z_2$ and $z_1 z_2$ as functions on a smooth complex manifold M of complex dimension 4, which is the is cartesian product of the smooth Riemann surface

$$\{(\mu_1, \mu_2) : \mu_1^2 - \mu_2^2 = 1\}$$

and the three dimensional complex manifold (recall that $(z_1-z_2)^2 = 2s_2 - s_1^2$)

$$\{(\mu_3,\mu_4,s_1,s_2) \in \mathbb{C}^4 : \ \mu_3^2 - (2s_2 - s_1^2)\mu_4^2 = 1\},$$

which is smooth since $\mu_3^2 - (2s_2-s_1^2)\mu_4^2 - 1$ is nondegenerate in a neighborhood of the manifold.

We remark that on M we have the relations

$$|\mu| \geq 1, \quad ||\mu_1| - |\mu_2|| \leq 1, \qquad \Big| |\mu_3| - |z_1 - z_2||\mu_4| \Big| \leq 1$$

Then we can bound

$$|w_0| \leq \frac{(1+2|\mu_1|)(1+2|\mu_4|)}{|\mu|^2}$$

if $|z_1 - z_2| \leq 1$, which can only be large if $|\mu_1| + 1 \approx |\mu_4| + 1$.

The function w above will describe the pointwise size of a soliton. Because of that, the next question we want to address is where is w large. Heuristically we expect to have two regions of interest

(i) **The one bump case** $|\mu_1|, |\alpha| \lesssim 1$ where the amplitude of w could get as high as $2\operatorname{Im}(z_1 + z_2) = 4\operatorname{Im} z$,with $2z = z_1 + z_2$. This value is attained when $\gamma_1 = \gamma_2 = \alpha = 0$.

(ii) **Separated bumps** $|\mu_1| \approx |\alpha| \gg 1$, where we have amplitudes closer to $2\operatorname{Im} z$ if $z_1 - z_2$ is small.

We are particularly interested in understanding this in the (near) degenerate case, when z_1 and z_2 are close but α is large and approximatively balances the $\cosh(4\operatorname{Re}\gamma)$ in the denominator, so that w has size $O(1)$. Toward that goal, we denote

$$\sigma = (z_1 - z_2)\coth(\gamma_1 - \gamma_2) \tag{9.12}$$

which is bounded when α is large, and has limit $\pm(z_1 - z_2)$ as $\operatorname{Re}(\gamma_1 - \gamma_2)$ goes to $\pm\infty$. Then we can rewrite D_0 as

$$D_0 = 2\cosh(4\operatorname{Re}\gamma) + 2[|z_1 + z_2|^2 - 2(z_1 z_2 + \bar{z}_1\bar{z}_2) + \sigma^2]|\alpha|^2.$$

Since

$$|\sigma|^2 = |z_1 - z_2|^2 + O(|\alpha|^{-2}),$$

it follows that

$$D_0 = 2\cosh(4\operatorname{Re}\gamma) + 2|z_1 - \bar{z}_2|^2|\alpha|^2 + O(1).$$

On the other hand we can rewrite the expression A_0 as

$$A_0 = c_+\bar{\alpha}e^{2\gamma} + c_-\alpha e^{-2\gamma}, \tag{9.13}$$

where c_+ and c_- are bounded,

$$c_+ = i[|z_1+z_2|^2 - 2(z_1z_2 + \bar{z}_1\bar{z}_2) - (\bar{z}_1 - \bar{z}_2)^2] + 2\operatorname{Im}(z_1+z_2)\bar{\sigma}$$
$$c_- = i[|z_1+z_2|^2 - 2(z_1z_2 + \bar{z}_1\bar{z}_2) - (z_1 - z_2)^2] + 2\operatorname{Im}(z_1+z_2)\sigma$$

Thus we get

$$w = -2e^{2i\operatorname{Im}\gamma}\frac{c_+\bar{\alpha}e^{2\operatorname{Re}\gamma} + c_-\alpha e^{-2\operatorname{Re}\gamma}}{2\cosh(4\operatorname{Re}\gamma) + 2|z_1 - \bar{z}_2|^2|\alpha|^2} + O(\frac{1}{\cosh(4\operatorname{Re}\gamma) + |\alpha|^2}). \tag{9.14}$$

With a slightly larger error we can further simplify this as

$$w = -\frac{8i(\operatorname{Im} z)^2(\bar{\alpha}e^{2\gamma} + \alpha e^{-2\bar{\gamma}})}{\cosh 4\operatorname{Re}\gamma + 8(\operatorname{Im} z)^2|\alpha|^2} + O(|z_1 - z_2|) + O(\alpha^{-1}). \tag{9.15}$$

This has near maximum amplitude $2\operatorname{Im} z$ when

$$\cosh(2\operatorname{Re}\gamma) = 2\operatorname{Im} z|\alpha|, \tag{9.16}$$

and phase

$$\pi/2 \mp (\arg\alpha - 2\operatorname{Im}\gamma),$$

where the sign depends on the sign of $\operatorname{Re}\gamma$.

It is also interesting to check the asymptotic behavior of (9.14) as $\operatorname{Re}(\gamma_1-\gamma_2)\to\infty$. There we can approximate

$$\cosh(\gamma_1-\gamma_2) \approx \sinh(\gamma_1-\gamma_2)\operatorname{sgn}\operatorname{Re}(\gamma_1-\gamma_2).$$

This yields

$$w = \frac{-\bar{c}_-\bar{\alpha}e^{2\gamma} + c_+\alpha e^{-2\bar{\gamma}}}{\cosh 4\operatorname{Re}\gamma + 8(\operatorname{Im} z)^2|\alpha|^2} + O(|z_1-z_2|^2) + O(\alpha^{-2}),$$

where

$$c_\pm = (8i(\operatorname{Im} z)^2 \pm 4\operatorname{Im} z(z_1-z_2)\operatorname{sgn}\operatorname{Re}(\gamma_1-\gamma_2)).$$

This gives factors of $\operatorname{Im} z\operatorname{Im} z_1$ respectively $\operatorname{Im} z\operatorname{Im} z_2$ at the numerator, which will select the different bump amplitudes $2\operatorname{Im} z_1$, respectively $2\operatorname{Im} z_2$. Compared with the prior computation we see the transition from the amplitude $2\operatorname{Im} z$ for one bump solitons to the amplitude $2\operatorname{Im} z_j$ as the distance tends to infinity.

9.3. **Two soliton states.** Separated two solitons are close to the sum of two 1-solitons. We study the general pure two soliton solution and estimate the difference to the algebraic sum of two solitons, whenever the two centers are far apart. This analysis is new, nontrivial and interesting in the case of two close eigenvalues. Asymptotically the eigenvalue parameters of the two solitons are the poles of the transmission coefficients. But as soon as their distance is closer than $\ln(2+\frac{\operatorname{Im}(z_1+z_2)}{|z_1-z_2|})$, the interaction is visible and we can see a transition regime via effective soliton parameters, which we describe. As a consequence we obtain a uniform parametrization of the two soliton manifold across multiplicities.

Following the pattern in the previous sections, we begin by considering γ_1, γ_2 of the form

$$\gamma_j = i(\beta_0 + \beta_1 z_j + \beta_2 z_j^2 + \beta_3 z_j^3) \tag{9.17}$$

with real coefficients β_k. Then in terms of the elementary sysmmetric polynomials s_1 and s_2

$$2\gamma = i(\beta_0 + \beta_1(z_1+z_2) + \beta_2(z_1^2+z_2^2) + \beta_3(z_1^3+z_2^3)) = i(\beta_0 + \beta_1 s_1 + \beta_2 s_2 + \beta_3 s_1(\frac{3}{2}s_2 - \frac{1}{2}s_1^2)$$

$$\gamma_1-\gamma_2 = i(\beta_1(z_1-z_2) + \beta_2(z_1^2-z_2^2) + \beta_3(z_1^3-z_2^3)) = i(z_1-z_2)(\beta_1 + \beta_2 s_1 + \beta_3(\frac{1}{2}s_2 + \frac{1}{2}s_1^2))$$

$$\mu_1 = \cosh(2\gamma), \quad \mu_2 = \sinh(2\gamma), \quad \mu_3 = \cosh(\gamma_1-\gamma_2), \quad \alpha = \frac{\sinh(\gamma_1-\gamma_2)}{z_1-z_2}.$$

Then we can view the above w as

$$w = w(\mathbf{s},\boldsymbol{\beta}), \qquad \mathbf{s} = (z_1+z_2, z_1^2+z_2^2), \quad \boldsymbol{\beta} = (\beta_0,\beta_1,\beta_2,\beta_3),$$

where for $\mathbf{z}$ we use the topology defined by the symmetric polynomials. For this function we have

Lemma 9.2. *For z_1, z_2 in a compact subset of the upper half-space, the function w is a uniformly smooth function of $(\mathbf{s},\boldsymbol{\beta})$. Furthermore, we have the uniform bound*

$$|\partial_{\mathbf{s}}^a \partial_{\boldsymbol{\beta}}^b w| \lesssim |w_0|, \tag{9.18}$$

where

$$w_0 = \frac{(1+|\alpha|)\cosh(2\operatorname{Re}\gamma)}{\cosh(4\operatorname{Re}\gamma) + |\alpha|^2}.$$

Proof. The proof is straightforward. On one hand we know that w_0 is of the same size as the one defined in (9.10). The bounds (9.10) and (9.11) and

$$|\partial^a_{\beta_j,s_1,s_2}\mu_j| \lesssim |\mu|$$

imply the uniform bounds. These bounds for μ_1 and μ_2 are obvious. Both $\cosh(\gamma_1 - \gamma_2)$ and $\alpha = \frac{\sin(\gamma_1-\gamma_2)}{z_1-z_2}$ are even and analytic as functions of z_2 and z_1, and hence they are holomorphic functions of s_1 and s_2. The bounds on derivatives then follow by Cauchy's integral formula on balls around s_1 and s_2. □

We now describe two soliton states. Relative to the $(\mathbf{z}, \boldsymbol{\beta})$ parametrization with $\beta_j \in \mathbb{R}$, this corresponds to choosing

$$Q_{\mathbf{z},\boldsymbol{\beta}}(x) = w(\mathbf{z}, \tilde{\boldsymbol{\beta}}),$$

where

$$\tilde{\boldsymbol{\beta}} = (\beta_0, \beta_1 + x, \beta_2, \beta_3),$$

and the corresponding γ_j's are

$$\gamma_j = i(\beta_0 + (\beta_1 + x)z_j + \beta_2 z_j^2 + \beta_3 z_j^3).$$

To compare with our general set-up, the associated scattering parameters κ_j are

$$\kappa_j = i(\beta_0 + \beta_1 z_j + \beta_2 z_j^2 + \beta_3 z_j^3).$$

In particular we have

$$2\gamma = i(2\beta_0 + (\beta_1 + x)(z_1 + z_2) + \beta_2(z_1^2 + z_2^2) + \beta_3(z_1^3 + z_2^3)) \tag{9.19}$$

and

$$\gamma_0 := \frac{\gamma_1 - \gamma_2}{z_1 - z_2} = i(\beta_1 + x + \beta_2(z_1 + z_2) + \beta_3(z_1^2 + z_1 z_2 + z_2^2)). \tag{9.20}$$

