Title
Conserved energies for the cubic nonlinear Schrödinger equation in one dimension

Journal
Duke Mathematical Journal, 167(17)

ISSN
0012-7094

Authors
Koch, H
Tataru, D

Publication Date
2018-11-01

DOI
10.1215/00127094-2018-0033

Peer reviewed

CONSERVED ENERGIES FOR THE CUBIC NLS IN 1-D.

HERBERT KOCH AND DANIEL TATARU

Abstract. We consider the cubic Nonlinear Schrödinger Equation (NLS) as well as the modified Korteweg-de Vries (mKdV) equation in one space dimension. We prove that for each $s > -\frac{1}{2}$ there exists a conserved energy which is equivalent to the H^s norm of the solution. For the Korteweg-de Vries (KdV) equation there is a similar conserved energy for every $s \geq -1$.

Contents

1. Introduction

We consider the (de)focusing cubic Nonlinear Schrödinger equation (NLS)

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

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

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

with real or complex solutions in one space dimension. The NLS equation (1.1) is invariant with respect to the scaling

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

and mKdV (1.2) 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,

$$u_0(x) \to \lambda u_0(\lambda x), \tag{1.3}$$

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

The reason we consider the two flows simultaneously is that they are commuting Hamiltonian flows. They are in effect part of an infinite family of commuting Hamiltonian flows with respect to the symplectic form

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

Each of these Hamiltonians yield joint conservation laws for all of these flows. The first several energies are as follows (see (4.6), Lemma 8.1 and Lemma 8.3 where formulas for the terms of degree 2,4, and 6 in u are given):

$$\begin{aligned} H_0 =& \int |u|^2 dx, \\ H_1 =& \frac{1}{i} \int u\partial_x \bar{u} dx, \\ H_2 =& \int |u_x|^2 + |u|^4 dx, \\ H_3 =& i \int\ u_x \partial_x \overline{u}_x + 3|u|^2 u\overline{u}_x dx \\ H_4 =& \int |u_{xx}|^2 + \frac{3}{2}|(u^2)_x|^2 + ||u|_x^2|^2 + 2|u|^6 dx. \end{aligned} \tag{1.4}$$

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 these commuting Hamiltonians generate flows as follows: H_0 generates the phase shifts, H_1 generates the group of translations, H_2 the NLS flow, H_3 the KdV flow, etc.

In this article we prove that we can extend this countable family of conservation laws to a continuous family. For all $s > -\frac{1}{2}$ we construct conserved energies E_s associated to H^s solutions for our equations. Our construction of these energies relies heavily on the scattering transform associated to these problems, which requires some extensive preliminaries. For the reader's benefit we will now state a preliminary form of our main result, which makes no reference to the scattering transform. A more complete version will be presented at the end of the next section, after a substantial review of the scattering transform; see Theorem 2.4.

Theorem 1.1. *There exists $\delta > 0$ so that for each $s > -\frac{1}{2}$ and both for the focusing and defocusing case there exists an energy functional*

$$E_s : \{u \in H^s,\ \|u\|_{l^2 DU^2} \leq \delta\} \to \mathbb{R}^+$$

with the following properties:

(1) *E_s is conserved along the NLS and mKdV flow.*
(2) *E_s agrees with the linear H^s energy up to quartic terms,*

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

(3) *The map*

$$\{u \in H^\sigma,\ \|u\|_{l^2 DU^2} \leq \delta\} \times (-\frac{1}{2}, \sigma] \ni (u, s) \to E_s(u)$$

is analytic in $u \in H^\sigma$, analytic in s for $s < \sigma$ and continuous in s at $s = \sigma$.

(4) *The functionials E_s interpolate the energies H_{2j} in the sense that*

$$E_n(u) = \sum_{j=1}^{n} \binom{n}{j} H_{2j}(u).$$

The Banach space $l^2DU^2 = L^2 + DU^2$ is the inhomogeneous version of the DU^2 space, and is first introduced in Section 5 and in greater detail in Appendix B. It is a replacement for the unusable scaling critical space $H^{-\frac{1}{2}}$ and satisfies

$$\|u\|_{l^2DU^2} \lesssim \|u\|_{B^{-\frac{1}{2}}_{2,1}} \lesssim \|u\|_{H^s}, \qquad s > -\frac{1}{2}. \tag{1.5}$$

We note that for half-integers the energies $E_{k+\frac{1}{2}}$ are not directly associated to the momenta. For the sake of completeness, in Section 8 we discuss also the construction of generalized momenta P_s, so that for half integers $P_{k+\frac{1}{2}}$ is a linear combination of the odd Hamiltonians $H_1, H_3, \cdots, H_{2k+1}$.

To further clarify the assertions in the theorem, we note that for simplicity we establish the energy conservation result 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 [4]. One key consequence of our 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 1.2. *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 (1.3). 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 $s < 0$.

This work is a natural continuation of earlier work of the authors[14], [13] and of Christ, Colliander and Tao [4] where apriori H^s bounds for NLS are obtained for a restricted range $-\frac{1}{4} \leq s < 0$, without explicit use of integrable structures, but also without providing conserved energies. The same ideas were later implemented for mKdV by the second author together with Christ and Holmer [5] for $s > -\frac{1}{8}$, as well as for KdV by Liu [16] for $s > -\frac{4}{5}$. Together with Buckmaster the first author has proven uniform in time H^{-1} bounds for the Korteweg-de Vries equation [3] using the Miura map but no inverse scattering techniques.

In contrast, Theorem 1.1 relies on inverse scattering methods via the AKNS formalism, see for instance [1, 19]. Likely our ideas here also extend to other completely integrable systems. As an illustration of this, in Section 9 we state the similar result for the KdV problem; we also outline its proof, which conveniently uses exactly the same algebraic structure as the NLS/mKdV equations.

In the periodic case Kappeler, Schaad and Topalov [10, 11] and Kappeler and Grebert [8] have proved related results for the Korteweg-de Vries equation, for modified Korteweg-de Vries and for the defocusing NLS relying on complex algebraic geometry. Also in the KdV case, there is independent work of Killip-Visan-Zhang [12] in this direction in the range $-1 \leq s < 1$, both in the periodic case and on the real line.

The structure of the paper is as follows. Sections 2-7 contain the proof of our main result, beginning with an outline of the scattering transform in Section 2, which concludes with a more complete form of our main result, see Theorem 2.4. This is followed by an analysis of the terms in the formal series for the conserved energies in the subsequent sections, and a final summation argument in Section 7. It also contains a further discussion of our main results, touching on issues such as conserved momenta, frequency envelope bounds. Section 8 provides detailed formulas for Taylor expansion of E_s of degree 4 for all s, and of degree 6 in u for the H_j. Section 9 discusses the corresponding results for the KdV equation.

There are also two appendices to the paper, both self-contained, establishing some results which are useful here but may also be of independent interest. The first appendix is devoted to a Hopf algebra structure arising in connection with the multilinear integrals appearing in the expansion for our conserved energies. We need it for the proof of Theorem 3.3. This drastically simplifies the analysis for large s. The connection of Theorem 3.3 to Hopf algebras was pointed out to us by Martin Hairer, which we gratefully acknowledge. The second appendix is concerned with U^p and V^p spaces and their properties, which are heavily used in all our estimates.

2. An outline of the construction

2.1. **An overview of the 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 defocusing NLS equation arises as 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 Schroedinger 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 $L\psi = z\psi$ 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 $L\psi = z\psi$,

$$\begin{aligned} \mathcal{P} &= 2i \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} L^2 + 2 \begin{pmatrix} 0 & u \\ \bar u & 0 \end{pmatrix} L + i \begin{pmatrix} -|u|^2 & u_x \\ -\bar u_x & |u|^2 \end{pmatrix} \\ &= -2i \begin{pmatrix} -\partial_x^2 + |u|^2 & u_x \\ -\bar u_x & \partial_x^2 - |u|^2 \end{pmatrix} + 2i \begin{pmatrix} |u|^2 & -u\partial_x \\ \bar u \partial_x & -|u|^2 \end{pmatrix} + i \begin{pmatrix} -|u|^2 & u_x \\ -\bar u_x & |u|^2 \end{pmatrix} \\ &= i \begin{pmatrix} 2\partial_x^2 - |u|^2 & -u\partial_x - \partial_x u \\ \bar u \partial_x + \partial_x \bar u & -2\partial_x^2 + |u|^2 \end{pmatrix} \end{aligned}$$

This is equivalent to the pair of Kappeler and Grebert [8]. The same applies for the defocusing mKdV problem with respect to the system

$$
\begin{aligned}
\psi_x &= \begin{pmatrix} -iz & u \\ \bar{u} & iz \end{pmatrix} \psi \\
\psi_t &= i \begin{pmatrix} -[4iz^3 + 2iz^2|u|^2] & 2z^2u + 2izu_x - u_{xx} + 2u^3 \\ 4z^2u - 2uu_x - u_{xx} + 2u^3 & 4iz^3 + 2iz^2|u|^2 \end{pmatrix} \psi
\end{aligned} \tag{2.2}
$$

The Lax pairs associated to the focusing equations are similar up to some sign twists resp. the replacement of $\bar{u}$ by $-\bar{u}$. Precisely, for the focusing NLS problem we have the system

$$
\begin{aligned}
\psi_x &= \begin{pmatrix} -iz & u \\ -\bar{u} & iz \end{pmatrix} \psi \\
\psi_t &= i \begin{pmatrix} -[2z^2 - |u|^2] & 2izu + u_x \\ 2iz\bar{u} + i\bar{u}_x & 2z^2 - |u|^2 \end{pmatrix} \psi
\end{aligned} \tag{2.3}
$$

while for the focusing mKdV we have the related system

$$
\begin{aligned}
\psi_x &= \begin{pmatrix} -iz & u \\ -\bar{u} & iz \end{pmatrix} \psi \\
\psi_t &= i \begin{pmatrix} -[4iz^3 - 2iz^2|u|^2] & 2z^2u + 2izu_x - u_{xx} + 2u^3 \\ 4z^2u - 2uu_x - u_{xx} + 2u^3 & 4iz^3 - 2iz^2|u|^2 \end{pmatrix} \psi
\end{aligned} \tag{2.4}
$$

Much of this formalism here and below can be found in the seminal paper by Ablowitz, Kaup, Newell and Segur [1]. The scattering transform associated to both the defocusing NLS and the defocusing MKdV is defined via the first equation of (2.1) resp. (2.2) which we write as linear system

$$
\begin{cases} \dfrac{d\psi_1}{dx} = -iz\psi_1 + u\psi_2 \\[2ex] \dfrac{d\psi_2}{dx} = iz\psi_2 + \bar{u}\psi_1 \end{cases} \tag{2.5}
$$

The last $+\bar{u}$ turns into $-\bar{u}$ for the focusing case. In the defocusing case, 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

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

respectively

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

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

$$
|T| \leq 1, \qquad |T|^2 = 1 - |R|^2 = 1 - |L|^2
$$

Further, we have the symmetry $(\psi_1, \psi_2) \to (\bar{\psi}_2, \bar{\psi}_1)$ which via the Wronskian leads to

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

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

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

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

This provides a holomorphic extension of T^{-1} to the upper half-space. In the defocusing case it also provides a holomorphic extension of T to the upper half space with $|T| \leq 1$. To see that this extension is holomorphic, one needs to show that there are no z for which $T^{-1}(z) = 0$. Indeed, a straightforward ODE analysis shows that such solutions ψ would have to decay exponentially at both ends,

$$\psi = \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix} (1 + o(1)) \text{ as } x \to -\infty, \qquad \psi = \begin{pmatrix} 0 \\ e^{izx} \end{pmatrix} (d + o(1)) \text{ as } x \to \infty$$

Then one can view z and ψ as an eigenvalue/eigenfunction for the problem

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

In the defocusing case $\mathcal{L}$ is self-adjoint, so no such eigenfunctions exist outside the continuous spectrum $\mathbb{R}$ of $\mathcal{L}$. To see that this extension satisfies $|T| \leq 1$ we begin with the simpler case $u \in L^1$, when one can easily show that in the upper half space we have

$$\lim_{|z|\to\infty} T(z) = 1$$

and then the property $|T| \leq 1$ on the real line extends to the upper half-space by the maximum principle. For more general $u \in l^2DU^2$ this property extends by density.