Next we consider the location of the two bumps for the 2-solitons. We begin with the location of the single bumps for the corresponding 1-solitons with the same spectral and scattering parameters, whose centers are given by x_1, x_2 determined by

$$\operatorname{Re}\gamma_j = 0 \iff x_j = -\frac{\operatorname{Im}\kappa_j}{\operatorname{Im} z_j} = -\frac{\operatorname{Im}(\beta_1 z_j + \beta_2 z_j^2 + \beta_3 z_j^3)}{\operatorname{Im} z_j}. \tag{9.21}$$

For later considerations we denote their phase at the center of soliton by

$$\theta_j = \beta_0 + (\beta_1 + x_j)\operatorname{Re} z_j + \beta_2 \operatorname{Re} z_j^2 + \beta_3 \operatorname{Re} z_j^3. \tag{9.22}$$

Then

$$\gamma_j = i(\theta_j + z_j(x - x_j))$$

and

$$\gamma_0 = i\frac{\theta_1 - \theta_2 + z_1(x - x_1) - z_2(x - x_2)}{z_1 - z_2}.$$

Recall that

$$\alpha = \frac{\sinh((z_1 - z_2)\gamma_0)}{z_1 - z_2}, \qquad \sigma = (z_1 - z_2)\coth((z_1 - z_2)\gamma_0).$$

The two bumps are centered (recall (9.16)) for $|z_1 - z_2| \ll \operatorname{Im} z_1)$ where

$$\cosh(2\operatorname{Re}\gamma) \approx |z_1 - \bar{z}_2||\alpha|. \tag{9.23}$$

Here the expression $\operatorname{Re}\gamma$ is linear in x, decreasing at a uniform rate, and vanishing at a point x_0, which is related to x_1 and x_2 by the relation

$$x_0 = -\beta_1 - \frac{\beta_2 \operatorname{Im}(z_1^2 + z_2^2) + \beta_3 \operatorname{Im}(z_1^3 + z_2^3)}{\operatorname{Im} z_1 + \operatorname{Im} z_2} = \frac{\operatorname{Im} z_1}{\operatorname{Im}(z_1 + z_2)}x_1 + \frac{\operatorname{Im} z_2}{\operatorname{Im}(z_1 + z_2)}x_2 \tag{9.24}$$

which can be seen as the center of mass of the 2-soliton state.

We also define an averaged phase at the center by

$$2\theta = \theta_1 + \theta_2 + \frac{\operatorname{Im} z_1 \operatorname{Re} z_2 - \operatorname{Im} z_2 \operatorname{Re} z_1}{\operatorname{Im}(z_1 + z_2)}(x_1 - x_2), \tag{9.25}$$

in order to have

$$2\gamma = i(2\theta + (z_1 + z_2)(x - x_0)). \tag{9.26}$$

Moreover, $\cosh(2\operatorname{Re}\gamma)$ grows at uniform exponential rates away from x_0. On the other hand, α has a smaller logarithmic derivative,

$$\frac{\partial \alpha}{\partial x} = i\alpha\sigma = O(|\alpha||z_1 - z_2| + 1)$$

As a consequence, if $|\alpha(x_0)|$ is large then there are exactly two unit size regions where (9.23) (where we consider both sides as functions of x) is satisfied. Furthermore, in this region the coefficients c_+ and c_- are slowly varying, as

$$\frac{\partial \sigma}{\partial x} = i\alpha^{-2}$$

This implies that, with $O(|\alpha^{-2}|)$ accuracy, the maximum points of w are described by the relation

$$\cosh(2\operatorname{Re}\gamma) = |z_1 - \bar{z}_2||\alpha|. \tag{9.27}$$

To accurately calculate the roots of (9.27) it is useful to consider the value of α at the center x_0. This is determined by

$$(z_1 - z_2)\gamma_{00} := (z_1 - z_2)\gamma_0(x_0) = i\left(\theta_1 - \theta_2 + \frac{z_1 \operatorname{Im} z_2 + z_2 \operatorname{Im} z_1}{2\operatorname{Im} z}(x_2 - x_1)\right), \tag{9.28}$$

and in particular

$$\operatorname{Re}((z_1 - z_2)\gamma_{00}) = \frac{\operatorname{Im} z_1 \operatorname{Im} z_2}{\operatorname{Im} z}(x_2 - x_1). \tag{9.29}$$

Thus we define

$$\alpha_0 = \frac{\sinh((z_1 - z_2)\gamma_{00})}{z_1 - z_2}, \qquad \sigma_0 = (z_1 - z_2)\coth((z_1 - z_2)\gamma_{00}), \tag{9.30}$$

where γ_{00} is linear in β_2 and β_3,

$$\gamma_{00} = \gamma_0(x_0) =: a_2\beta_2 + a_3\beta_3$$

where the coefficients a_2, a_3 are symmetric functions in z_1, z_2, given by (denoting $z_1 + z_2 = 2z$)

$$a_2 = i(z_1 + z_2) - i\frac{\operatorname{Im}(z_1^2 + z_2^2)}{\operatorname{Im}(z_1 + z_2)} = -2\operatorname{Im} z - i\frac{\operatorname{Re}(z_2 - z_1)\operatorname{Im}(z_2 - z_1)}{2\operatorname{Im} z} \tag{9.31}$$

$$\begin{aligned} a_3 &= i(z_1^2 + z_1 z_2 + z_2^2) - i\frac{\operatorname{Im}(z_1^3 + z_2^3)}{\operatorname{Im}(z_1 + z_2)} \\ &= -\frac{3}{2}\operatorname{Re}(z_1 + z_2)\operatorname{Im}(z_1 + z_2) - \frac{i}{2}[\operatorname{Im}(z_1 + z_2)]^2 + \frac{i}{4}(z_1 - z_2)^2 - \frac{3i}{4}\frac{\operatorname{Im}[(z_1 + z_2)(z_1 - z_2)^2]}{\operatorname{Im}(z_1 + z_2)} \end{aligned} \tag{9.32}$$

These can be checked to be linearly independent over $\mathbb{R}$, for instance by verifying that the following expression is nonzero (we write $z_j = x_j + iy_j$)

$$\begin{aligned}
J =& \operatorname{Im}(z_1+z_2)\operatorname{Im}(a_2\bar{a}_3)\\
=& \operatorname{Im}(z_1+z_2)\Big[\operatorname{Im}(z_1+z_2)\operatorname{Re}(z_1^2+z_1z_2+z_2^2)-\operatorname{Im}(z_1^3+z_2^3)\Big]\\
&-\Big[\operatorname{Re}(z_1+z_2)\operatorname{Im}(z_1+z_2)-\operatorname{Im}(z_1^2+z_2^2)\Big]\operatorname{Im}(z_1^2+z_1z_2+z_2^2)\\
=& (y_1+y_2)^2(-\frac{1}{2}(x_1-x_2)^2-2y_1y_2)-\frac{3}{2}(y_1+y_2)(x_1-x_2)(y_1-y_2)(x_1+x_2)\\
&+(x_1-x_2)(y_1-y_2)\Big[\frac{3}{2}(y_1+y_2)(x_1+x_2)+\frac{1}{2}(y_1-y_2)(x_1-x_2)\Big]\\
=& -2y_1y_2((x_1-x_2)^2+(y_1+y_2)^2)\\
=& -2\operatorname{Im}z_1\operatorname{Im}z_2|z_1-\bar{z}_2|^2.
\end{aligned}$$

This implies that

$$|\gamma_{00}| \approx |\beta_2| + |\beta_3|.$$

In particular we will be interested in $\operatorname{Re}((z_1-z_2)\gamma_{00})$ (see (9.28)) , which can be alternatively expressed in the form

$$\operatorname{Re}((z_1-z_2)\gamma_{00}) = \frac{\operatorname{Im}z_1\operatorname{Im}z_2}{\operatorname{Im}z}(x_1-x_2).$$

On the other hand for the imaginary part we get

$$\operatorname{Im}((z_1-z_2)\gamma_{00}) = \theta_1-\theta_2+\frac{\operatorname{Im}(z_1z_2)}{2\operatorname{Im}z}(x_2-x_1).$$

As above we distinguish two scenarios, still assuming $|z_1-z_2| \ll 1$:

(i) **Single bump case.** This becomes

$$\text{dist }(\gamma_{00}, i\pi(z_1-z_2)^{-1}\mathbb{Z}) \lesssim 1$$

and corresponds to 2-solitons $Q_{\mathbf{z},\boldsymbol{\beta}}$ which have two overlapping solitons. The amplitude of w could get as high as $2\operatorname{Im}(z_1+z_2)$, which value is attained at $x = x_0$ when $\beta_2 = \beta_3 = 0$. In this case we have exponential decay away from x_0 and the two soliton is close to a 2-soliton with $z_1 = z_2$ since the 2-soliton depends smoothly on s_1 and s_2.

(ii) **Two bumps case.**

$$\text{dist }(\gamma_{00}, i\pi(z_1-z_2)^{-1}\mathbb{Z}) \gg 1,$$

which corresponds to 2-solitons $Q_{\mathbf{z},\boldsymbol{\beta}}$ which have two simple bumps, with amplitudes closer to the range between $2\operatorname{Im}z_1$ and $2\operatorname{Im}z_2$.

In the second case above, we seek a more accurate description of the location of the two bumps, which are given by the relation (9.27), which translates to

$$\cosh(\operatorname{Im}(z_1+z_2)(x-x_0)) = |z_1-\bar{z}_2|\left|\frac{\sinh((z_1-z_2)(\gamma_{00}+i(x-x_0)))}{z_1-z_2}\right|.$$

Based on the discussion above, this equation will have two roots, one above and one below x_0.