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

$$T_t = 0, \qquad L_t = -4i\xi^2 L, \qquad R_t = 4i\xi^2 R.$$

and if u evolves according to (1.2)

$$T_t = 0, \qquad L_t = -8i\xi^3 L, \qquad R_t = 8i\xi^3 R.$$

The scattering map for u is given by

$$u \to R,$$

and the map $u \to R$ conjugates the NLS flow (1.1) to the (Fourier transform of) the linear Schroedinger equation, and simultaneously the mKdV flow to the linear Airy flow. Reconstructing u from R requires solving a Riemann-Hilbert problem, see [7] for this approach for the modified Korteweg-de Vries equation.

A key difference between real and nonreal z is that for real ξ, 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 $T(z)$. We greatly exploit this property, as our fractional conserved energies are all defined in terms of only the values of T away from the real axis.

The situation is more complicated in the focusing case, due to the fact that T is no longer holomorphic in the upper half space, which is closely related to the fact that the corresponding Lax operator $\mathcal{L}$ is no longer self-adjoint. Now the quantity $|\psi_1|^2 + |\psi_2|^2$ is conserved, which shows that

$$|T|^2 - |R|^2 = 1$$

on the real line. The function T^{-1} still has a holomorphic extension to the upper half-plane, so as above we get the relation $|T| \geq 1$ in the upper half-plane.. But now T may have poles in the upper half-space, which correspond to nonreal eigenvalues of $\mathcal{L}$. Arguing either by the holomorphy of T^{-1} or by standard Fredholm theory, the poles of T must be isolated in the open upper half space, though they can accumulate on the real line. At this point the understanding of the spectrum of the AKNS operator is limited. Zhou [22] has constructed an example of a Schwartz potential with infinitely many eigenvalues, showing that the situation is much more complex than for Schrödinger operators.

For data u for which T is holomorphic in the upper half-space, the analysis is similar to the defocusing case. 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 [22], 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 of the real line

$$\{0 \leq \operatorname{Im} z \lesssim_{\|u\|_{H^s}} (1 + |\operatorname{Re} z|)^{-2s}\}$$

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

$$\{0 \leq \operatorname{Im} z \ll (1 + |\operatorname{Re} z|)\}.$$

This is what allows us to still construct the fractional Sobolev conservation laws even in the focusing case.

2.2. **The transmission coefficient in the upper half-plane and conservation laws.** Our construction of fractional Sobolev conserved quantities 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. In particular, for any function $\eta : \mathbb{R} \to \mathbb{R}$ the expression

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

is formally conserved. It is useful to consider at first special functions η. In the defocusing case, for instance, trace formulas show that the (real) conserved quantities

$$H_k = -\int \xi^k \ln |T(-\xi/2)| d\xi \tag{2.6}$$

can be explicitly expressed in terms of u if j is an nonnegative integer. More precisely, if u is a Schwartz function then $\ln |T|$ is a Schwartz function on the real line, and has a Taylor expansion

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

at infinity in the upper half plane. This provides easy access to the conserved energies (1.4).

In the focusing case the expansion (2.7) remains valid up to a sign twist,

$$\ln T(z) \approx \frac{1}{2\pi i} \sum_{j=0}^{\infty} H_j (2z)^{-j-1} \tag{2.8}$$

However, $\ln T$ may have poles in the upper half plane, and the right hand side in the formula (2.6) 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 (2.6) 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 nonnegative measure. Given the above discussion, a natural candidate for a fractional Sobolev conservation law would 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 will be to use the weights

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

which we extend as holomorphic functions to the subdomain $D = U \setminus i[1,\infty)$ of the upper half-space U. Thus in the defocusing case we take

$$E_s(u) = -2\int_{\mathbb{R}} (1+\xi^2)^s \ln|T(\xi/2)| d\xi, \tag{2.10}$$

rewrite it as the real part of a contour integral

$$E_s(u) = -2\operatorname{Re}\int_{\mathbb{R}} (1+\xi^2)^s \ln T(\xi/2)\, d\xi$$

and then switch the integration contour to the double half-line $i[2,\infty)$. For small s this gives directly

$$E_s(u) = 4\sin(\pi s)\int_1^\infty (t^2-1)^s \operatorname{Re}\ln T(it/2) dt, \qquad -\frac{1}{2} < s < 0. \tag{2.11}$$

For larger s the above integral will diverge, and to remedy that we need to remove the appropriate number of terms in the asymptotic expansion of $\ln T$ at infinity. Precisely, in the range

$$s < N+1, \qquad u \in H^{\max\{N,s\}}$$

we obtain

$$E_s(u) = 4\sin(\pi s)\int_1^\infty (t^2-1)^s \left(\operatorname{Re}\ln T(it/2) - \frac{1}{2\pi}\sum_{j=0}^N (-1)^j H_{2j}\, t^{-2j-1}\right) dt + \sum_{j=0}^N \binom{s}{j} H_{2j}, \tag{2.12}$$

where all terms turn out to be well-defined, and which is independent of the choice of N since for $j > s-1$ we can undo the change of the contour in the integral

$$-2\sin(\pi s)(-1)^j \int_1^\infty (t^2-1)^s t^{-2j-1} dt = i\int_{-\infty+i0}^{\infty+i0} (1+z^2)^s z^{-2j-1} dz = \pi\binom{s}{j}. \tag{2.13}$$

For the last equality we used the Taylor expansion

$$(1+z^2)^s z^{-2j-1} = \sum_{k=0}^\infty \binom{s}{k} z^{2k-2j-1}.$$

For definiteness we will almost always choose $N \le s < N+1$. Thus we set

Definition 2.1 (Energies in the defocusing case)**.** *Let $s > -\frac{1}{2}$ and $u \in H^s$. Then*
a) For $-\frac{1}{2} < s < 0$, the energy E_s is defined by (2.11)*.*
b) For $N \geq 0$ and $N \leq s < N+1$, the energy E_s is defined by (2.12)*.*

If u is Schwartz then the two definitions (2.10), respectively (2.11), (2.12) are equivalent. However, the expressions (2.11), (2.12) are much more robust, and, as we shall see, are defined directly as convergent integrals for $u \in H^s$.

Consider now the focusing case. The above discussion still applies provided that there are no poles for T in the upper half-plane. However, if there are poles then the definitions (2.10), respectively (2.11), (2.12) are no longer equivalent. Instead we will use (2.11), (2.12) directly to define the energies:

Definition 2.2 (Energies in the focusing case)**.** *Let $s > -\frac{1}{2}$ and $u \in H^s$. Then*
a) For $-\frac{1}{2} < s < 0$, the energy E_s is defined by (2.11)*, taken with the opposite sign.*
b) For $N \geq 0$ and $N \leq s < N+1$, the energy E_s is defined by (2.12)*, taken with the opposite sign.*

Here we allow for T to have (finitely many) poles on the half-line $i[1,\infty)$. We also note the role played by the smallness condition for u in l^2DU^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.

We now summarize the above discussion concerning the relation between (2.10), respectively (2.11), (2.12) . 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 we have the following relations:

Proposition 2.3 (Trace formulas)**.** *Let $N > [s]$ and $u \in \mathcal{S}$. In the defocusing case*

$$(2.14)\qquad \begin{aligned} E_s &= \int_{-\infty}^{\infty} (1+\xi^2)^s (-\operatorname{Re}\ln T(\xi/2)) d\xi \\ &= 4\sin(\pi s) \int_1^\infty (\tau^1 - 1)^s \Big[\operatorname{Re}\ln T(i\tau/2) + \frac{1}{2\pi} \sum_{j=0}^N (-1)^j H_{2j} \tau^{-2j-1} \Big] d\tau + \sum_{j=0}^N \binom{s}{j} H_{2j} \end{aligned}$$

and in the focusing case

(2.15)

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

where the k sum runs over all the poles $2z_k$ of T with multiplicity m_j.

For the focusing case we remark on the case when there are infinitely many poles for T in the upper half-space. The second expression above is always a convergent integral, whereas in the first expression we have a nonnegative 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 nonnegative 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. Now we are ready to state a more complete version of our mai result in Theorem 1.1:

Theorem 2.4. *For each $s > -\frac{1}{2}$ and both for the focusing and defocusing case the energy functionals E_s 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 trace at the real line of*[1] *$\mp \log |T|$ exists as a positive measure, and the trace formulas* (2.14) *and* (2.15) *hold with absolute convergence in all sums and integrals.*
(3) *If $\|u\|_{l^2_1 DU^2} \le 1$ then*

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

(4) *The map*

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

is analytic in $u \in H^\sigma$ in the defocusing case. In the focusing case it 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 $s < \sigma$.

The proof of the above theorem is completed in Section 7. We remark now on several direct consequences of our trace formulas:

- The energies E_s are nonnegative in the defocusing case, and in the focusing case if $s \le 0$.
- In the focusing case the eigenvalues are localized in a strip

$$(1 + |\operatorname{Re} z_j|)^s \operatorname{Im} z_j \lesssim_{\|u\|_{H^s}} 1$$

Furthermore, the following sum is bounded:

$$\sum m_j \operatorname{Im} z_j (1 + |\operatorname{Re} z_j|)^s \lesssim_{\|u\|_{H^s}} 1 \qquad (2.16)$$

This is an immediate consequence of the trace formulas if $s \le 0$ since then $\operatorname{Im} z^s > 0$ in the upper half plane. If $s \ge 0$ the trace formula for $s = 0$ implies that the imaginary part of the eigenvalues is bounded. Hence there are at most finitely many eigenvalues not giving a positive contribution to the sum if $\operatorname{Im} z_j (1 + |\operatorname{Re} z_j|)^s$. Their contribution is bounded by the L^2 norm and we arrive at the bound (2.16).

One can view the small data part of our result 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 soliton/breather solutions, where T has a finite number of poles in the upper half-space and $|T| = 1$ on the real line. As a direct byproduct of our results, one can also obtain also the H^s stability of these solutions. We consider this in detail in a subsequent paper.

2.3. **Estimates for the transmission coefficient.** Given the definition of the conserved energies in the previous subsection, the proof of our main result hinges on obtaining precise estimates for the transmission coefficient. The steps in our analysis are as follows:

I. Our analysis begins in Section 3, where we compute a formal homogeneous expansion of inverse of the transmission coefficient T, namely

$$T^{-1}(z) = 1 + \sum_{j=1}^{\infty} T_{2j}(z)$$

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

where $T_{2j}(z)$ are multilinear integral forms, homogeneous of degree $2j$ in $u, \bar{u}$. There is a similar though less explicit expansion for $\ln T$,

$$-\ln T(z) = \sum_{j=1}^{\infty} \tilde{T}_{2j}(z)$$

We have $\tilde{T}_2 = T_2$, while $\tilde{T}_{2j}$ are still multilinear integral forms of degree $2j$ in $u, \bar{u}$.