In a first approximation we evaluate the size of $|x-x_0|$ for the two roots by

$$|x-x_0| \approx \log\left|\frac{\sinh((z_1-z_2)\gamma_{00})}{z_1-z_2}\right|.$$

In this region we take a Taylor expansion of $\ln\alpha$,

$$\ln\alpha = \ln\alpha_0 + i(x-x_0)\sigma_0 + O(|x-x_0|^2|\alpha_0|^{-2}).$$

At the roots x this leads to

$$2\operatorname{Im} z|x-x_0| = \ln|z_1-\bar{z}_2| + \ln(2|\alpha_0|) - (x-x_0)\operatorname{Im}\sigma_0 + O(\frac{\ln^2|\alpha_0|}{|\alpha_0|^2}),$$

and finally to

$$x - x_0 = \pm\frac{\ln|z_1-\bar{z}_2| + \ln(2|\alpha_0|)}{2\operatorname{Im} z \pm \operatorname{Im}\sigma_0} + O(\frac{\ln^2|\alpha_0|}{|\alpha_0|^2}).$$

This gives the approximate locations of the centers for the two bumps as

$$x^{\pm} = x_0 \pm \frac{\ln|z_1-\bar{z}_2| + \ln(2|\alpha_0|)}{2\operatorname{Im} z \pm \operatorname{Im}\sigma_0}, \tag{9.33}$$

with accuracy $O(\epsilon)$ where

$$\epsilon = \frac{\ln^2|\alpha_0|}{|\alpha_0|^2}. \tag{9.34}$$

We remark that when $\operatorname{Re}((z_1-z_2)\gamma_{00}) \gg 1$ (which corresponds to $x_1 - x_2 \gg 1$) we can approximate

$$\sigma_0 \approx z_1 - z_2, \qquad \ln(2\alpha_0) \approx (z_1-z_2)\gamma_{00} - \ln(z_1-z_2)$$

at the expense of allowing larger errors of size

$$e^{-2\operatorname{Re}(z_1-z_2)\gamma_{00}} \approx (|z_1-z_2||\alpha_0|)^{-2}.$$

This yields with $z_+ = z_1 \approx \frac{1}{2}(z+\sigma_0)$, $z_- = z_2 \approx \frac{1}{2}(z-\sigma_0)$

$$x - x_0 = \pm\frac{\ln|z_1-\bar{z}_2| + \ln 2(|\alpha_0|)}{2\operatorname{Im} z_\pm} + O\Big(\frac{\ln^2|\alpha_0| + |z_1-z_2|^{-2}}{|\alpha_0|^2}\Big).$$

Then the larger solution can be associated to z_1,

$$\hat{x}_1 \approx x_0 + \frac{\ln|z_1-\bar{z}_2| - \ln|z_1-z_2| + \operatorname{Re}[(z_1-z_2)\gamma_{00}]}{2\operatorname{Im} z_2} = x_2 + \frac{\ln|z_1-\bar{z}_2| - \ln|z_1-z_2|}{2\operatorname{Im} z_2},$$

and the smaller one can be associated to z_1, using (9.24) and (9.29),

$$\hat{x}_2 \approx x_0 - \frac{\ln|z_1-\bar{z}_2| - \ln|z_1-z_2| + \operatorname{Re}[(z_1-z_2)\gamma_{00}]}{2\operatorname{Im} z_1} = x_1 - \frac{\ln|z_1-\bar{z}_2| - \ln|z_1-z_2|}{2\operatorname{Im} z_1}.$$

These formulas agree with (9.2) and the ones in the introduction.

Our next objective is to determine the associated effective spectral parameters $z^{\pm}$ and phase parameter $\theta^{\pm}$. These are determined also with ϵ accuracy as follows:

(i) The imaginary parts $\operatorname{Im} z^{\pm}$ correspond to the amplitudes of the two bumps,

$$2\operatorname{Im} z^{\pm} \approx |Q(x^{\pm})|.$$

(ii) The real parts correspond to the frequencies near the two bumps,

$$2\operatorname{Re} z^{\pm} \approx |Q^{-1}(x^{\pm})| \operatorname{Im}\partial_x Q(x^{\pm}).$$

(iii) The phases correspond to the arguments of Q at $x^{\pm}$,

$$2\theta^{\pm} = \arg Q(x^{\pm}).$$

We now proceed to compute the three quantities. We recall that, with

$$c_+ = i[|z_1+z_2|^2 - 2(z_1z_2+\bar z_1\bar z_2) - (\bar z_1 - \bar z_2)^2] + 2\operatorname{Im}(z_1+z_2)\bar\sigma,$$
$$c_- = i[|z_1+z_2|^2 - 2(z_1z_2+\bar z_1\bar z_2) - (z_1 - z_2)^2] + 2\operatorname{Im}(z_1+z_2)\sigma,$$

we have

$$Q(x) = e^{2i\operatorname{Im}\gamma}\frac{c_+\bar\alpha e^{2\operatorname{Re}\gamma} + c_-\alpha e^{-2\operatorname{Re}\gamma}}{2\cosh(4\operatorname{Re}\gamma) + 2|z_1-\bar z_2|^2|\alpha|^2} + O\Big(\frac{1}{\cosh(4\operatorname{Re}\gamma)+|\alpha|^2}\Big). \tag{9.35}$$

For the amplitude near x^+ we have $e^{-2\operatorname{Re}\gamma}$ as the leading factor at the numerator, so we further simplify this as

$$Q = e^{2i\operatorname{Im}\gamma}c_-\frac{\alpha}{|\alpha|}\frac{\cosh(2\operatorname{Re}\gamma)|\alpha|}{\cosh(2\operatorname{Re}\gamma)^2 + |z_1-\bar z_2|^2|\alpha|^2} + O\Big(\frac{1}{\cosh(4\operatorname{Re}\gamma)+|\alpha|^2}\Big).$$

Similarly, near x^- we have

$$Q = e^{2i\operatorname{Im}\gamma}c_+\frac{\bar\alpha}{|\alpha|}\frac{\cosh(2\operatorname{Re}\gamma)|\alpha|}{\cosh(2\operatorname{Re}\gamma)^2 + |z_1-\bar z_2|^2|\alpha|^2} + O\Big(\frac{1}{\cosh(4\operatorname{Re}\gamma)+|\alpha|^2}\Big).$$

At the bump center x^+ (the approximate center is accurate enough) we can also replace σ by σ_0 given by

$$\sigma_0 := \sigma(x_0) = (z_1-z_2)\coth((z_1-z_2)\gamma_{00}),$$

to obtain, with slightly larger ϵ errors,

$$|Q(x^+)| \approx \frac{|c_-|}{2|z_1-\bar z_2|} \approx \frac{|[|z_1+z_2|^2 - 2(z_1z_2+\bar z_1\bar z_2) - (z_1-z_2)^2] - 2i\operatorname{Im}(z_1+z_2)\sigma_0|}{2|z_1-\bar z_2|}.$$

Here we can rewrite the expression at the numerator as (using again $2z = z_1+z_2$)

$$4\operatorname{Im}z(2\operatorname{Im}z + \operatorname{Im}\sigma_0) + 2i[(\operatorname{Re}(z_1-z_2)\operatorname{Im}(z_1-z_2) + 2\operatorname{Re}\sigma_0\operatorname{Im}z],$$

and, with ϵ errors, we replace $(\operatorname{Re}z_1 - \operatorname{Re}z_2)(\operatorname{Im}z_1 - \operatorname{Im}z_2)$ by $\operatorname{Re}\sigma_0\operatorname{Im}\sigma_0$ to get

$$Q(x_+) \approx \frac{|2(2\operatorname{Im}z + \operatorname{Im}\sigma_0)(2\operatorname{Im}z - i\operatorname{Re}\sigma_0)|}{|z_1-\bar z_2|}.$$

Taking the square norm we replace back $|\operatorname{Re}\sigma_0|^2$ by $|\operatorname{Re}(z_1-z_2)|^2$. Then we get

$$|Q(x^+)| = 2\operatorname{Im}z + \operatorname{Im}\sigma_0 + O(\epsilon). \tag{9.36}$$

A similar computation yields

$$|Q(x^-)| = 2\operatorname{Im}z - \operatorname{Im}\sigma_0 + O(\epsilon).$$

Based on this, we define the imaginary part of the effective spectral parameter for the bumps as

$$2\operatorname{Im}z^\pm = 2\operatorname{Im}z \pm \operatorname{Im}\sigma_0. \tag{9.37}$$

Next we consider the effective frequency parameter. Near x^+ we have

$$\operatorname{Im}Q^{-1}Q_x = 2\operatorname{Re}z + \operatorname{Im}(\alpha^{-1}\partial_x\alpha) + \operatorname{Im}(c_+^{-1}\partial_x c_+) + O(\epsilon).$$

The last term has size $O(\epsilon)$ and can be placed into the error. For the middle term we compute

$$\alpha^{-1}\partial_x\alpha = \frac{i(z_1-z_2)\cosh((z_1-z_2)\gamma_0)}{\sinh((z_1-z_2)\gamma_0)} = i\sigma_0.$$

We can again freeze σ to σ_0. This yields the approximate effective frequencies

$$z^\pm := z \pm \frac{\sigma_0}{2}. \tag{9.38}$$

Finally, we consider the phase, which at the maximal amplitude near x^+ respectively x^- is given by

$$2\theta^\pm \approx 2\operatorname{Im}\gamma \pm \arg(\alpha) + \arg(c_\mp).$$

We evaluate the three components. For γ we have

$$2\operatorname{Im}\gamma(x^\pm) = 2(\theta + (x^\pm - x_0)\operatorname{Re} z),$$

For α, using logarithmic derivatives,

$$\arg(\alpha) \approx \arg\alpha_0 + (x^\pm - x_0)\operatorname{Re}\sigma \approx \arg\alpha_0 + (x^\pm - x_0)\operatorname{Re}\sigma_0.$$

Finally, for $c_\mp$, reusing some of the computations we did for the effective frequency, we have

$$\arg(c_\mp) \approx \arg(2i\operatorname{Im} z \pm \operatorname{Re}\sigma_0) = \arg(z_\pm - \bar{z}_\mp).$$

We conclude that the phase is

$$2\theta^\pm \approx 2\theta \pm \arg(\alpha_0) + (x^\pm - x_0)[\operatorname{Re}(z_1 + z_2) \pm \operatorname{Re}\sigma_0] + \arg(z_\pm - \bar{z}_\mp)$$

which we rewrite as

$$2\theta^\pm = 2\theta \pm \arg(\alpha_0) + 2(x^\pm - x_0)\operatorname{Re} z^\pm \pm \arg(z^\pm - \bar{z}^\mp). \tag{9.39}$$

One can now match this with the known asymptotics for separated x_1, x_2 in [15], see (9.38). There we have

$$\operatorname{Re}\sigma_0 = \pm\operatorname{Re}(z_1 - z_2).$$

- The expression coming from the last term

  $$\operatorname{Im}\ln(z_1 - \bar{z}_2) \text{ or } \operatorname{Im}\ln(z_2 - \bar{z}_1)$$

  is one part of what we expect.
- The term $\arg\alpha_0$ has two components, $\operatorname{Im}\ln(z_1 - z_2)$ which we expect, and

  $$\pm\operatorname{Im}[(z_1 - z_2)\gamma_{00}] \approx \pm[\theta_1 - \theta_2 + \frac{\operatorname{Im}(z_1 z_2)}{\operatorname{Im}(z_1 + z_2)}(x_1 - x_2)].$$

  The first term combines with the first term of 2θ to give θ_1 or θ_2. The second term combines with the second term of θ and with the expression $(x^\pm - x_0)[\operatorname{Re}(z_1 + z_2) \pm \operatorname{Re}\sigma_0]$ with $x^\pm$ replaced by x_1 or x_2 and σ_0 as above, and they all cancel.
- We are left with the extra error coming from the substitution $x^\pm$ by x_1 (or x_2) which is

  $$\operatorname{Re} z_1(x_1 - \hat{x}_1)$$

  which does not appear in the earlier asymptotics. But this is simply a matter of notations, i.e. in our computations the new phase is evaluated at $\hat{x}_1$, whereas in [15] it is evaluated at x_1. One could also choosing the center of mass as a reference point, in which case the phase adjustment would be

  $$(x^\pm - x_0)z^\pm + i(\theta^\pm - \theta) \approx -\frac{\pi}{2} \pm \ln\alpha_0 \pm \ln 2(z^\pm - \bar{z}^\mp)$$

  But this is a less stable computation.

Summarizing the outcome of the analysis in this section, we have proved the following:

Theorem 9.3. *Let σ_0 and α_0 be defined as in* (9.30)*,*

$$\alpha_0 = \frac{\sinh((z_1 - z_2)\gamma_{00})}{z_1 - z_2}, \quad \sigma_0 = (z_1 - z_2)\frac{\cosh((z_1 - z_2)\gamma_{00})}{\sinh((z_1 - z_2)\gamma_{00})},$$

where

$$\gamma_{00} = \Big(- 2\operatorname{Im} z - i\frac{\operatorname{Im}(z_1 - z_2)\operatorname{Re}(z_1 - z_2)}{2\operatorname{Im} z}\Big)\beta_2$$
$$+ \Big(- 3\operatorname{Im} z^2 - 2i(\operatorname{Im} z)^2 + \frac{i}{4}(z_1 - z_2)^2 - \frac{3i}{4}\frac{\operatorname{Im}\Big(z(z_1 - z_2)^2\Big)}{\operatorname{Im} z}\Big)\beta_3.$$

Let $z_\pm$*,* $x_\pm$ *and* $\theta_\pm$ *be defined by* (9.38)*,* (9.33) *respectively* (9.39)*, i.e.*

$$z_\pm = z \pm \frac{\sigma_0}{2},$$
$$x_\pm = x_0 \pm \frac{\ln|z_1 - \bar{z}_2| + \ln 2|\alpha_0|}{2\operatorname{Im} z_\pm},$$
$$\theta^\pm = \theta + (x_\pm - x_0)\operatorname{Re} z^\pm \pm \frac{\arg\alpha_0 + \arg(z_\pm - \bar{z}_\mp)}{2}.$$

Then Q *is a sum of two solitons with a small error.*

(9.40)

$$Q(x) = \frac{2e^{i\theta_+ + 2i\operatorname{Re} z_+(x - x_+)}}{\operatorname{Im} z_+}\operatorname{sech}\Big(\frac{2(x - x_+)}{\operatorname{Im} z_+}\Big) + \frac{2e^{i\theta_- + 2i\operatorname{Re} z_-(x - x_-)}}{\operatorname{Im} z_-}\operatorname{sech}\Big(\frac{2(x - x_-)}{\operatorname{Im} z_-}\Big) + O\Big(\frac{\ln^2|\alpha_0|}{|\alpha_0|^2}\Big).$$

The solitons are given as a function of x by the quotient of A_0 in (9.8) and D_0 in (9.9) with γ_j and γ defined in (9.17) and (9.19), γ_0 in (9.20), γ_{00} in (9.28), (9.31), (9.32), α_0 and σ_0 in (9.30) , x_j and θ_j in (9.21) and (9.22), x_0 in (9.24) and θ in (9.25).

In other words, the above approximation has errors which are not only exponentially small in the distance between the bumps, but also uniformly small as $z_1 - z_2 \to 0$, and accurate enough to capture the leading order interaction between the two bumps.

9.4. **A uniform parametrization of the 2-soliton manifold.** We have seen that the set of pure 2-solitons is a uniformly smooth manifold in L^2, or more generally in H^s for $s > -\frac{1}{2}$. In this section we will use Proposition 9.3 to provide concrete uniform parametrizations. We will also discuss nonuniform parametrizations. We begin by discussing several ways we can smoothly parametrize the 2-soliton manifold.

a) Using the variables
$$(\mathbf{z}, \boldsymbol{\beta}),$$
employed earlier in the paper in the general case of N-solitons. Here $\boldsymbol{\beta}$ describes the correspondence between $Q_{\mathbf{z},0}$ and $Q_{\mathbf{z},\boldsymbol{\beta}}$ using the first four flows. This is the simplest description, but it is only uniform in the region $|\boldsymbol{\beta}| \lesssim 1$. Here we need to take a double quotient space for $\boldsymbol{\beta}$, namely modulo $\kappa_j \in \pi i\mathbb{Z}$.

b) Centering the 2-soliton around the center of mass and phase[3] (x_0, θ) (which can be viewed as associated to the global translation and phase shift symmetries), we can instead use the following set of parameters:
$$(\mathbf{z}, x_0, \theta, \gamma_{00}).$$
Here γ_{00} is linearly equivalent to β_2 and β_3. In this case the quotient structure decouples partially. Precisely, we have
$$\theta \in \mathbb{R} \quad (\mod \pi), \qquad \gamma_{00} \in \mathbb{C} \qquad (\mod (z_1 - z_2)^{-1}\pi i),$$

[3]We use this terminology for convenience here, but the notion of *center of phase* does not seem to be well-defined outside of the 2-soliton manifold.

but with the nontrivial gluing

$$(\theta, \gamma_{00} + \frac{\pi i}{z_1 - z_2}) \leftrightarrow (\theta + \frac{\pi}{2}, \gamma_{00}).$$

c) By the set of parameters P of pairs:

$$(z, x_0, \theta, (z_1 - z_2)^2, \alpha_0, \mu_0)$$

obtained by replacing the parameter γ_{00} by its hyperbolic functions,

$$\alpha_0 = \frac{\sinh((z_1 - z_2)\gamma_{00})}{z_1 - z_2}, \qquad \mu_0 = \cosh((z_1 - z_2)\gamma_{00})$$

which lie on the smooth manifold

$$\mu_0^2 - (z_1 - z_2)^2 \alpha_0^2 = 1.$$

Here we have one remaining symmetry

$$(\theta, \alpha_0, \mu_0) \to (\theta + \frac{\pi}{2}, -\alpha_0, -\mu_0).$$

Alternatively, away from $\alpha_0 = 0$ one can replace μ_0 by σ_0 given by

$$\sigma_0 = (z_1 - z_2)\coth((z_1 - z_2)\gamma_{00}).$$

d) We can also parametrize the 2-solitons by the set of (approximate) effective parameters

$$\Big(z_-, z_+, x_-, x_+, \theta_-, \theta_+\Big) \tag{9.41}$$

where $\theta_\pm \in \mathbb{R}/(\pi\mathbb{Z})$, provided the solitons are well separated. Here we describe the two soliton set for separated solitons by their approximate position and their phases.

The set of pure two solitons is a uniformly smooth manifold by Theorem 6.2. The sum of two solitons with the effective parameters is clearly a uniformly smooth manifold with the uniform parametrization by these parameters. Since the Hausdorff distance between the set of pure 2 solitons and the sum of the two solitons is close in L^2 as well as in any other Sobolev space H^s we see that we obtain a uniformly smooth parametrization of the pure 2 solitons in the well separated regime.

This will turn out to be uniform, but it is defined only for separated solitons. We can also use half of the above parameters along with the center parameters

$$(z, z_+, x_0, x_+, \theta, \theta_+).$$

which can be defined via the relations

$$z = \frac{z_+ + z_-}{2}$$

$$x_0 = \frac{x_+ \operatorname{Im} z_+ + x_- \operatorname{Im} z_-}{2 \operatorname{Im} z}$$

$$\theta = \frac{\theta_+ + \theta_-}{2} - (x_+ - x_0)\operatorname{Re} z^+ - (x_- - x_0)\operatorname{Re} z^- - \arg(z_+ - z_-) - \frac{\pi}{2}$$

In order to describe a uniform parametrization we distinguish two cases.

I) **The double bump region.** This corresponds to $|\beta_2| + |\beta_3| \lesssim 1$ in (a) or equivalently to $|\gamma_{00}| \lesssim 1$ in (b), to $|\alpha_0| \lesssim 1$ in (c) but is not covered by (d). Here matters are simple because the uniform topology is used in the three cases (a), (b) and (c).

II) **Separated bumps.** This is the region covered in (d), where we the metric is simply equivalent to the euclidean metric,

$$g = dz_\pm^2 + dx_\pm^2 + d\theta_\pm^2$$

We next recast this metric in terms of the parametrization in (c). We have

$$2z_\pm = 2z \pm \sigma_0$$
$$x_\pm = x_0 \pm \frac{\frac{1}{2}\ln(4\operatorname{Im} z^2 + \operatorname{Re}(\frac{1}{\alpha_0^2} - \sigma_0^2)) + \ln(2|\alpha_0|)}{2\operatorname{Im} z_\pm}$$
$$\theta_\pm = \theta \pm \arg\alpha_0 + 2x^\pm \operatorname{Re} z^\pm \pm \arg(z^\pm - z^\mp)$$

We take the uniform coordinates (9.41) and the corresponding standard metric in these coordinates - recall that search for uniform estimates for z_1, z_2 in compact subset of the open upper half plane. We write the metric on an 8 dimensional set in a schematic fashion as

$$dz_\pm^2 + dx_\pm^2 + d\theta_\pm^2$$

and seek to express it in equivalent form in terms of the variables z, α_0 and μ_0. To be more precise we set up some notation for this section. For a real function f df denotes the differential and df^2 the quadratic form

$$(y_1, y_2) \to df^2(x)(y_1, y_2) := (df(x)y_1)df(x)y_2$$

and similarly for vector valued functions F

$$dF^2(y_1, y_2) = \langle dF(x)y_1, dF(x)y_2\rangle.$$

We identify maps to $\mathbb{C}$ with the corresponding map to $\mathbb{R}^2$.

We shall see in the end that the metric tensor is at least as large as the standard metric. This will allow to neglect some terms. we write

$$dF \sim dG \Longleftrightarrow dF^2 \sim dG^2$$

if the Gram matrices have small distance, i.e. $\|dF^T dF - dG^T dG\| \lesssim 1$.