Nevertheless, it turns out that the $\tilde{T}_{2j}$ are much more localized than T_{2j}, which is the key to better decay and better algebraic properties. This observation is an interesting result in itself. Its proof is purely algebraic, and is based on a Hopf algebra structure underlying the multilinear integral calculus. This is explained in detail in Appendix A. The relevant structure of the integrals $\tilde{T}_{2j}$ is encoded in algebraic statement of Theorem A.3.

II. The first term $T_2(z)$ in the formal expansions is studied in Section 4. This is a quadratic form in u, and its contribution to the conserved norms is exactly $\|u\|_{H^s}^2$. To prove this, we estimate $T_2(z)$ in the full upper half-space.

III. The remaining terms T_{2j} in the expansion of T are only needed on the imaginary half-line $[i, i\infty)$. They are estimated in Section 5 in terms of the H^s norm of u and the DU^2 norm of u. Here the DU^2 norm arises as a convenient proxy for the scale invariant $\dot{H}^{-\frac{1}{2}}$ norm of u. Appendix B summarizes and proves some useful and some new properties of the U^2 and V^2 spaces. The bounds for T_{2j} suffice in order to establish the convergence of the formal series for T, and also they easily carry over to $\tilde{T}_{2j}$. In particular they suffice in order to bound the corresponding contributions to the energy for large enough j (namely $j > 2s+1$). Thus the proof of the theorem reduces to estimating finitely many $\tilde{T}_{2j}$ terms.

IV. The main bound for $\tilde{T}_{2j}$ is proved in Section 6, taking advantage of the better decay properties due to the better structure of $\tilde{T}_{2j}$. This suffices for $j > s+1$. For smaller j (which is only occurs if $s \geq 1$) we also need to consider the expansion of $\tilde{T}_{2j}(z)$ in powers of z^{-1} as $z \to i\infty$, and estimate the errors. This is done in Section 6 and used in Section 7 to complete the proof of our main result.

V. The Hopf algebra structure allows to explicitly calculate $\tilde{T}_{2j}$(Proposition A.2). We do these calculations for $\tilde{T}_4$ and $\tilde{T}_6$ in Section 8.

VI. There are close connections to the Korteweg- de Vries equation. The first relies on the observation that, if u is a real function and if ψ satisfies the first equation in (2.1) then $\phi = \psi_1 + \psi_2$ satisfies with $v = u_x + u^2$

$$-\phi_{xx} + v\phi = z^2\phi$$

which is the eigenvalue equation of the standard Lax operator for the Korteweg-de Vries equation. As a consequence the transmission coefficients for the Lax operator for mKdV with real potential u coincide with the transmission coefficient for the Lax operator for $v = u_x + u^2$ for KdV.

We will use a Lax-pair similar to the one for the NLS, which allows to reuse most of the constructions developed for the NLS case. One difficulty in this case is that

$$\tilde{T}_2(z) = \frac{1}{2z} \int u dx$$

which can never be bounded by an H^s norm of u. However we are able to prove that

$$\tilde{T}(i\tau) - \tilde{T}_2(i\tau)$$

is well-defined on a small ball in H^{-1}, and that $\tilde{T}_4(z)$ plays the same role as $T_2(z)$ for NLS. A similar though technically different and interesting removal procedure for the linear term is implemented in [12], which in contrast to our approach also applies to the periodic KdV equation.

3. The transmission coefficient in the upper half-plane

As seen above, the transmission coefficient in the upper half plane plays a key role in our construction of the conserved energies. Here we consider the multilinear expansion of the transmission coefficient in the upper half-plane. This section is only concerned with the algebraic aspects of this expansion, setting aside for now the questions of bounds and convergence of the formal series.

For the defocusing problem, $\ln T$ extends to a holomorphic function on the upper half plane with nonpositive real part. For the focusing problem, we expect it to be holomorphic at least outside a strip around the real line. To cover both cases, we consider a more general system of the form

$$\begin{cases} \dfrac{d\psi_1}{dx} = -iz\psi_1 + u\psi_2 \\ \dfrac{d\psi_2}{dx} = iz\psi_2 + \overline{v}\psi_1 \end{cases} \tag{3.1}$$

and construct the Jost solutions recursively in the upper half plane. The scattering problems for the defocusing, respectively the focusing case are obtained by taking $v = \pm u$. This plays no role in any of the estimates in the following sections, but for definiteness we consider the defocusing case.

Lemma 3.1. *There is a formal homogeneous expansion of T^{-1} in terms of u, v*

$$T^{-1}(z) = 1 + \sum_{j=1}^{\infty} T_{2j}(z)$$

with

$$T_{2j}(z) = \int_{x_1<y_1<x_2<y_2<\dots<x_j<y_j} \prod_{l=1}^{j} e^{2iz(y_l - x_l)} u(y_l)\overline{v(x_l)} dx_1 dy_1 \dots dx_j dy_j. \tag{3.2}$$

We remark that, at least as long as $u, v \in L^2$, each term T_{2j} is pointwise defined for $\operatorname{Im} z > 0$.

Proof. We solve (3.1) iteratively. We first compute the quadratic approximation to $T(z) - 1$. Suppose we use ψ_l and begin the iteration with $\psi_l^0 = \begin{pmatrix} e^{-izx} \\ 0 \end{pmatrix}$. The first iteration gives

$$\psi_{l,2}^1(x) = -\int_{-\infty}^{x} e^{-izx_1}\overline{v(x_1)} e^{iz(x-x_1)} dx_1$$

which inserted into the first equation yields

$$\begin{aligned} \psi_{l,1}^2(x) &= e^{izx} + \int_{-\infty}^{x} e^{-iz(x-y_1)} u(y_1) \int_{-\infty}^{y_1} e^{-izx_1}\overline{v(x_1)} e^{iz(y_1-x_1)} dx_1 dy_1 \\ &= e^{-izx}\left(1 + \int_{x_1<y_1<x} u(y_1)\overline{v(x_1)} e^{2iz(y_1-x_1)} dx_1 dy_1\right). \end{aligned}$$

This shows that the quadratic approximation to $T^{-1} - 1$ is given by

$$T_2(z) = \int_{x_1<y_1} u(y_1)\overline{v(x_1)} e^{2iz(y_1-x_1)} dx_1 dy_1 \tag{3.3}$$

We iterate the procedure and arrive at (3.2). □

For our modified energies we need to work not with T directly, but rather $\ln T$. Since the log is analytic near 1, the formal series for T will yield a formal series for $\ln T$:

Lemma 3.2. *There is a formal homogeneous expansion of* $\ln T$ *in terms of* u, v

$$-\ln T = \sum_{j=1}^{\infty} \tilde{T}_{2j}$$

where each term $\tilde{T}_{2j}(z)$ *is a linear combination of expressions of the form*

$$\int_{\Sigma} \prod_{l=1}^{j} e^{2iz(y_l - x_l)} u(y_l)\overline{v(x_l)} dx_1 dy_1 \dots dx_j dy_j. \tag{3.4}$$

where Σ *can be any domain which can be represented as a linear ordering of* x_l *and* y_l *which obeys the constraint* $x_l < y_l$ *for all* l.

The proof of the lemma is straightforward, we simply observe that each term $\tilde{T}_{2j}$ is a homogeneous polynomial of appropriate homogeneity in T_{2l} with $l \leq j$, and that such products can be expressed as iterated integrals as above. The first terms have the form

$$\tilde{T}_2 = T_2, \qquad \tilde{T}_4 = T_4(z) - \frac{1}{2}T_2(z)^2, \qquad \tilde{T}_6 = T_6(z) - T_2(z)T_4(z) + \frac{1}{3}T_3(z)^3$$

For the purpose of studying the convergence of this series, such considerations are sufficient. However, if we want at the same time to capture also the expansion of $\ln T$ in powers of z^{-1} as $z \to i\infty$, then this is no longer enough. The reason is quite simple, namely that, even if u and v are Schwartz functions, the terms T_{2n} in the expansion of T only have z^{-n} decay, whereas for the expansion of $\ln T$ the corresponding term has as leading term a constant times z^{-2n+1}, at least near the imaginary axis for Sobolev data. Thus our goal is now twofold. On one hand we want to understand the decay rates at infinity for each of these integrals, and on the other hand we want to see which of them are present in the expansion of $\ln T$.

It is useful to have a more graphical representation for these integrals. This is achieved by connecting x's and y's by non-intersecting arcs. There is exactly one way to do that and we identify the integral with symbols like . For example

 $$= \int_{x_1 < x_2 < y_1 < x_3 < y_2 < y_3} e^{2iz(y_1+y_2+y_3-x_1-x_2-x_3)} \prod u(y_j)\overline{v(x_j)} dx_j dy_j.$$

It is not difficult to see that the decay rate as $z \to i\infty$ of these integrals depends on how connected they are. Precisely, if $z = i\tau$ then, due to the exponential factor, the bulk of the integral comes from the region where $|x_j - y_j| \lesssim \tau^{-1}$, and the integral can essentially be estimated by the volume of this region. If such an integral has k connected components then this volume is comparable to τ^{-2n+k}. Thus the desired decay rate z^{-2n+1} is also the best possible decay rate, and is only achieved for integrals with fully connected symbols.

In the sequel we call an integral (3.4) connected if for every nontrivial decomposition of the sequence there are more x to the left than y, or, equivalently, if in the graphical representation there is an arc from the first x to the last y. We denote it by symbols like . We arrive at the main result of this section, namely Theorem A.3 which we repeat here.

Theorem 3.3. *The terms* $\tilde{T}_{2j}$ *of the homogeneous expansion of* $\ln T$ *in* u, v *are formal linear combinations of connected integrals.*

The first terms are (see Proposition A.2)

$-\ln T =$  $- 2$  $+ 12$  $+ 4$  $+ \dots$

There is even a polyhomogeneous formal expansion, with respect to powers of z and u resp. $\bar{u}$ which we will explore below. The proof of this proposition is nontrivial, and requires the introduction of an additional Hopf algebra type structure. For this reason, we relegate this part of the proof to Section A of the appendix. The classical conserved quantities are the coefficients of the expansion of $\ln T$ in powers of z^{-1} at infinity, which are called energies. The above theorem has the consequence that these energies are given as one dimensional integrals, and not as polynomials of one dimensional integrals. This is of course well-known and typically proven by very different techniques.

4. Positive harmonic functions in the the upper half-space

A significant step in our analysis is to transfer information about the transmission coefficient from the positive imaginary axis to the real line. Given enough a-priori information on $\log |T|$, this is simply a matter of applying the appropriate form of the residue theorem. However, in our setting such a-priori information is not freely available, and instead we want to obtain it as a conclusion of our results (namely in our trace formulas). This section is devoted to considering such issues in an uniform fashion. We will phrase our results in terms of a nonnegative subharmonic function G in the upper half-space. One should think of G as $\mp \log |T|$, which is harmonic in the defocusing case but may have singularities at the poles z_k of T in the focusing case,

$$-\Delta \log |T| = \sum m_k \delta z_k \geq 0$$

where m_k denote the corresponding multiplicities. We begin with a first result which provides an integral representation for such functions.