For $z_\pm$ we have the obvious relation

$$dz_-^2 + dz_+^2 \approx dz^2 + d\sigma_0^2.$$

Next we consider $x_\pm$, for which we harmlessly discard the middle component,

$$dx_\pm \sim dx_0 \pm d(\frac{\ln|\alpha_0|}{\operatorname{Im} z_\pm}) = dx_0 \pm \left(\frac{1}{\operatorname{Im} z_\pm} d\ln|\alpha_0| - \frac{\ln|\alpha_0|}{\operatorname{Im}^2 z_\pm} d\operatorname{Im} z_\pm\right).$$

We multiply the '+' equation by $\operatorname{Im} z_+$, the '-' equation by $\operatorname{Im} z_-$ and add

$$2\operatorname{Im} z dx_0 - \ln|\alpha_0|\left(\frac{1}{\operatorname{Im} z_+} d\operatorname{Im} z_+ - \frac{1}{\operatorname{Im} z_-} d\operatorname{Im} z_-\right) = 2\operatorname{Im} z dx_0 - \frac{\ln|\alpha_0|}{\operatorname{Im} z_+ \operatorname{Im} z_-}(2\operatorname{Im} z d\sigma_0 - \sigma_0 d\operatorname{Im} z)$$

Discarding the $\operatorname{Im} z_\pm$ denominators we are left with

$$\operatorname{Im} z_- dx_- + \operatorname{Im} z_+ dx_+ \sim \operatorname{Im} z(4\operatorname{Im}^2 z - \operatorname{Im}^2\sigma)dx_0 - 2\ln|\alpha_0|(2\operatorname{Im} z d\operatorname{Im}\sigma_0 - \operatorname{Im}\sigma_0 d\operatorname{Im} z) \tag{9.42}$$

Next we take the difference to obtain

$$dx_+ - dx_- \sim \frac{1}{\operatorname{Im}^2 z_+ \operatorname{Im}^2 z_-}\left(2\operatorname{Im} z(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma)d\ln|\alpha_0| - \ln|\alpha_0|(\operatorname{Im}^2 z_- d\operatorname{Im} z_- + \operatorname{Im}^2 z_+ d\operatorname{Im} z_-)\right)$$

and discarding the fraction

$$\sim 2\operatorname{Im} z(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma)d\ln|\alpha_0| - 2\ln|\alpha_0|\left((\operatorname{Im}^2 z + \frac{1}{4}\operatorname{Im}^2\sigma_0)d\operatorname{Im} z - \operatorname{Im} z\operatorname{Im}\sigma_0 d\operatorname{Im}\sigma_0\right) \tag{9.43}$$

We repeat the same computation for $\theta_\pm$. We can harmlessly discard the last term, as well as the $dx_\pm$ component, leaving us with the equivalent form

$$d\theta_\pm \sim d\theta \pm (d\arg\alpha_0 + \frac{\log|\alpha_0|}{\operatorname{Im} z_\pm} d\operatorname{Re} z_\pm)$$

Adding the $\pm$ forms yields

$$2d\theta + \ln|\alpha_0|\left(\frac{1}{\operatorname{Im} z_+} d\operatorname{Re} z_+ - \frac{1}{\operatorname{Im} z_-} d\operatorname{Re} z_-\right) = 2d\theta + \frac{\ln|\alpha_0|}{\operatorname{Im} z_+ \operatorname{Im} z_-}\left(2\operatorname{Im} z d\operatorname{Re}\sigma_0 - \operatorname{Im}\sigma_0 d\operatorname{Re} z\right)$$

Discarding the $\operatorname{Im} z_\pm$ denominators we are left with

$$d\theta_+ + d\theta_- \sim 2(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma_0)d\theta + \ln|\alpha_0|\left(2\operatorname{Im} z d\operatorname{Re}\sigma_0 - \operatorname{Im}\sigma_0 d\operatorname{Re} z\right) \tag{9.44}$$

On the other hand taking the difference we obtain

$$d\theta_+ - d\theta_- \sim 2d\arg(\alpha_0) + \frac{2\ln|\alpha_0|}{\operatorname{Im} z_+ \operatorname{Im} z_-}(\operatorname{Im} z d\operatorname{Re} z - \frac{1}{4}\operatorname{Im}\sigma_0 d\operatorname{Re}\sigma_0)$$

and eliminating the denominators

$$d\theta_+ - d\theta_- \sim 2(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma_0)d\arg(\alpha_0) + 2\ln|\alpha_0|(\operatorname{Im} z d\operatorname{Re} z - \frac{1}{4}\operatorname{Im}\sigma_0 d\operatorname{Re}\sigma_0) \tag{9.45}$$

Combining (9.42) with (9.44) and (9.43) with (9.45) we obtain the set of forms

$$e_1 = 2(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma_0)(\operatorname{Im} z dx_0 + id\theta) + \ln|\alpha_0|\left(2\operatorname{Im} z d\sigma_0 - \operatorname{Im}\sigma_0 dz\right) \tag{9.46}$$

respectively

$$e_2 = 2(\operatorname{Im}^2 z - \frac{1}{4}\operatorname{Im}^2\sigma_0)d\ln\alpha_0 + 2i\ln|\alpha_0|(\operatorname{Im} z dz - \frac{1}{4}\operatorname{Im}\sigma_0 d\sigma_0) \tag{9.47}$$

which is equivalent to $(dx_\pm, d\theta_\pm)$,

$$dx_+^2 + dx_-^2 + d\theta_+^2 + d\theta_-^2 \sim e_1^2 + e_2^2.$$

Then the correct metric is

$$g = dz^2 + d\sigma_0^2 + e_1^2 + e_2^2. \tag{9.48}$$

Thus we can uniformly characterize the two soliton manifold:

Theorem 9.4. *The two soliton manifold is smoothly and uniformly parametrized by the parameters $(z, x_0, \theta_0, \alpha_0, \sigma_0)$ endowed with the metric (9.48) in the range when $|\alpha_0| \geq 1$, and by the parameters $(z, x_0, \theta_0, \alpha_0, \mu_0)$ endowed with the euclidean metric in the range when $|\alpha_0| \lesssim 1$.*

9.5. **Double solitons.** These are the limiting solitons where we have a double eigenvalue. They are parametrized by the eigenvalue z and the flow parameters β_0, β_1, β_2 and β_3. To better describe the bump locations, we translate these into the alternative set consisting of the spectral parameter z, the center of mass/momentum x_0, θ and γ_{00}.

For the center (x_0, θ) it is easiest to use the formula (9.26), which yields

$$\theta - zx_0 = \beta_0 + \beta_1 z + \beta_2 z^2 + \beta_3 z^3.$$

Matching imaginary parts we get

$$x_0 = -\beta_1 - 2\operatorname{Re} z\beta_2 - (3\operatorname{Re}^2 z - \operatorname{Im}^2 z)\beta_3,$$

and matching real parts,

$$\theta = \beta_0 - |z|^2\beta_2 - 2\operatorname{Re} z|z|^2\beta_3.$$

On the other hand by (9.31) and (9.32) we have

$$\alpha_0 = 1/\sigma_0 = \gamma_{00} = -2\operatorname{Im} z\beta_2 - (6\operatorname{Re} z\operatorname{Im} z + 2i\operatorname{Im}^2 z)\beta_3,$$

and

$$\alpha = -2\operatorname{Im} z\beta_2 - (3\operatorname{Im} z^2 + 2i\operatorname{Im}^2 z)\beta_3 + i(x - x_0).$$

For simplicity we set $x_0 = \theta = 0$, which amount to a shift in x plus adjusting the phase. We plug these values into the expression of Proposition 9.3 to obtain the corresponding approximate spectral and scattering parameters

$$z_{\pm} = z \pm \frac{1}{2\gamma_{00}},$$

$$x_{\pm} = x_0 \pm \frac{2\ln 2 + \ln \operatorname{Im} z + \ln |\gamma_{00}|}{2 \operatorname{Im} z_{\pm}},$$

and

$$\theta_{\pm} = \theta \pm \frac{\operatorname{Re} z_{\pm}}{2\operatorname{Im} z_{\pm}}\Big(2\ln 2 + \ln \operatorname{Im} z + \ln|\gamma_{00}|\Big) \pm \frac{1}{2}(\arg \gamma_{00} + \arg(z_{\pm} - \bar z_{\mp})).$$

These formulas are valid when $|\gamma_{00}| \gg 1$, which corresponds to $|\beta_2| + |\beta_3| \gg 1$.

It may also be interesting to write down the exact formula for the 2-soliton, namely

$$Q = -4\operatorname{Im} z \frac{e^{2\gamma} + e^{-2\bar\gamma} + 2i\operatorname{Im} z(\bar\alpha e^{2\gamma} + \alpha e^{-2\bar\gamma})}{|\cosh(2\gamma)|^2 + |\sinh(2\gamma)|^2 + 1 + 8|\operatorname{Im} z|^2|\alpha|^2}, \tag{9.49}$$

with

$$2\alpha = 2\gamma_0 = 2\gamma_{00} + 2i(x - x_0) = 2(a_2\beta_2 + a_3\beta_3) + 2i(x - x_0)$$

We scale and apply a Galilean transform to normalize to $z = i$. Then $\gamma_{00} = -2\beta_2 - 2i\beta_3$. After a translation and a phase change we have $x_0 = \theta = 0$ which leads to $\beta_1 = \beta_3$, $\beta_0 = \beta_2$ and with (9.19) $\gamma = -2x$. Then $\alpha = -2\beta_2 + i(x - 2\beta_3)$, and the normalized double soliton has the form

$$Q = 4\frac{(1 - 4i\beta_2)\cosh(2x) - 2(x - 2\beta_3)\sinh(2x)}{\cosh^2(2x) + 4(4\beta_2^2 + (x - 2\beta_3)^2)}. \tag{9.50}$$

which we plot for selected parameters. First we show real 2-soliton functions corresponding to $\beta_2 = 0$:

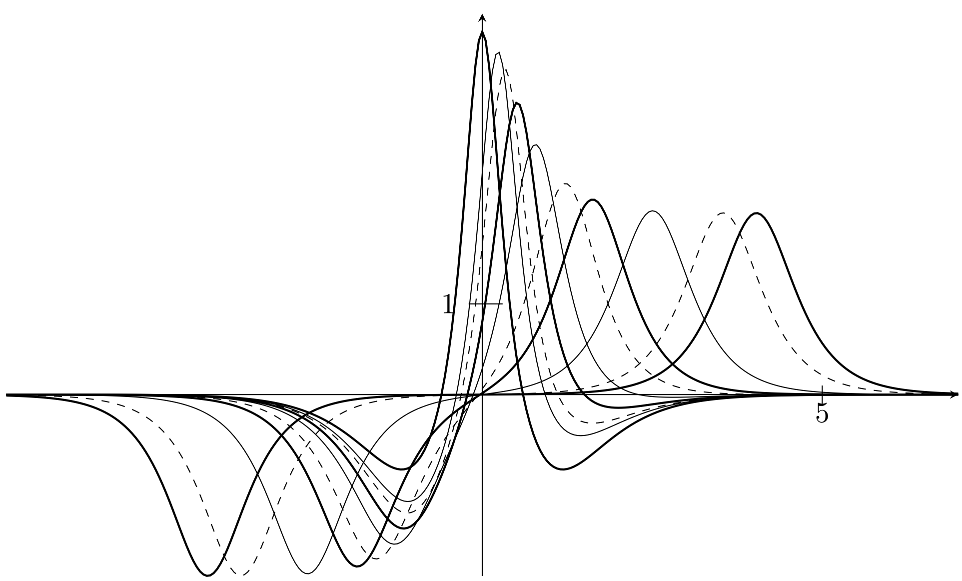

FIGURE 1. $\beta_2 = 0, \beta_3 = 0, 0.2, 0.3, 0.5, 0.9, 2.4, 6, 20, 150, 400.$

In the general case we plot real and imaginary values for a few values of β_2 and β_3, see Figure 2.