Lemma 4.1. *a) Let G be a nonnegative harmonic function in the upper half-space, which is bounded on the positive imaginary axis. Then it has a trace on the real line, which is a locally finite nonnegative measure μ. Furthermore, we can represent G via the Poisson kernel as*

$$G(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\operatorname{Im} z}{|z-\xi|^2} d\mu(\xi) \tag{4.1}$$

b) Similarly, suppose G is a nonnegative subharmonic function in the upper half-space, which is bounded on the positive imaginary axis. Then it has a trace on the real line, which is a locally finite nonnegative measure μ, and $\nu = -\Delta G$ is a nonnegative locally finite measure in the upper half-plane so that

$$G(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\operatorname{Im} z}{|z-\xi|^2} d\mu(\xi) + \frac{1}{2\pi} \int_H \log \left| \frac{z-\bar{z}_0}{z-z_0} \right| d\nu(z_0) \tag{4.2}$$

Proof. a) The trace of nonnegative harmonic functions are measures, see [2]. For large R denote

$$G_R(z) = \frac{1}{\pi} \int_{-R}^{R} \frac{\operatorname{Im} z}{|z-\xi|^2} d\mu$$

and let G_∞ be the expression on the right in (4.1). Then G_R is a harmonic function in the upper half-space which decays at infinity. Hence $G - G_R$ is nonnegative at infinity in the limit. It is also nonnegative on the real axis, therefore by the maximum principle we get $G \geq G_R$. Letting $R \to \infty$, by Fatou's lemma we get

$$\lim_{R\to\infty} G_R = G_\infty \leq G$$

Now consider the difference $H = G - G_\infty$, which is nonnegative, harmonic, vanishes on the real line and is bounded on the purely imaginary axis. Applying the Harnack inequality (see [15]) we conclude that H is bounded in the upper half-plane. Since it vanishes on the real line, the odd extension of H is a global bounded harmonic function, and thus constant. Hence we get $H = 0$.

b) The negative Laplacian of a subharmonic function is a measure. We denote this measure by ν. For $K \subset \{z : \operatorname{Im} z > 0\}$ compact we define

$$G_K = \frac{1}{2\pi} \int_K \ln \left| \frac{z - \bar{z}_0}{z - z_0} \right| d\nu(z_0).$$

By the maximum principle $G_K \leq G$. Similarly, if $K_0 \subset K_1$ then $G_{K_0} \leq G_{K_1}$ and

$$G_\infty = \sup_K G_K \leq G$$

and it satisfies

$$-\Delta G_\infty = \nu.$$

Now we apply part a) to the nonnegative harmonic function $G - G_\infty$. □

Consider a subharmonic function G as above and $j \geq 0$. Then we define the expressions

$$G_{2j} = \frac{1}{\pi} \int_{\mathbb{R}} \xi^{2j} d\mu + \frac{1}{2\pi} \int_U \frac{1}{2N+1} \operatorname{Im} z^{2j+1} d\nu$$

which can formally be interpreted as the coefficients of the formal expansion of G at infinity,

$$G(i\tau) \approx \sum_{j=0}^{\infty} (-1)^j \tau^{-2j-1} G_{2j}$$

We make this correspondence rigurous in the following:

Lemma 4.2. *Assume that the support of ν is contained in a strip $0 < \operatorname{Im} z < c$ and that on the imaginary axis the function G admits the finite expansion*

$$G(i\tau) = \sum_{j=0}^{N-1} (-1)^j \tau^{-2j-1} \tilde{G}_{2j} + O(\tau^{-2N-1})$$

for some real constants $\tilde{G}_{2j}$. Then the measures $(1+\xi^2)^N \mu$, $\operatorname{Im} z(1+|z|^2)^N \nu$ are finite, and we have

$$G(i\tau) = \sum_{j=0}^{N} (-1)^j \tau^{-2j-1} G_{2j} + o(\tau^{-2N-1})$$

Proof. We argue by induction on N. For $N = 0$ we use the representation (4.2) to compute

$$\lim_{\tau \to \infty} \tau G(i\tau) = G_0$$

where the integrand is positive and increasing in τ so the limit always exists. For the induction step, we assume that the result holds for $N-1$ and prove it for N. To achieve this we consider the difference

$$G_{\geq N}(i\tau) = G(i\tau) - \sum_{j=0}^{N-1} (-1)^j \tau^{-2j-1} G_{2j}$$

which can be represented as

$$(-1)^N G_{\geq N}(i\tau) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\xi^{2N}}{\tau^{2N-1}(\tau^2 + \xi^2)} d\mu(\xi) + \frac{1}{2\pi} \int_H \Xi_{\geq N}(\tau, z) d\nu(z)$$

where

$$\Xi_{\geq N}(z) = \operatorname{Im} \int_0^z \frac{\zeta^{2N}}{\tau^{2N-1}(\tau^2 + \zeta^2)} d\zeta$$

Then the integrand is nonnegative outside a compact set, so using again monotonicity we compute the limit

$$\lim_{\tau\to\infty} \tau^{2N+1}(-1)^N G_{\geq N}(i\tau) = G_{2N}$$

which may be either finite or $+\infty$. By hypothesis this is finite and the conclusion of the lemma follows. □

Next we turn our attention to noninteger s, for which we define the energy type quantities

$$E_s(G) = 4\sin(\pi s)\int_1^\infty (\tau^2-1)^s\Big[-G(i\tau)+\sum_{j=0}^N (-1)^j G_{2j}\tau^{-2j-1}\Big]d\tau + 2\pi\sum_{j=0}^N \binom{s}{j} G_{2j}$$

For these we have the following:

Lemma 4.3. *Let G be a subharmonic function in the upper half-space, with the associated measures μ, ν as above, and $N \leq s < N+1$ with the following properties:*

(1) *The quantities G_{2j}, $0 \leq j \leq N$ are finite.*
(2) *The support of ν is contained in the region $\operatorname{Im} z \lesssim (1+\operatorname{Re} z)^{-2s}$ for $s \leq 0$, respectively $\operatorname{Im} z \lesssim 1$ for $s \geq 0$.*

Then the following trace formula holds:

$$E_s(G) = \int_{\mathbb{R}} (1+\xi^2)^s d\mu + \int \Xi_s d\nu \tag{4.3}$$

in the sense that the equality holds whenever the one side is finite. If one of the sides is finite then the integrand of $E_s(G)$ is bounded by $C\tau^{-2s-1}$.

Proof. We use the representation of G in (4.2) and expansions of the integrands to rewrite $E_s(G)$ by introducing summands to the integrand to get better decay

$$\frac{1}{\pi}\int_{\mathbb{R}} \frac{\tau}{\tau^2+\xi^2} - \sum_{j=0}^N (-1)^j \xi^{2j}\tau^{-2j-1} d\mu(\xi) + \frac{1}{2\pi}\int \log\left|\frac{i\tau-\bar z}{i\tau - z}\right| + \sum_{j=0}^N \frac{1}{2j+1}\operatorname{Im} z^{2j+1}\tau^{-2j-1} d\nu(z)$$
$$= \frac{1}{\pi}\int_{\mathbb{R}} \frac{(-1)^N\xi^{2N+2}}{(\tau^2+\xi^2)\tau^{2N+1}} d\mu(\xi) + \frac{1}{2\pi}\int (-1)^N \Xi_{\geq N+1}(z) d\nu(z)$$

where

$$\Xi_{\geq N+1}(z) = \operatorname{Im}\int_0^z \frac{\zeta^{2N+2}}{(\tau^2+\zeta^2)\tau^{2N+1}} d\zeta$$

is independent of the contour of integration.

We verify the claimed identity first for Dirac measures $\mu = \delta_\xi$:

$$4\sin(\pi s)\int_1^\infty (\tau^2-1)^s \frac{\xi^{2N+2}}{(\tau^2+\xi^2)\tau^{2N+1}} d\tau + \sum_{j=0}^N \binom{s}{j}\xi^{2j} = (1+\xi^2)^s$$

which is nothing but (the real part of) Cauchy's formula applied to the integral, which we consider as contour integral from $i\infty - 0$ to i and then to $i\infty + 0$. We move this contour of integration to the real line, Taylor expand $(z^2+1)^s$ at 0, and evaluate

$$\frac{1}{\pi}\int_{-\infty}^\infty (1+(\eta+i0)^2)^s \frac{\xi^{2N+2}}{(\eta+i0-\xi^2)^2(\eta+i0)^{2N+1}} d\tau = (1+\xi^2)^s - \sum_{j=0}^N \binom{s}{j}\xi^{2j}.$$

The same argument applies for the case when $\nu = \delta_z$ is a Dirac mass. Then we need to verify the relation

$$4\sin(\pi s)\int_1^\infty (\tau^2-1)^s \Xi_{\geq N+1}(z)d\tau + \sum_{j=0}^N \binom{s}{j}\Xi_j(z) = 2\Xi_s(z)$$

Both sides vanish at $z=0$, so it suffices to verify that their differentials are equal, i.e. that

$$4\sin(\pi s)\int_1^\infty (\tau^2-1)^s \frac{\xi^{2N+2}}{(\tau^2+z^2)\tau^{2N+1}}d\tau + \sum_{j=0}^N \binom{s}{j} z^{2j} = 2(1+z^2)^s$$

This is the same residue theorem computation as above, except that at z we now get a full residue rather than the prior half residue.

We first establish the desired bound (4.3) if both measures μ and ν are compactly supported. Then all integrals in (4.3) are absolutely convergent, and it suffices to consider the case when both μ and ν are Dirac masses. Hence, to conclude the proof it suffices to assume that both measures are supported outside a compact set. Then we can use the localization of ν to conclude that the contribution of ν to both the left hand side and the right hand side of (4.3) are nonnegative. By Fatou's lemma the two expressions are either both finite and equal, or both infinite. □

The leading term in both $T^{-1}-1$ and $\ln T$ away from the real axis is $T_2(z)$. It is by far the simplest and most important to analyze, and provides the quadratic term in all our conserved energies. All the techniques for its study will remain relevant for the higher order terms. In the focusing case $v=-u$ so only the sign of T_2 changes. There is however no sign change in $\tilde{T}_2$ since we also replace $-\ln|T|$ by $\ln|T|$.

In both cases we study

$$T_2(z) = \int_{x<y} e^{2iz(y-x)}u(y)\overline{v(x)}dxdy.$$

We choose an approach which generalizes to the higher order terms. For definiteness we work with the unitary Fourier transform

$$\hat{u}(\xi) = (2\pi)^{-\frac{1}{2}}\int u(x)e^{-ix\xi}dx$$

and use Fourier inversion

$$\begin{aligned}
T_2(z) =& \frac{1}{2\pi}\int_{x<y} e^{2iz(y-x)-i(x\xi-y\eta)}u(\eta)\overline{\hat{v}(\xi)}d\xi d\eta \\
=& -\frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}}\frac{1}{2iz+i\xi}e^{iy(\eta-\xi)}\hat{u}(\eta)\overline{\hat{v}(\xi)}dyd\xi d\eta \\
=& -\frac{1}{2\pi i}\int_{\mathbb{R}}\frac{1}{2z+\xi}|\hat{u}(\xi)|^2 d\xi \\
=& -\frac{1}{2\pi i}\sum_{j=0}^N(-1)^j\int \xi^j|\hat{u}(\xi)|^2d\xi(2z)^{-j-1} + (-2z)^{-N}\int\frac{\xi^{N+1}}{(2z-\xi)}|\hat{u}(\xi)|^2d\xi.
\end{aligned} \tag{4.4}$$

We denote $f(\xi) = |\hat{u}(-\xi)|^2$. Then

$$T_2(z) = -\frac{1}{2\pi i}\int_{\mathbb{R}}\frac{1}{2z-\xi}f(\xi)d\xi.$$

and we apply Lemma 4.2 and 4.3.