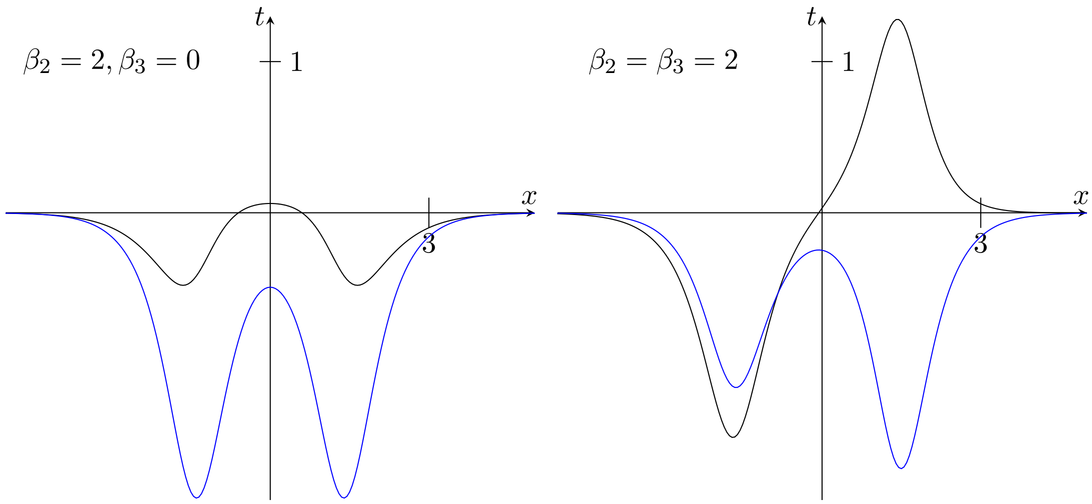

FIGURE 2. The real part is shown with a black line, the imaginary part by a blue line.

9.6. **A description of the two soliton dynamics for NLS.** In this section we describe the possible patterns for the interaction of two solitons with nearby spectral parameters along the NLS flow.

Along the flow the two spectral parameters z_1 and z_2 stay fixed, while the scattering parameters κ_1 and κ_2 evolve along the NLS flow according to

$$\dot{\kappa}_1 = iz_1^2, \qquad \dot{\kappa}_2 = iz_2^2$$

which expressed in terms of the β's becomes

$$\dot{\beta}_0 = 0, \dot{\beta}_1 = 0, \dot{\beta}_2 = 1, \dot{\beta}_3 = 0,$$

i.e. β_3 is the NLS time, which we redenote by t, and the others stay fixed. We set the trivial parameters β_0 and β_1 to zero, and we work out the formulas for the approximate effective position of Proposition 9.3 in this case. We begin with the center of mass, which moves with velocity

$$\dot{x}_0 = -\frac{\operatorname{Im}(z_1^2 + z_2^2)}{\operatorname{Im}(z_1 + z_2)}.$$

The remaining interesting parameter is γ_{00}, which also moves linearly, with velocity

$$\dot{\gamma}_{00} = a_2,$$

where we recall that the coefficient a_2 is given by

$$a_2 = i\frac{(z_1 + z_2)\operatorname{Im}(z_1 + z_2) - \operatorname{Im}(z_1^2 + z_2^2)}{\operatorname{Im}(z_1 + z_2)} = -\operatorname{Im}(z_1 + z_2) - i\frac{(\operatorname{Re} z_1 - \operatorname{Re} z_2)(\operatorname{Im} z_1 - \operatorname{Im} z_2)}{\operatorname{Im} z_1 + \operatorname{Im} z_2}.$$

Assuming that z_1 and z_2 are close, this has a small real part, and an imaginary part which is away from zero. Then we write γ_{00} in the form

$$\gamma_{00}(t) = a + a_2 t,$$

where the complex parameter a is our remaining degree of freedom. We can use time translations to further normalize a, e.g. by choosing it purely real, and then periodicity to insure that $|a| \lesssim |z_1 - z_2|^{-1}$.

With these notations we have

$$\alpha_0(t) = \frac{\sinh\left((z_1 - z_2)(a + ta_2)\right)}{z_1 - z_2},$$

$$\sigma_0(t) = (z_1 - z_2)\coth\left((z_1 - z_2)(a + ta_2)\right),$$

and the approximate effective bump position is

$$x_\pm(t) = x_0(t) \pm \frac{\ln|z_1 - \bar z_2| + \ln 2|\alpha_0(t)|}{\operatorname{Im} z_1 + \operatorname{Im} z_2 \pm \sigma_0(t)}.$$

To understand the behavior of the two bumps in time, we need to look at the location of the line

$$L: \qquad t \to (z_1 - z_2)\gamma_{00} = (z_1 - z_2)(a + ta_2)$$

relative to the imaginary axis, and, more importantly, relative to $i\pi\mathbb{Z}$. Based on this relative position we distinguish two main scenarios, with several interesting subcases each. These are described in terms of the difference $\delta z = z_2 - z_1$ of the two spectral parameters.

(a) The double soliton case, $\delta z = 0$, which will be viewed both separately and as a limit of the scenarios below.

(b) Split velocities, where L is fully transversal to the imaginary axis. This corresponds to $|\operatorname{Im}\delta z| \gtrsim |\operatorname{Re}\delta z|$. Depending on how how close L gets to $i\pi\mathbb{Z}$ we have two subcases:
   (i) Nonresonant, where $d(L, i\pi\mathbb{Z}) \approx 1$, where the two bumps stay as far as possible from each other, i.e. $|\log|\delta z||$ at the closest approach.
   (ii) Resonant, where $d(L, i\pi\mathbb{Z}) \ll 1$, and the bumps approach closer than the above threshold. The double soliton case can be seen as a limit of this scenario where a is fixed, and the closest approach is $|\log|a||$.

(c) Split scales, where L is close to parallel to the imaginary axis. This corresponds to $|\operatorname{Im}\delta z| \ll |\operatorname{Re}\delta z|$. Depending on how close L gets to $i\pi\mathbb{Z}$ we also have two main subcases, and an interesting limiting case:
   (i) Nonresonant, where $d(L, i\pi\mathbb{Z}) \approx d_0 = |\operatorname{Im}\delta z|/|\operatorname{Re}\delta z|$, where the two bumps stay as far as possible from each other, i.e. $|\log|\operatorname{Im}\delta z||$ at the closest approach.
   (ii) Resonant, where $d(L, i\pi\mathbb{Z}) = d_1 \ll d_0$, and the bumps approach closer than the above threshold. The double soliton case can also be seen as a limit of this scenario where a is fixed and the closest approach is $|\log|a||$.
   (iii) Quasiperiodic, where $\operatorname{Im}\delta z = 0$ and L is parallel to the imaginary axis, at distance d. There the soliton distance oscillates between $\log|a|$ and $\log|z_1 - z_2|$.

We successively discuss each of these scenarios in turn.

(a) The double solitons $z_1 = z_2$. There we have

$$\dot x_0 = -2\operatorname{Re} z, \qquad \dot\gamma_{00} = 2i\operatorname{Im} z.$$

Hence after suitable space and time translations we can set

$$x_0 = -2t\operatorname{Re} z$$

$$\gamma_{00} = 2it\operatorname{Im} z + a, \qquad a \in \mathbb{R}$$

Hence if $|a| \gg 1$ we have the approximate bump locations

$$z_\pm = z \pm \frac{1}{4it\operatorname{Im} z + 2a}$$

$$x_\pm = x_0 \pm \frac{2\ln 2 + \ln\operatorname{Im} z + \ln|2it\operatorname{Im} z + a|}{2\operatorname{Im} z_\pm}$$

The separation between the two bumps is $O(\ln|\gamma_{00}|)$, with a $\log|a|$ minimum. The trajectories of the bumps are like in the following picture:

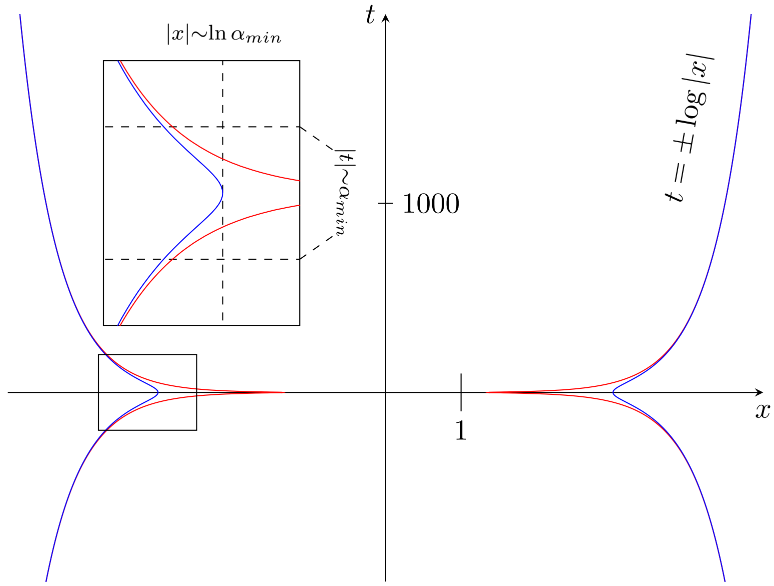

FIGURE 3. The path $x_{\pm}(t)$ of the solitons for an eigenvalue $z = i$ of multiplicity 2. The red curve has $a = 2$ and the blue curve $a = 100$. Turning is smooth, as seen in the enlarged window.

More general 2*-solitons.* Now we consider the case of two different but close spectral parameters. We use a galilean transformation and a translation to normalize so that the center of mass is time independent at $x_0 = 0$, and set

$$z_1^2 + z_2^2 = -2$$

so that both z_1 and z_2 are close to i. To describe the dynamics we will use the small parameter $\delta z = z_1 - z_2$. This parameter will play a major role for in the region where $|\operatorname{Re}\alpha_0| \lesssim 1$, which happens for a time range

$$T \approx \frac{1}{\operatorname{Re}\delta z},$$

after which the interaction of the two bumps trivializes, in the sense that the two bumps will evolve linearly but with a spatial shift as predicted by Proposition 9.3.

(b)(i) Split velocities $|\operatorname{Im}\delta z| \lesssim |\operatorname{Re}\delta z|$ *and nonresonant* $\operatorname{dist}(\delta z\gamma_{00}, i\pi\mathbb{Z}) \gtrsim 1$. Then solitons come together with their respective speed, until they reach distance $-\log|\delta z|$. Then they exchange spectral parameters and move away. The effective scattering parameters are shifted between the asymptotes at $\pm\infty$ by $\ln|z|$.