Proposition 4.4. *a) For z in the upper half-plane we have*

$$|T_2(z)| \leq \frac{1}{\pi}\int_{\mathbb{R}}\frac{\operatorname{Im} z}{|\operatorname{Im} z|^2+|2\operatorname{Re} z+\xi|^2}|\hat{u}(\xi)|^2d\xi. \tag{4.5}$$

Further, for $n \geq 0$ let

$$H_{j,2} = (-1)^j \int_{\mathbb{R}} \xi^j |\hat{u}(\xi)|^2 d\xi = \begin{cases} \int_{\mathbb{R}} |u^{(k)}|^2 dx & \text{if } j = 2k \\ \operatorname{Im} \int_{\mathbb{R}} u^{(k+1)} \overline{u^{(k)}}\, dx & \text{if } j = 2k+1. \end{cases} \tag{4.6}$$

then

$$\left| T_2(z) + \frac{1}{2\pi i} \sum_{j=0}^{n-1} H_{j,2} (2z)^{-1-j} \right| \leq \frac{1}{|2z|^{n+1}\pi} \int_{\mathbb{R}} |\xi|^n \frac{\operatorname{Im} z (2|z| + 2n|2z+\xi|)}{|\operatorname{Im} z|^2 + |2\operatorname{Re} z + \xi|^2} |\hat{u}(\xi)|^2 d\xi \tag{4.7}$$

b) If $s \notin \mathbb{N}$, $s > -\frac{1}{2}$ then

$$\|u\|_{H^s}^2 = 4\sin \pi s \int_1^\infty (\tau^2 - 1)^s \Big[\operatorname{Re} T_2\Big(\frac{\tau i}{2}\Big) - \frac{1}{2\pi} \sum_{j=0}^{[s]} H_{2j,2} (-1)^j \tau^{-2j-1} \Big] d\tau + \sum_{j=0}^{[s]} \binom{s}{j} H_{2j,2}. \tag{4.8}$$

Proof. Only the bound (4.7) and its special case (4.5) remain to be proven. They are immediate consequences of (4.4). □

5. Bounding the iterative integrals T_{2j}

Here we consider the question of estimating the integrals T_{2j}, with the goal of establishing the convergence of the formal series for T. Since our energies are expressed in terms of the values of $T(z)$ on the positive imaginary axis, we will focus on this case, but also comment on the region of validity of the expansion in the upper half-plane.

The natural setting here is to work at scaling. Naively one might start with u and v in the space $\dot{H}^{-\frac{1}{2}}$. However, this does not seem to work so well, so instead we use its cousins DU^2 and DV^2. Appendix B provides a self contained introduction for these spaces, including some new results.

We list the properties which we need in this section. We use a suggestive formal notation which will be made precise (and proven) in Appendix B. For simplicity in the discussion below we consider spaces of functions defined on $\mathbb{R}$, but similar properties apply in any bounded or unbounded interval $(a, b) \subset \mathbb{R}$.

(1) U^2 and V^2 have the same scaling properties as BV and L^∞ and satisfy

$$U^2 \subset V^2 \subset \mathcal{R}. \tag{5.1}$$

where $\mathcal{R}$ is the set of ruled functions, .i.e. functions with limits from the left and the right everywhere including one sided limits at the endpoints. Ruled functions are bounded. Further, we have the embedding relation with the homogeneous Besov spaces $\dot{B}^{-\frac{1}{2}}_{2,q}$, (see Corollary B.21)

$$L^1 + \dot{B}^{-\frac{1}{2}}_{2,1} \subset DU^2 \subset DV^2 \subset \dot{B}^{-\frac{1}{2}}_{2,\infty} \tag{5.2}$$

(2) DU^2 and DV^2 are the spaces of distributional derivatives of U^2 and V^2 functions. In particular $u \in U^2$ if and only if $\lim_{t\to-\infty} u(t) = 0$ and $u' \in DU^2$, see definition B.16. Then

$$\|u\|_{U^p} = \|u'\|_{DU^p}$$

which is less than infinity if and only if $u \in U^p$. Moreover, if v is left continuous and $\lim_{t\to\infty} v(t) = 0$ then

$$\|v\|_{V^p} = \|v'\|_{DV^p}.$$

(3) The bilinear estimates

$$\begin{aligned} \|vu\|_{DU^2} &\leq 2\|v\|_{V^2}\|u\|_{DU^2} \\ \|vu\|_{DV^2} &\leq \|v\|_{DV^2}\|u\|_{U^2} \end{aligned} \tag{5.3}$$

hold (see Definition B.15 and the discussion thereafter).

To iteratively solve the system (3.1) we define the one step operator

$$\phi \to L\phi(t) = \int_{x<y<t} e^{iz(y-x)}u(x)\overline{v(y)}\phi(x)dydx.$$

Our first bound for L is as follows:

Lemma 5.1. *Let* $\operatorname{Im} z > 0$. *Then the operator* L *satisfies*

$$\|L\|_{V^2\to U^2} \leq 8\|e^{-i\operatorname{Re} zx}v\|_{DU^2}\|e^{-i\operatorname{Re} zx}u\|_{DU^2}.$$

Proof. It suffices to consider $z = i$, by rescaling and including the oscillatory factor $e^{i\operatorname{Re} zx}$ into u. Then, applying the second and third property above several times, with $\eta = \chi_{t<0}e^t$

$$\begin{aligned} \|L\psi\|_{U^2} &= \|(L\psi)'\|_{DU^2} \\ &= \left\|\overline{v(y)}\int_{-\infty}^{y} e^{-(y-x)}u(x)\psi(x)dx\right\|_{DU^2} \\ &\leq 2\|v\|_{DU^2}\|\eta * (u\psi)\|_{V^2} \\ &\leq 4\|v\|_{DU^2}\|\eta * (u\psi)\|_{U^2} \\ &\leq 4\|v\|_{DU^2}\|u\psi\|_{DU^2} \\ &\leq 8\|v\|_{DU^2}\|u\|_{DU^2}\|\psi\|_{V^2}. \end{aligned}$$

In addition we have used

$$\|\eta * w\|_{U^2} = \|(\eta * w)'\|_{DU^2} = \|\eta' * w\|_{DU^2} \leq \|w\|_{DU^2}$$

□

This bound is quite sharp on the real line, but as we move z into the upper half-space we can do better than this. This is done using some more localized versions of the DU^2 spaces.

Definition 5.2. *Given a frequency scale* $\sigma > 0$ *we define*

$$\|u\|_{l^p_\sigma DU^2} = \left\|\|\chi_{[k/\sigma,(k+1)/\sigma]}u\|_{DU^2}\right\|_{l^p_k}$$

and similarly for U^2,

$$\|v\|_{l^p_\sigma U^2} = \left\|\|\chi_{[k/\sigma,(k+1)/\sigma]}u\|_{U^2}\right\|_{l^p_k}.$$

where, for an interval I, χ_I *is a smooth cutoff associated to the interval* I, *and the above functions* $\chi_{[k/\sigma,(k+1)/\sigma]}$ *form a partition of unit.*

These spaces are translation invariant, and the norm of a translated function is at most a fixed constant time the norm of the original function. Connecting the two spaces we have the following

Lemma 5.3. *We have*

$$\|u\|_{l^p_\sigma U^2} \lesssim \|\partial u\|_{l^p_\sigma DU^2} + \sigma\|u\|_{l^p_\sigma DU^2} \tag{5.4}$$

Proof. Since $\|u\|_{U^2} = \|\partial u\|_{DU^2}$, we can write

$$\begin{aligned} \|\chi_{[k/\sigma,(k+1)/\sigma]}u\|_{U^2} &\lesssim \|\chi_{[k/\sigma,(k+1)/\sigma]}\partial u\|_{DU^2} + \|\partial_x\chi_{[k/\sigma,(k+1)/\sigma]}u\|_{DU^2} \\ &\lesssim \|\chi_{[k/\sigma,(k+1)/\sigma]}\partial u\|_{DU^2} + \sigma\|\tilde\chi_{[k/\sigma,(k+1)/\sigma]}u\|_{DU^2} \end{aligned}$$

where $\tilde\chi$ is a cutoff selecting a slightly larger set. Then the conclusion follows after l^p summation. □

In the case $p=2$ we have an alternative simple characterization of these spaces:

Lemma 5.4. *The spaces $l^2_\sigma DU^2$ can be characterized as follows*

$$l^2_\sigma DU^2 = DU^2 + \sigma^{\frac{1}{2}} L^2. \tag{5.5}$$

We postpone the proof for the end of Appendix B. As a corollary we have::

Corollary 5.5. *The following embeddings hold*

$$B^{-\frac{1}{2}}_{2,1} \subset l^2_1 DU^2 \subset B^{-\frac{1}{2}}_{2,\infty}. \tag{5.6}$$

Proof. This is an immediate consequence of Lemma 5.4 and the analogous homogeneous result in Corollary B.21. □

On the other hand, for $p > 2$ we will use the following embeddings:

Lemma 5.6. *For $p \geq 2$ we have the bounds*

$$\|u\|_{l^p_\tau DU^2} \lesssim \tau^{\frac{1}{p}-1} \|u\|_{\dot H^{\frac{1}{2}-\frac{1}{p}}}, \tag{5.7}$$

Also if $0 \leq \tau_1 \leq \tau_2$ then

$$\|u\|_{l^p_{\tau_2} DU^2} \lesssim \|u\|_{l^p_{\tau_1} DU^2}. \tag{5.8}$$

Proof. The inequality (5.7) follows via the Sobolev embedding

$$\|u\|_{l^p_\tau DU^2} \lesssim \tau^{\frac{1}{p}-1} \|u\|_{l^p_\tau L^p} = \tau^{\frac{1}{p}-1} \|u\|_{L^p} \lesssim \tau^{\frac{1}{p}-1} \|u\|_{\dot H^{\frac{1}{2}-\frac{1}{p}}}.$$

The second inequality is a consequence of Lemma 5.4. □

We also need to understand the effect of phase shifts:

Lemma 5.7. *Suppose that $|\xi| \lesssim \sigma$. Then*

$$\|e^{ix\xi} u\|_{l^2_\sigma DU^2} \lesssim \|u\|_{l^2_\sigma DU^2}, \qquad \|e^{ix\xi} u\|_{l^2_\sigma U^2} \lesssim \|u\|_{l^2_\sigma U^2}. \tag{5.9}$$

Proof. Since $|\xi| \lesssim \sigma$, it follows that the function $e^{ix\xi}$ is uniformly smooth and bounded on the supports of the cutoff functions $\chi_{[k/\sigma,(k+1)/\sigma]}$ in Definition 5.2. The desired conclusion immediately follows. □

As a consequence we have

Corollary 5.8. *Let $\sigma > 0$ and $\xi \in \mathbb{R}$. Then the following estimate holds*

$$\|e^{i\xi x} u\|_{l^2_\sigma DU^2} \lesssim \left(\frac{1+\sigma+|\xi|}{\sigma} \right)^{\frac{1}{2}} \|u\|_{l^2_1 DU^2}.$$

Proof. If $|\xi| \leq \sigma$ then we by the previous Lemma we have

$$\|e^{i\xi x} u\|_{l^2_\sigma DU^2} \lesssim \|u\|_{l^2_\sigma DU^2} \lesssim \max\{1, \sigma^{-\frac{1}{2}}\} \|u\|_{l^2_1 DU^2}.$$

Else we use $|\xi|$ as an intermediate threshold,

$$\begin{aligned} \|e^{i\xi x} u\|_{l^2_\sigma DU^2} &\lesssim (|\xi|/\sigma)^{\frac{1}{2}} \|e^{i\xi x} u\|_{l^2_{|\xi|} DU^2} \lesssim (|\xi|/\sigma)^{\frac{1}{2}} \|u\|_{l^2_{|\xi|} DU^2} \\ &\lesssim (|\xi|/\sigma)^{\frac{1}{2}} \max\{1, |\xi|^{-\frac{1}{2}}\} \|u\|_{l^2_1 DU^2}. \end{aligned}$$

□

Using the above spatially localized norms we can prove a stronger form of Lemma 5.1:

Lemma 5.9. *Let* $\operatorname{Im} z > 0$. *Then the operator* L *satisfies*

$$\|L\|_{U^2\to U^2} \lesssim \|e^{-i\operatorname{Re} zx}u\|_{l^2_{\operatorname{Im} z}DU^2}\|e^{-i\operatorname{Re} zx}v\|_{l^2_{\operatorname{Im} z}DU^2}.$$

Proof. As above is suffices to prove this for $z = i$. We repeat the previous argument, but with some obvious changes:

$$\|L\psi\|_{U^2} \lesssim \|v\|_{l^2DU^2}\|\eta * (u\psi)\|_{l^2U^2} \lesssim \|v\|_{l^2DU^2}\|u\psi\|_{l^2DU^2} \lesssim \|v\|_{l^2DU^2}\|u\|_{l^2DU^2}\|\psi\|_{U^2}.$$

□

The main estimate of this section follows:

Proposition 5.10. *The iterated integrals* $T_{2j}(z)$ *and* $\tilde{T}_{2j}$ *satisfy the following bounds in the upper half-plane:*

$$|T_{2j}(z)| + |T_{2j}(z)| \leq \left(c\|e^{i\operatorname{Re} zx}u\|_{l^2_{\operatorname{Im} z}DU^2}\right)^j \left(c\|e^{-i\operatorname{Re} zx}v\|_{l^2_{\operatorname{Im} z}DU^2}\right)^j \tag{5.10}$$

Proof of Proposition 5.10. We begin with the bound for T_{2j}. The first component of the solution to (3.1) is defined by

$$\psi_1(t) = e^{izt}\sum_{j=0}^{\infty}(L^j 1)(t)$$

provided that this series converges uniformly. Thus, the transmission coefficient T is given by

$$T^{-1}(z) = \sum_{j=0}^{\infty}\lim_{t\to\infty}(L^j(1))(t).$$

Functions in U^2 have left and right limits everywhere which together with Lemma 5.9 completes the proof of the T_{2j} part of (5.10). To switch to $\tilde{T}_{2j}$ we begin with the relation

$$\sum_{j=1}^{\infty}\tilde{T}_{2j} = \log\left(1 + \sum_{j=1}^{\infty}T_{2j}\right)$$

Recalling that both T_{2j} and $\tilde{T}_{2j}$ are homogeneous multilinear forms of degree $2j$, replacing u by zu and $v = \bar{u}$ by zv we must also have the formal series relation

$$\sum_{j=1}^{\infty}\zeta^j\tilde{T}_{2j} = \log\left(1 + \sum_{j=1}^{\infty}\zeta^j T_{2j}\right)$$

We can use (5.10) to bound the size of the coefficients for the series on the right. To achieve that we introduce a partial order "$\preceq$" on holomorphic functions near zero, where $g \preceq h$ means that the absolute value of every coefficient of the Taylor series of g at zero is bounded by the corresponding coefficient of the Taylor series of h at zero. This order is easily seen to be compatible with the addition, multiplication and composition of holomorphic functions.

In particular we note that

$$\ln(1+\zeta) \preceq \frac{\zeta}{1-\zeta} =: f(\zeta)$$

Bounding the coefficients in the Taylor series for the logarithm by 1 and T_{2j} as in (5.10), it follows that the series on the left is dominated by

$$\sum_{j=1}^{\infty}\zeta^j\tilde{T}_{2j} \preceq f\circ f(C\zeta).$$

and

$$C = c^2 \|e^{i \operatorname{Re} zx} u\|_{l^2_{\operatorname{Im} z} DU^2} \|e^{-i \operatorname{Re} zx} v\|_{l^2_{\operatorname{Im} z} DU^2}.$$

Computing the Taylor series for the function on the right we obtain

$$\sum_{j=1}^{\infty} \zeta^j \tilde{T}_{2j} \preceq \sum_{j=1}^{\infty} 2^{j-1} C^j \zeta^j$$

which yields the inequality

$$|\tilde{T}_{2j}| \leq 2^{j-1} C^j$$

thus proving the $\tilde{T}_{2j}$ part of (5.10) with c replaced by $2c$.

□

As an immediate corollary to Proposition 5.10 we have:

Corollary 5.11. *Assume that*

$$\|(u, v)\|_{l^2_1 DU^2} \ll 1. \tag{5.11}$$

Then the formal series for T^{-1} and for $\ln T$ converge uniformly on the half-line $[i, i\infty)$, and more generally in the region $\{\operatorname{Im} z \geq 1 + |\operatorname{Re} z|\}$.

The latter statement follows from Corollary 5.8, which yields smallness for $z = \xi + i\tau$ in the above region.

Furthermore, we note that the function T^{-1} is in effect well defined in the entire upper half-plane:

Corollary 5.12. *Assume that $(u, v) \in l^2_1 DU^2$. In both the defocusing and the focusing case the map T^{-1} is well defined as a holomorphic function in the whole upper half plane.*

Proof. Consider the ODE (3.1) with $z = \xi + i\tau$ and $\tau > 0$. By Corollary 5.8 we have

$$\|e^{ix\xi} u\|_{l^2_\tau DU^2} \lesssim \left(\frac{1 + |\xi| + \tau}{\tau}\right)^{\frac{1}{2}} \|u\|_{l^2 DU^2}.$$

Here we cannot use the argument in Proposition 5.10 because we lack smallness. To remedy this we partition the real line into three intervals

$$\mathbb{R} = (-\infty, x_0] \cup (x_0, x_1) \cup (x_1, \infty)$$

so that on the first and the last interval we have smallness for $\|e^{ix\xi} u\|_{l^2_\tau DU^2}$. In particular we can apply the argument in Proposition 5.10 in order to solve (3.1) up to x_0. From x_0 we continue by solving the ODE up to x_1, and from there on we repeat the previous argument. We obtain a global solution (ψ_1, ψ_2) for (3.1) with $\psi_1 \in U^2$ and $\psi_2 \in DU^2$, and the inverse transmission coefficient $T^{-1}(z)$ is obtained as the limit of ψ_1 at infinity. Holomorphy in z is obvious. □

We note that T^{-1} can have no zeroes in the region of applicability of Corollary 5.11. However, in the focusing case T^{-1} may have zeroes in the upper half-plane. If $T^{-1}(z) = 0$ then z is an eigenvalue of the Lax operator. Eigenvalues are isolated since $T^{-1}(z)$ is analytic. The geometric multiplicity is always one, which is a consequence of the structure of the Jost solutions. The order of the zero of T^{-1} is known to be the algebraic multiplicity of the eigenvalue. This can be seen using the Backlund transform, and is beyond our goals in here.

The next step is to assume in addition that $u, v \in H^s$, in which case we expect that there is additional decay for T_{2j}. Precisely, we have

Proposition 5.13. *Assume that $u, v \in H^s$. Then for s in the range*

$$-\frac{1}{2} < s \leq \frac{j-1}{2} \tag{5.12}$$

we have the pointwise bounds

$$|T_{2j}(i\tau)| + |\tilde{T}_{2j}(i\tau)| \leq \left(1 + \frac{1}{2s+1}\right) \tau^{-2s-1} \|(u,v)\|^2_{H^s} (c\|(u,v)\|_{l^2_1 DU^2})^{2j-2}, \tag{5.13}$$

as well as the integrated bound

$$\int_1^\infty \tau^{2s} |T_{2j}(i\tau)| + |\tilde{T}_{2j}(i\tau)| d\tau \leq \left(1 + \frac{1}{j-1-2s} + \frac{1}{(2s+1)^2}\right) \|(u,v)\|^2_{H^s} (c\|(u,v)\|_{l^2_1 DU^2})^{2j-2}. \tag{5.14}$$

Here c is a large universal constant, and the above bounds are uniform in j and s.

Proof of Proposition 5.13. The starting point for this proof is the bound (5.10), so it makes no difference whether we work with T_{2j} or $\tilde{T}_{2j}$; we choose the former. Greek letters λ and μ are powers of 2. Sums $\sum_\lambda \dots$ are understood as $\sum_{j=0,\lambda=2^j}^\infty \dots$. We define u_λ by the Fourier multiplication by a characteristic function $\hat{u}_1 = \chi_{|\xi|<1}\hat{u}$, $\hat{u}_\lambda = \chi_{\lambda \leq |\xi| < 2\lambda} \hat{u}$, $\hat{u}_{<\lambda} = \chi_{|\xi|<\lambda}\hat{u}$ and by an abuse of notation we write $(u,v)_\lambda = (u_\lambda, v_\lambda)$. Then $u = \sum_\lambda u_\lambda$. We expand the functions in $\tilde{T}_{2j}$ and recall that we only consider $\tau \geq 1$. We separate the two highest dyadic frequencies, which we denote by $\lambda_1 \geq \lambda_2 \geq 1$. Each such pair occurs $O(j^2)$ times, which is subexponential and thus can be neglected. We carry out the summation in the other terms up to frequency λ_2. Using Proposition 5.10 we obtain the bound

$$|T_{2j}(i\tau)| \lesssim \sum_{\lambda_1 \geq \lambda_2} \|(u,v)_{\lambda_1}\|_{l^2_\tau DU^2} \|(u,v)_{\lambda_2}\|_{l^2_\tau DU^2} \|(u,v)_{\leq \lambda_2}\|^{2j-2}_{l^2_\tau DU^2}.$$

To prove the claim for T_{2j} we will bound the summands on the right hand side and carry out the summation. The first two factors we estimate in the H^s norm using Lemma 5.4. If $1 < \lambda \leq \tau$ we estimate

$$\|u_\lambda\|_{l^2_\tau DU^2} \lesssim \tau^{-1/2} \|u_\lambda\|_{L^2} \lesssim \lambda^{-s} \tau^{-\frac{1}{2}} \|u\|_{H^s},$$

and if $\lambda \geq \tau$ we also use Corollary B.21 in the form (5.2)

$$\|u_\lambda\|_{l^2_\tau DU^2} \lesssim \|u_\lambda\|_{DU^2} \lesssim \lambda^{-1/2} \|u_\lambda\|_{L^2} \leq \lambda^{-s-\frac{1}{2}} \|u\|_{H^s}.$$

The remaining factors we estimate in terms of the $l^2_1 DU^2$ norm, using (5.8) if $\lambda \geq \tau$,

$$\|u_{<\lambda}\|_{l^2_\tau DU^2} \lesssim \|u_{<\lambda}\|_{l^2_1 DU^2},$$

and if $\lambda < \tau$ it follows from Lemma 5.4 and (5.5) that

$$\|u_{<\lambda}\|_{l^2_\tau DU^2} \lesssim \tau^{-1/2} \|u_{<\lambda}\|_{L^2} \lesssim (\lambda/\tau)^{\frac{1}{2}} \|u_{<\lambda}\|_{l^2_1 DU^2}.$$