(b)(ii) Split velocities $|\operatorname{Im}\delta z| \lesssim |\operatorname{Re}\delta z|$ *and resonant* $\operatorname{dist}(\delta z\gamma_{00}, i\pi\mathbb{Z}) = r \ll 1$. Then solitons come together with their respective speed, until they reach distance $-\log|\delta z|$ but then they continue to approach logarithmically for another $-\log r$ before turnaround. In the limiting case $z_1 = z_2$ then this reduces to only the logarithmic pattern which comes up to minimal distance $-\log|a|$.

(c)(i) Split scales $|\operatorname{Im}\delta z| \gg |\operatorname{Re}\delta z|$ *and nonresonant* $r_0 = \operatorname{dist}(\delta z\gamma_{00}, i\pi\mathbb{Z}) \approx \dfrac{\operatorname{Re}\delta z}{\operatorname{Im}\delta z}$. Then solitons come together with their respective speed, until they reach distance $-\log|\delta z|$ but then they continue to approach logarithmically until distance $-\log r_0$ before turnaround; this pattern repeats until distance grows again above $-\log|\delta z|$, for a time $T \approx 1/\Re\delta z$.

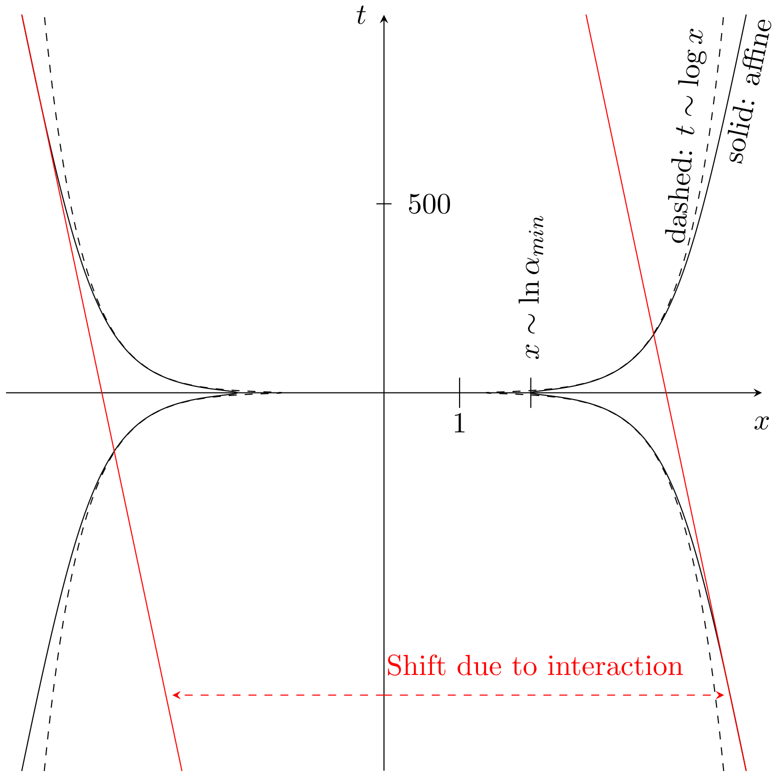

FIGURE 4. For reference, the dashed lines are the path for the double eigenvalue i. The solid line is the path of the two solitons with $z_1 = 0.0005 + i$, $z_2 = -0.0005 + i$ and $a = 10$, which switches back and forth from affine to logarithmic shape. The red lines show the asymptotic shift.

(c)(ii) Split scales $|\operatorname{Im} z| \gg |\Re z|$ and resonant $r = \operatorname{dist}(\delta z \gamma_{00}, i\pi\mathbb{Z}) \lesssim \dfrac{\operatorname{Re} \delta z}{\operatorname{Im} \delta z}$. This is the same periodic pattern as above but it comes closer in exactly once, to distance $-\log r$.

(c)(iii) The quasiperiodic solutions $\Im \delta z = 0$. In this case we can set

$$\gamma_{00} = 2i\Im z + a, \qquad a \in \mathbb{R}$$

which leads to the time frequency

$$\omega = 2 \operatorname{Re} \delta z \Im z$$

and the time period $2\pi/\omega$.

Assuming that

$$1 \ll |a| \lesssim |z_1 - z_2|^{-1},$$

the maximal distance from the center of mass axis is approximately $\log |z_1 - z_2|$ and the minimal distance is about $\log |a|$. Else we get a uniform $\log |a|$ distance.

9.7. **A description of the 2-soliton dynamics for mKdV.** Here we are interested in real solutions to

$$u_t + u_{xxx} + 6u^2 u_x = 0.$$

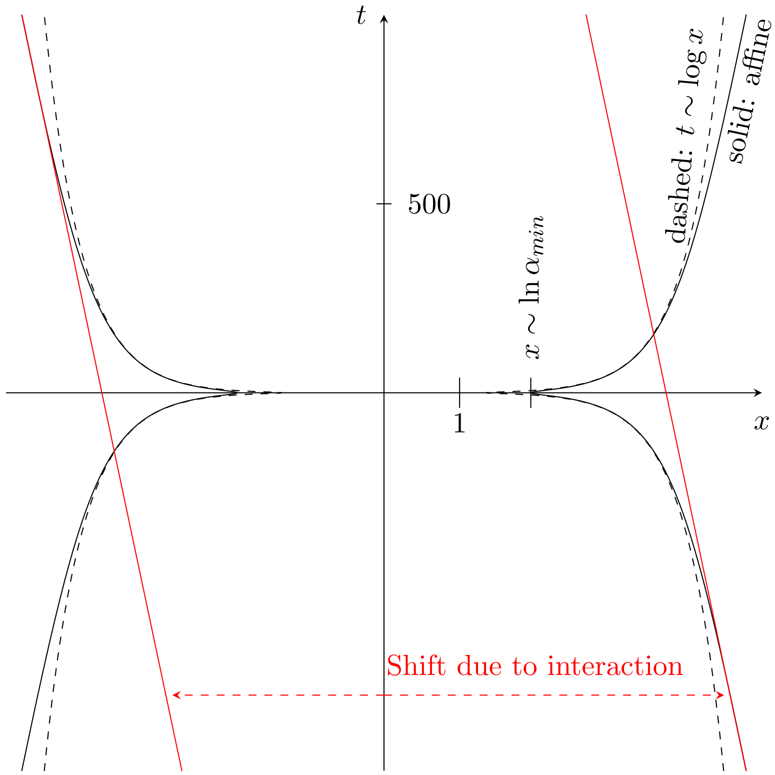

FIGURE 5. Here $a = 1$. The dashed curve corresponds to the double eigenvalue $z = i$ for reference. The other curves correspond to $z_{0,1} = (1 \pm 0.01)i \mp h$ where $h = 0, 0.002, 0.004$ and 0.016.

In this case the spectrum is symmetric under the reflection on the imaginary axis $z \to -\bar{z}$. Hence 1-soliton solutions correspond to pure imaginary eigenvalues iw with $w > 0$. Then $\theta \in \pi/2\mathbb{Z}$, the unbounded iw wave is

$$\begin{pmatrix} e^{-w(x-x_0)-w^3 t} \\ e^{w(x-x_0)-w^3 t} \end{pmatrix}$$

and the soliton has speed w^2 and the explicit form

$$w \cosh(w(x - x_0 - w^2 t)).$$

For the two soliton case we first specialize the previous formulas. We are interested in real solutions with close eigenvalues, which we write as

$$z_{1,2} = i(w \pm \rho)$$

with $\rho^2 \in \mathbb{R}$. Then we have

$$\dot{x}_0 = w^2 + 3\rho^2, \qquad \dot{\gamma}_{00} = 2i(\rho^2 - w^2)$$
$$\gamma = -w[(x - x_0) - (w^2 + 3\rho^2)t]$$
$$\gamma_1 - \gamma_2 = i\tilde{\gamma} = 2\rho(x - x_0 - t(3w^2 + \rho^2)) + i\mu.$$

for some fixed constant μ. With these parameters, the two soliton Q is given by

$$Q = -2\frac{e^{2\gamma}\Big[2w\cosh(\overline{\gamma_1 - \gamma_2}) - \dfrac{4w^2}{\bar{\rho}}\sinh(\overline{\gamma_1 - \gamma_2})\Big] + e^{-2\gamma}\Big[2w\cosh(\gamma_1 - \gamma_2) + \dfrac{4w^2}{\rho}\sinh(\gamma_1 - \gamma_2)\Big]}{4\cosh^2(2\gamma) + 2|\cosh(\gamma_1 - \gamma_2)|^2 - 2 + 8(2(w^2/\rho^2) - 1)|\sinh(\gamma_1 - \gamma_2)|^2.}$$

Since we are interested in real solitons, the parameter μ cannot be chosen arbitrarily. Precisely, there are three possibilities for real solutions:

a) Both eigenvalues are distinct and purely imaginary, $\rho \in (0.w)$. Here we have four connected components, which up to symmetries are grouped in two subcases:

(i) $\mu \in 2\pi\mathbb{Z}$, which gives the formula

$$Q = -\frac{4w\cosh(2\gamma)\cosh(\gamma_1-\gamma_2) - 4w^2\sinh(2\gamma)\rho^{-1}\sinh(\gamma_1-\gamma_2)}{2\cosh^2(2\gamma) + (8(w^2/\rho)-3)\sinh^2(\gamma_1-\gamma_2)}.$$

which corresponds to two bumps of opposite signs, and $\mu \in \pi + 2\pi\mathbb{Z}$ which yields the soliton $-Q$ with two negative bumps.

(ii) If $\mu \in 2\pi\mathbb{Z} + \pi/2$ then we divide by i and get

$$Q = -\frac{4w\sinh(2\gamma)\sinh(\gamma_1-\gamma_2) - 4w^2\cosh(2\gamma)\rho^{-1}\cosh(\gamma_1-\gamma_2)}{2\sinh^2(2\gamma) + (8(w^2\rho^{-2})-3)\cosh^2(\gamma_1-\gamma_2)}.$$

which corresponds to two positive bumps, and $\mu \in \frac{3}{2}\pi + 2\pi\mathbb{Z}$ which yields $-Q$.

b) Distinct complex conjugate eigenvalues $z_{1,2} = iw \pm \tilde{\rho}$ for some $\tilde{\rho} > 0$. This is the breather solution, which is time periodic in a moving frame. Then $\tilde{\gamma}$ is real valued, so after a time translation we can set $\mu = 0$, and

$$Q = -\frac{4w\cosh(2\gamma)\cos(\tilde{\gamma}) - 4w^2\sinh(2\gamma)\tilde{\rho}^{-1}\sin(\tilde{\gamma})}{2\cosh^2(2\gamma) + (8w^2/\tilde{\rho}^2+3)\sin^2(\tilde{\gamma})}.$$

c) The double eigenvalue 2-soliton, with eigenvalues $z_{1,2} = iw$, which is the transition between the regimes a) i) and b) above. The formula for the 2 soliton is obtained as limit of the formulas above.

$$Q = \frac{4w\cosh(2w(x-x_0-tw^2)) - 4w^2(x-x_0-3tw^2)\sinh(2(x-x_0-tw^2))}{2\cosh^2(2w(x-x_0-tw^2)) + 8w^2(x-x_0-3tw^2)^2}.$$

Precisely, we can see the solitons in (a)(i), (b) and (c) above as a single, analytic function of ρ^2. If $\rho^2 < 0$ we have breathers, and if $\rho^2 > 0$ we have two soliton states with eigenvalues $i(w \pm |\rho|)$. By contrast, the solitons in (a)(ii) do not connect to the double eigenvalue and breather case. We briefly discuss the three cases further below.