We obtain a bound of the form

$$|T_{2j}(i\tau)| \lesssim \tau^{-2s+1} \sum_{\lambda_1 \geq \lambda_2} C(\tau, \lambda_1, \lambda_2) \|(u,v)_{\lambda_1}\|_{H^s} \|(u, v_{\lambda_2})\|_{H^s} \|(u,v)\|^{2j-2}_{l^2_1 DU^2} \tag{5.15}$$

where the constant $C(\tau,\lambda_1,\lambda_2)$ depends on the relative position of the entries as follows:

$$C(\tau,\lambda_1,\lambda_2)=\begin{cases}\left(\frac{\lambda_1}{\tau}\right)^{-2s+j-1}\left(\frac{\lambda_2}{\lambda_1}\right)^{-s+j-\frac{1}{2}} & \lambda_2\le\lambda_1\le\tau\\ \left(\frac{\lambda_1}{\tau}\right)^{-(s+\frac{1}{2})}\left(\frac{\lambda_2}{\tau}\right)^{-s+j-1} & \lambda_2\le\tau\le\lambda_1\\ \left(\frac{\lambda_1}{\tau}\right)^{-(s+\frac{1}{2})}\left(\frac{\lambda_2}{\tau}\right)^{-(s+\frac{1}{2})} & \tau\le\lambda_2\le\lambda_1\end{cases}$$

More precisely, by an application of Schur's lemma and the Cauchy-Schwarz inequality

$$\tau^{2s+1}|T_{2j}(i\tau)|\lesssim\sqrt{A_pB_p}\|(u,v)\|_{H^s}^2\|(u,v)\|_{l_1^2(DU^2)}^{2j-2}$$

$$\int_1^\infty\tau^{2s}|T_{2j}(i\tau)|d\tau\lesssim A_i\|(u,v)\|_{H^s}^2\|(u,v)\|_{l_1^2(DU^2)}^{2j-2} \tag{5.16}$$

with

$$A_p=\sup_{\tau,\lambda_1}\sum_{\lambda_2}C(\tau,\lambda_1,\lambda_2),\quad B_p=\sup_{\tau,\lambda_2}\sum_{\lambda_1}C(\tau,\lambda_1,\lambda_2)$$

$$A_i=\max\left\{\sup_{\lambda_1}\sum_{\tau,\lambda_2}C(\tau,\lambda_1,\lambda_2),\sup_{\lambda_2}\sum_{\tau,\lambda_1}C(\tau,\lambda_1,\lambda_2)\right\}.$$

We break the sum in (5.16) up into three sums given by the constraints

$$\lambda_2\le\lambda_1\le\tau,\quad\lambda_2\le\tau\le\lambda_2,\quad\tau\le\lambda_2\le\lambda_1.$$

They are all given by geometric sums. The pointwise bounds are easy consequences and we provide more details for the slightly more involved integrated bounds of (5.14) and bound the A_i up to a multiplicative constant by

$$1+\frac{1}{(j-\frac{1}{2}-s)(j-1-2s)}\qquad\text{for the case }\lambda_1\le\lambda_2\le\tau \tag{5.17}$$

$$1+\frac{1}{(s+\frac{1}{2})(s-\frac{1}{2}-j)}\qquad\text{for the case }\lambda_1\le\tau\le\lambda_2 \tag{5.18}$$

$$1+\frac{1}{(s+\frac{1}{2})^2}\qquad\text{for the case }\tau\le\lambda_1\le\lambda_2. \tag{5.19}$$

This completes the proof. □

We remark that, as a consequence of this last proposition, the proof of our main theorem reduces to considering the terms $\tilde{T}_{2j}$ for $j\le 2s+1$ since we obtain the following corollary.

Corollary 5.14. *If* $\|u\|_{l^2DU^2}\ll 1$ *then we have the bound*

$$\int_1^\infty\tau^{2s}\left|\ln T-\sum_{j=1}^{[2s+1]}\tilde{T}_{2j}(i\tau)\right|d\tau\le c\left(1+\frac{1}{j-1-2s}+\frac{1}{(2s+1)^2}\right)\|(u,v)\|_{\dot{H}^s}^2(c\|(u,v)\|_{l^2DU^2})^{2[2s+1]}. \tag{5.20}$$

In view of the identity for T_2 of the previous section, this bound suffices for the proof of our main theorem in the range $-\frac{1}{2}<s<\frac{1}{2}$. Restating our main theorem in this case, we have

Proposition 5.15. *Let $-\frac{1}{2} < s < \frac{1}{2}$, and $u \in H^s$ with $\|u\|_{l^2 DU^2} \ll 1$. Then $E_s(u)$ is well defined, and*

$$\big|E_s(u) - \|u\|_{H^s}^2\big| \lesssim \|u\|_{H^s}^2 \|u\|_{l_1^2 DU^2}^2. \tag{5.21}$$

We conclude the section with a further discussion of the case $-\frac{1}{2} < s < 0$, which was our original goal in this paper. First, we note that, as a consequence of (5.10), we have the bound

$$|T_{2j}(z)|^{\frac{1}{j}} \leq c\frac{|\operatorname{Re} z|^{-2s}}{\operatorname{Im} z}\|u\|_{H^s}^2$$

Thus, the expansion for $T(z)$ converges for $\operatorname{Im} z \geq (\operatorname{Re} z)^{-2s} c\|u\|_{H^s}^2$. In particular, if smallness is assumed, $c\|u\|_{H^s}^2 \leq 1/2$ then the series converges for $\operatorname{Im} z \geq (\operatorname{Re} z)^{-2s}$, and the following estimate and also the corresponding integrated bound hold:

$$|\ln T(z) - T_2(z)| \leq 2\frac{(\operatorname{Re} z)^{-2s}}{\operatorname{Im} z}\|u\|_{H^s}^2.$$

6. Expansions for the iterative integrals $\tilde{T}_{2j}$

In view of the bounds of the previous section, it remains to separately consider the terms $\tilde{T}_{2j}$ for $j \leq 2s+1$. To avoid the degeneracy at $s = -\frac{1}{2}$ we will harmlessly assume throughout the section that $s \geq 0$. The implicit constants in the present section depend on j, in contrast to the previous section.

It is crucial in this section that we will take advantage of the fact, proved in Theorem 3.3, that the iterated integrals in T_{2j} have fully connected symbols. Our first result here is as follows:

Proposition 6.1. *The iterated integrals $\tilde{T}_{2j}$ satisfy the following bounds:*

$$|\tilde{T}_{2j}(z)| \lesssim \|e^{i\operatorname{Re} zx}u\|_{l^{2j}_{\operatorname{Im} z} DU^2}^{2j} \tag{6.1}$$

As before, the next step is to assume in addition that $u, v \in H^s$, in which case we expect that there is additional decay for $\tilde{T}_{2j}$. Precisely, we have

Proposition 6.2. *Assume that $u, v \in H^s$. Then we have the pointwise bounds*

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-2s-1}\|(u,v)\|_{H^s}^2 \|(u,v)\|_{l^2_\tau DU^2}^{2j-2}, \qquad s \leq j-1 \tag{6.2}$$

as well as the integrated bound

$$\int_1^\infty \tau^{2s}|\tilde{T}_{2j}(i\tau)|ds \lesssim \left(1 + \frac{1}{j-1-s}\right)\|(u,v)\|_{H^s}^2 \|(u,v)\|_{l^2_1 DU^2}^{2j-2}, \qquad 0 \leq s < j-1 \tag{6.3}$$

We remark that, as a consequence of this last proposition, the proof of our main theorem reduces to considering the terms $\tilde{T}_{2j}$ for $0 \leq j \leq s+1$:

Corollary 6.3. *We have the bound*

$$\int_1^\infty \tau^{2s}\Big|\ln T + \|u\|_{H^s}^2 - \sum_{j=2}^{[s]} \tilde{T}_{2j}(i\tau)\Big|d\tau \lesssim \left(\frac{1}{(s+\frac{1}{2})^2} + \frac{1}{[s]+1-s}\right)\|(u,v)\|_{\dot{H}^s}^2 \|(u,v)\|_{l^2 DU^2}^{2[s]}. \tag{6.4}$$

We note that this (and the obvious improvement for $s < 1$ since then there is no term T_4) suffices for the proof of our main theorem in the range $-\frac{1}{2} < s < 1$. We continue with the proof of the above results.

Proof of Proposition 6.1. The proof is a direct consequence of Proposition 5.10, taking into account Theorem 3.3. We localize spatially on the $(\operatorname{Im} z)^{-1}$ scale. The key fact is that the multilinear forms in $\tilde{T}_{2j}$ have connected symbols, therefore their kernels have exponential off-diagonal decay on the $(\operatorname{Im} z)^{-1}$ scale. Thus only the l^{2j} summability is needed, in order to account for the diagonal sum. □

Proof of Proposition 6.2. In order to apply the same method as in the proof of Proposition 5.13 we need to relate the space $l^{2j}_\tau DU^2$ to an L^2 based Sobolev space. For this we use the Sobolev type embedding (5.7) for $p = 2j$ and frequencies less than τ, and simply the DU^2 norm for larger frequencies: If $1 \le \lambda < \tau$

$$\|u_\lambda\|_{l^{2j}_\tau DU^2} \lesssim \tau^{\frac{1}{2j}-1}\|u_\lambda\|_{L^{2j}} \lesssim \left(\frac{\lambda}{\tau}\right)^{1-\frac{1}{2j}} \lambda^{-s-\frac12}\|u\|_{H^s}$$

and

$$\|u_{\le\lambda}\|_{l^{2j}DU^2} \lesssim \tau^{\frac{1}{2j}-1}\|u_{<\lambda}\|_{L^p} \lesssim \tau^{\frac{1}{2j}-1}\lambda^{\frac12-\frac{1}{2j}}\|u_{<\lambda}\|_{L^2} \le \left(\frac{\lambda}{\tau}\right)^{1-\frac{1}{2j}} \|u\|_{l^2_1 DU^2}.$$

If $\lambda \ge \tau$ we obtain

$$\|u_\lambda\|_{l^{2j}_\tau DU^2} \le \|u_\lambda\|_{l^2_\tau DU^2} \lesssim \|u_\lambda\|_{DU^2} \lesssim \lambda^{-s-\frac12}\|u\|_{H^s}$$

and

$$\|u_{\le\lambda}\|_{l^{2j}_\tau DU^2} \lesssim \|u\|_{l^2_1 DU^2}.$$

Using this we obtain again a bound of the form (5.15), but with an improved factor for all frequencies below τ:

$$C(\tau,\lambda_1,\lambda_2) = \begin{cases} \left(\dfrac{\lambda_1}{\tau}\right)^{2(-s+j-1)} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{-s+2j-\frac52+\frac{1}{2j}} & \lambda_2 \le \lambda_1 \le \tau \\ \left(\dfrac{\lambda_1}{\tau}\right)^{-(s+\frac12)} \left(\dfrac{\lambda_2}{\tau}\right)^{-s+2j-\frac52+\frac{1}{2j}} & \lambda_2 \le \tau \le \lambda_1 \\ \left(\dfrac{\lambda_1}{\tau}\right)^{-(s+\frac12)} \left(\dfrac{\lambda_2}{\tau}\right)^{-(s+\frac12)} & \tau \le \lambda_2 \le \lambda_1 \end{cases} \tag{6.5}$$

This allows us to conclude the proof in the same manner as in Proposition 5.13, but with the improved threshold $s < j-1$ instead of $s < (j-1)/2$. Because of their importance later on we give the bounds for the constants A_i, again up to a multiplicative constant.

$$\frac{1}{(2j-\frac52+\frac{1}{2j}-s)(j-1-s)} \qquad \text{for } \lambda_1 \le \lambda_2 \le \tau \tag{6.6}$$

$$\frac{1}{(2j-\frac52+\frac{1}{2j}-s)(s+\frac12)} \qquad \text{for } \lambda_1 \le \tau \le \lambda_2 \tag{6.7}$$

$$\frac{1}{(s+\frac12)^2} \qquad \text{for } \tau \le \lambda_1 \le \lambda_2. \tag{6.8}$$

where $2j - \frac52 + \frac{1}{2j} - s$ is uniformly bounded from below. □

In order to understand the contributions of $\tilde{T}_{2j}$ to the conserved energies E_s in the remaining range $s \ge j-1$ we need to take into account the expansion of $\tilde{T}_{2j}$ at $i\infty$ in powers of z^{-1} in a similar but more complicated fashion as for T_2. Precisely, given Σ a connected symbol of length $2j$, we will consider higher order expansions and bounds for the iterated integral

$$T_\Sigma(z) = \int_\Sigma \prod_{l=1}^{j} e^{2iz(y_l - x_l)} u(x_l) v(y_l) dx_l dy_l.$$

We will do this on the positive imaginary axis $z = i\tau$, $\tau > 0$, though the results can be easily translated to any z in the upper half plane.