Case c) The double eigenvalue $z_{1,2} = iw$. Here the center of mass evolves according to

$$\dot{x}_0 = w^2$$

whereas α_0 is purely imaginary, and is given by

$$\dot{\alpha}_0 = -2iw^2$$

After a time translation we can set

$$\alpha_0 = -2itw^2,$$

and then the two bumps are nearly symmetric around the center of mass, at distance

$$x_\pm - x_0 \approx \pm\frac{\ln\langle |t|w^2\rangle}{2w}, \qquad |t| \gg 1.$$

The two soliton looks like a sum of two simple solitons unless the two are close. Several time sections of the graph are sketched in Figure 1.

Case a) The 2*-soliton* $z_{1,2} = iw_{1,2}$. Here the center of mass evolves according to

$$\dot{x}_0 = w_1^2 + w_2^2 - w_1 w_2 = \left(\frac{w_1 + w_2}{2}\right)^2 + \frac{3}{4}(w_1 - w_2)^2$$

whereas in case (i) γ_{00} is purely imaginary, and is given by

$$\dot{\gamma}_{00} = -2iw_1 w_2$$

but in case (ii) it is shifted by $\frac{i\pi}{w_1 - w_2}$. Here we have two different pictures.

In case (i), which corresponds to bumps of opposite sign, the solitons cross each other and, relative to the center of mass, their centers move as in the similar NLS picture in Figure 4 but with a smaller a.

In case (ii), on the other hand, the two solitons approach only to distance $\log |w_1 - w_2|$, where they exchange the effective spectral parameters, and then separate back.

Case b) The breather $z_{1,2} = iw \pm \rho$. Here the center of mass evolves according to

$$\dot{x}_0 = w^2 - 3\rho^2$$

which is the same as the speed of each of the two corresponding solitons, taken separately.

On the other hand γ_{00} is again purely imaginary, and is given by

$$\dot{\gamma}_{00} = -2i(w^2 + \rho^2)$$

Here this implies that

$$(z_1 - z_2)\gamma_{00} = -4i\rho(w^2 + \rho^2)$$

which yields the period

$$T = \frac{\pi}{2\rho(w^2 + \rho^2)}$$

which is presumably also the time scale on which the breather matches the double soliton.

The path of the soliton relative to the path of the center is periodic. If

$$\mathrm{dist}(\gamma_{00}, i\pi\mathbb{Z}) \ll |z_1 - z_2|$$

then the two-soliton is close a 2 soliton for the double eigenvalue. In the opposite regime the path is similar to figure 9.6, relative to the uniform movement of the center.

References

[1] Mark J. Ablowitz, David J. Kaup, Alan C. Newell, and Harvey Segur. "The inverse scattering transform-Fourier analysis for nonlinear problems". *Studies in Appl. Math.* 53.4 (1974), pp. 249–315.

[2] Tuncay Aktosun and Cornelis van der Mee. "A unified approach to Darboux transformations". *Inverse Problems* 25.10 (2009), pp. 105003, 22. ISSN: 0266-5611,1361-6420.

[3] Miguel A. Alejo and Claudio Muñoz. "Dynamics of complex-valued modified KdV solitons with applications to the stability of breathers". *Anal. PDE* 8.3 (2015), pp. 629–674. ISSN: 2157-5045.

[4] Miguel A. Alejo and Claudio Muñoz. "Nonlinear stability of MKdV breathers". *Comm. Math. Phys.* 324.1 (2013), pp. 233–262. ISSN: 0010-3616.

[5] Miguel A. Alejo and Claudio Muñoz. "On the nonlinear stability of mKdV breathers". *J. Phys. A* 45.43 (2012), pp. 432001, 7. ISSN: 1751-8113.

[6] Richard Beals and R. R. Coifman. "Scattering and inverse scattering for first-order systems. II". *Inverse Problems* 3.4 (1987), pp. 577–593. ISSN: 0266-5611.

[7] Deniz Bilman and Peter D. Miller. "A robust inverse scattering transform for the focusing nonlinear Schrödinger equation". *Comm. Pure Appl. Math.* 72.8 (2019), pp. 1722–1805. ISSN: 0010-3640,1097-0312.

[8] Michael Borghese, Robert Jenkins, and Kenneth D. T.-R. McLaughlin. "Long time asymptotic behavior of the focusing nonlinear Schrödinger equation". *Ann. Inst. H. Poincaré C Anal. Non Linéaire* 35.4 (2018), pp. 887–920. ISSN: 0294-1449,1873-1430.

[9] Radu C. Cascaval, Fritz Gesztesy, Helge Holden, and Yuri Latushkin. "Spectral analysis of Darboux transformations for the focusing NLS hierarchy." English. *J. Anal. Math.* 93 (2004), pp. 139–197. ISSN: 0021-7670; 1565-8538/e.

[10] Michael Christ, James Colliander, and Terence Tao. "A priori bounds and weak solutions for the nonlinear Schrödinger equation in Sobolev spaces of negative order". *J. Funct. Anal.* 254.2 (2008), pp. 368–395. ISSN: 0022-1236.

[11] Andres Contreras and Dmitry Pelinovsky. "Stability of multi-solitons in the cubic NLS equation". *J. Hyperbolic Differ. Equ.* 11.2 (2014), pp. 329–353. ISSN: 0219-8916.

[12] Scipio Cuccagna and Dmitry E. Pelinovsky. "The asymptotic stability of solitons in the cubic NLS equation on the line". *Appl. Anal.* 93.4 (2014), pp. 791–822. ISSN: 0003-6811.

[13] P. Deift and X. Zhou. "A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation". *Ann. of Math. (2)* 137.2 (1993), pp. 295–368. ISSN: 0003-486X.

[14] P. A. Deift. "Applications of a commutation formula". *Duke Math. J.* 45.2 (1978), pp. 267–310. ISSN: 0012-7094,1547-7398.

[15] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian methods in the theory of solitons*. Springer Series in Soviet Mathematics. Translated from the Russian by A. G. Reyman [A. G. Reĭman]. Springer-Verlag, Berlin, 1987, pp. x+592. ISBN: 3-540-15579-1.

[16] Fritz Gesztesy and Helge Holden. *Soliton equations and their algebro-geometric solutions. Vol. I.* Vol. 79. Cambridge Studies in Advanced Mathematics. $(1+1)$-dimensional continuous models. Cambridge University Press, Cambridge, 2003, pp. xii+505. ISBN: 0-521-75307-4.

[17] Benoit Grébert and Thomas Kappeler. *The defocusing NLS equation and its normal form.* EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2014, pp. x+166. ISBN: 978-3-03719-131-6.

[18] Axel Grünrock. "On the hierarchies of higher order mKdV and KdV equations". *Cent. Eur. J. Math.* 8.3 (2010), pp. 500–536. ISSN: 1895-1074.

[19] Benjamin Harrop-Griffith, Rowan Killip, and Monica Visan. "Sharp wellposednes for the cubic NLS and mKdV in $H^s(\mathbb{R})$". *arXiv:2003.05011* (2020).

[20] A. Hoffman and C. E. Wayne. "Orbital stability of localized structures via Bäcklund transformations". *Differential Integral Equations* 26.3-4 (2013), pp. 303–320. ISSN: 0893-4983.

[21] Todd Kapitula. "On the stability of N-solitons in integrable systems". *Nonlinearity* 20.4 (2007), pp. 879–907. ISSN: 0951-7715.

[22] Rowan Killip, Monica Vişan, and Xiaoyi Zhang. "Low regularity conservation laws for integrable PDE". *Geom. Funct. Anal.* 28.4 (2018), pp. 1062–1090. ISSN: 1016-443X.

[23] Herbert Koch and Daniel Tataru. "Conserved energies for the cubic nonlinear Schrödinger equation in one dimension". *Duke Math. J.* 167.17 (2018), pp. 3207–3313. ISSN: 0012-7094.

[24] MD Kruskal, CS Gardner, JM Green, and RM Miura. "Method for solving Korteweg–de Vries equation". *Phys. Rev. Lett* 19 (1967), pp. 1095–1098.

[25] Peter D. Lax. "Integrals of nonlinear equations of evolution and solitary waves". *Comm. Pure Appl. Math.* 21 (1968), pp. 467–490. ISSN: 0010-3640.

[26] Liming Ling, Li-Chen Zhao, and Boling Guo. "Darboux transformation and multi-dark soliton for N-component nonlinear Schrödinger equations". *Nonlinearity* 28.9 (2015), pp. 3243–3261. ISSN: 0951-7715,1361-6544.

[27] F. Magri. “Eight lectures on integrable systems”. *Integrability of nonlinear systems (Pondicherry, 1996)*. Vol. 495. Lecture Notes in Phys. Written in collaboration with P. Casati, G. Falqui and M. Pedroni. Springer, Berlin, 1997, pp. 256–296.

[28] V. B. Matveev and M. A. Salle. *Darboux transformations and solitons*. Springer Series in Nonlinear Dynamics. Springer-Verlag, Berlin, 1991, pp. x+120. ISBN: 3-540-50660-8.

[29] Tetsu Mizumachi and Dmitry Pelinovsky. “Bäcklund transformation and L^2-stability of NLS solitons”. *Int. Math. Res. Not. IMRN* 9 (2012), pp. 2034–2067. ISSN: 1073-7928.

[30] E. Olmedilla. “Multiple pole solutions of the nonlinear Schrödinger equation”. *Phys. D* 25.1-3 (1987), pp. 330–346. ISSN: 0167-2789,1872-8022.

[31] V. E. Zakharov and A. B. Shabat. “Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media”. *Ž. Èksper. Teoret. Fiz.* 61.1 (1971), pp. 118–134.

[32] Xin Zhou. “Direct and inverse scattering transforms with arbitrary spectral singularities”. *Comm. Pure Appl. Math.* 42.7 (1989), pp. 895–938. ISSN: 0010-3640.

[33] Xin Zhou. “L^2-Sobolev space bijectivity of the scattering and inverse scattering transforms”. *Comm. Pure Appl. Math.* 51.7 (1998), pp. 697–731. ISSN: 0010-3640.

Mathematisches Institut, Universität Bonn
Email address: koch@math.uni-bonn.de

Department of Mathematics, University of California, Berkeley
Email address: tataru@math.berkeley.edu