For regular enough u (precisely $u \in H^{\frac{1}{2}-\frac{1}{2j}}$) the integral T_σ decays like $|z|^{-2j+1}$ on the positive imaginary axis. Our goal here is to add to this a formal expansion, with errors which have better decay at infinity. Precisely, we have the following:

Proposition 6.4. *For any connected symbol Σ of degree $2j$, the integral $T_\Sigma(z)$ admits a formal expansion*

$$T_\Sigma(z) \approx \sum_{l=0}^{\infty} T_\Sigma^l z^{-(2j-1+l)}$$

where $T_{\Sigma,l}$ are linear combinations of integrals of the form

$$T_\Sigma^l = \sum_{|\alpha|+|\beta|=l} c_{\alpha\beta} \int \prod_{l=1}^{j} \partial^{\alpha_l} u_l \partial^{\beta_l} v_l dx$$

so that the errors in the above partial expansion satisfy the bounds

$$\left| T_\Sigma(z) - \sum_{l=0}^{k} z^{-(2j-1+l)} T_\Sigma^l \right| \lesssim \sum_{k+1 \le |\alpha|+|\beta| \le 2j-1+k}^{\max\{\alpha_l,\beta_l\} \le [\frac{k}{2}]+1} |z|^{-|\alpha|-|\beta|} \prod_k \|\partial^{\alpha_k} u_k\|_{l^{2j}_\tau DU^2} \|\partial^{\beta_k} v_k\|_{l^{2j}_\tau DU^2}. \tag{6.9}$$

on the positive imaginary axis.

We remark that the formal series above is easily obtained by taking a Taylor expansion of the T_Σ integrand. To simplify the notation we expand at the left point x_1 up to a certain order and obtain for multiindices α and β

$$T_\Sigma^{\alpha\beta} = \int_\Sigma \frac{1}{\alpha!\beta!} \prod e^{2iz(y_j - x_j)} \partial^{\alpha_j} u(x_1) \partial_{\beta_j} \overline{v(x_1)} (x_j - x_1)^{\alpha_j} (y_j - x_1)^{\beta_j} dx_j dy_j$$

By an application of Fubini

$$T_\Sigma^{\alpha\beta} = c_{\alpha\beta} (2z)^{1-2j-|\alpha|-|\beta|} \int \frac{1}{\alpha!\beta!} \prod \partial^{\alpha_j} u(x) \partial_{\beta_j} \overline{v(x)} dx$$

where

$$c_{\alpha\beta} = \frac{1}{\alpha!\beta!} \int_{\Sigma, x_1=0} \prod e^{-(x_j - y_j)} x_j^\alpha y_j^\alpha dx_j dy_j$$

However, this is not the way we will proceed in the proof of the proposition.

Proof of Proposition 6.4. Let $j > 1$. For this proof we relabel in a monotone fashion the set

$$\{x_1, y_1, \cdots x_j, y_j\} = \{t_1, \cdots, t_{2j}\}$$

and the functions

$$\{(u, v, \cdots, u, v)\} = \{v_1, \cdots, v_{2n}\}$$

We first consider the t_{2j} dependent part of the integral T_Σ, which we rewrite by an integration by parts as

$$\int_{t_{2j-1}}^{\infty} e^{-\tau t_{2j}} v(t_{2j}) dt_{2j} = -\frac{1}{\tau} e^{-\tau t_{2j-1}} v(t_{2j-1}) - \frac{1}{\tau} \int_{t_{2j-1}}^{\infty} e^{-\tau t_{2j}} v'(t_{2j}) dt_{2j}.$$

We do the same at the left endpoint t_1.

To obtain the expansion of T_Σ up to degree $2j-1+k$ we repeatedly apply this computation. At each step we get a factor of $\dfrac{1}{\tau}$ and two terms,

(1) The boundary term which leaves as with an integral with a dimension lowered by 1, and two functions evaluated at the same point.
(2) A term where we differentiate one of the functions.

We repeat this algorithm until either of the following stopping criteria is fulfilled:

(1) If step (i) is applied $2j-1$ times; then all integration variables are equal. These terms give exactly the coefficients T^l_Σ, $l \le k$, in the expansion of T_Σ up to degree k.
(2) If step (ii) is applied $k+1$ times. These terms are those where the integration variables are not all equal; they are part of the error term which we need to estimate.

We remark on the critical role of the assumption that Σ is connected. This is what guarantees that at each step we are still integrating a decaying exponential. In order to best balance derivatives, we alternately apply the above steps at the left and at the right. Starting on the right, this implies that we do it $\alpha^- = [\frac{k+2}{2}]$ times on the left, respectively $\alpha^+ = k+1-[\frac{k+2}{2}]$ times on the right. To describe a general error term that we need to estimate, we denote by j^-, respectively j^+ the number of times the option (i) is taken on the left, respectively on the right. These must satisfy

$$0 \le j^- \le \alpha^-, \qquad 0 \le j^+ \le \alpha^+, \qquad j^- + j^+ \le 2j-2$$

Our restricted domain of integration is $\tilde\Sigma = \{t_1 = \cdots = t_{j^-+1} < \cdots < t_{2j-j^+} = \cdots = t_{2j}\}$, and the phase $\phi = \tau \sum_{i=1}^j x_i - y_i$ rests unchanged but is now restricted to $\tilde\Sigma$. Thus, the corresponding terms in the error are linear combinations of integrals of the form

$$I = \frac{1}{\tau^{2j-1+l}} \int_{\tilde\Sigma} e^\phi \prod_{i=1}^{2j} \partial^{\alpha_i} v_i(t_i) dt_{j^-+1} \cdots dt_{2j-j^+}$$

where the differentiation indices α_i must satisfy the constraints

$$\sum_{i=1}^{1+j^-} \alpha_i = [(k+1)/2], \qquad \sum_{i=2j-j^+}^{2j} \alpha_i = [(k+2)/2].$$

The exponential e^ϕ decays fast in the maximum of the distances of the points. Denoting the first and the last function by

$$v^- = \prod_{i=1}^{1+j^-} \partial^{\alpha_i} v_i, \qquad v^+ = \prod_{i=2j-j^+}^{2j} \partial^{\alpha_i} v_i,$$

we can view I as a connected integral form applied to the functions

$$v^-, v_{2+j^-}, \cdots, v_{2j-j^+-1}, v^+$$

Then the same argument as in the proof of Proposition 5.10 yields the bound

$$|I| \lesssim \frac{1}{\tau^{2j-1+l}} \|v^-\|_{l_\tau^{\frac{2j}{j^-+1}} DU^2} \|v^+\|_{l_\tau^{\frac{2j}{j^++1}} DU^2} \prod_{i=2+j^-}^{2j-j^+-1} \|v_i\|_{l_\tau^{2j} DU^2} \tag{6.10}$$

where we have appropriately rebalanced the l^p indices. We recall the definition Definition 5.2, the bilinear estimate (5.3)

$$\|uv\|_{l^p_\tau DU^2} = \| \|\chi_{[k/\tau,(k+1)/\tau)}(uv)\|_{DU^2}\|_{l^p} \lesssim \| \|\chi_{[k/\tau,(k+1)/\tau)} u\|_{V^2}\|_{l^q} \| \|\chi_{[k/\tau,(k+1)/\tau)} v\|_{DU^2}\|_{l^r}$$

whenever $2 \le p$, $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$ by Hölder's inequality. We apply the embedding (5.1) and Lemma 5.4 to bound

$$\| \|u\|_{l^q_\tau V^2} \lesssim \|u'\|_{l^q DU^2} + \tau \|u\|_{l^q DU^2}.$$

After reordering the α_j so that α_1 is the largest among α_j, $j \le j^- + 1$ we recursively apply this estimate:

$$\|v^-\|_{l_\tau^{\frac{2j}{j^-+1}} DU^2} \lesssim \|\partial^{\alpha_1} v_1\|_{l^{2j} DU^2} \prod_{i=2}^{j^- -1} (\|\partial^{\alpha_i+1} v_i\|_{l^{2j} DU^2} + \tau^{-1} \|\partial^{\alpha_i} v_i\|_{l^{2j} DU^2}).$$

We argue similarly for v^+. Summing up the results, the total number of derivatives is at most $\alpha^+ + \alpha^- = 2j - 1 + k$, and the largest is at most $[(k+2)/2]$. Thus the conclusion of the Proposition 6.4 follows. □

Our last task is to convert the above bound into an estimate for H^s functions. We relate the indices s and k by the relation

$$j - 1 + \frac{k}{2} \le s \le j - 1 + \frac{k+1}{2} \tag{6.11}$$

Then we have

Proposition 6.5. *Let s be as in* (6.11)*. Then the error estimates in the above expansion satisfy the pointwise bounds*

$$\left| T_\Sigma(i\tau) - \sum_{l=0}^{k} T_\Sigma^l (i\tau)^{-(2j-1+l)} \right| \lesssim \tau^{-2s-1} \|(u,v)\|^2_{\dot{H}^s} \|(u,v)\|^{2j-2}_{l_1^2 DU^2}, \tag{6.12}$$

as well as the integrated bound

$$\int_1^\infty \tau^{2s} \left| T_\Sigma(i\tau) - \sum_{l=0}^{k} T_\Sigma^l (i\tau)^{-(2j-1+l)} \right| d\tau \lesssim \frac{1}{|\sin(2\pi s)|} \|(u,v)\|^2_{\dot{H}^s} \|(u,v)\|^{2j-2}_{l_1^2 DU^2}. \tag{6.13}$$

Moreover the error estimates for $\tilde{T}_{2j}$ satisfy the integrated bound

$$\int_1^\infty \tau^{2s} \left| \mathrm{Re}(\tilde{T}_{2j}(i\tau) + i \sum_{l=0}^{k} (i\tau)^{-(2j-1+l)} T_{2j}^l) \right| d\tau \lesssim \frac{1}{|\sin(\pi s)|} \|(u,v)\|^2_{\dot{H}^s} \|(u,v)\|^{2j-2}_{l_1^2 DU^2}. \tag{6.14}$$

The first two estimates transfer directly to $\tilde{T}_{2j}$, which is a linear combination of primitive integrals of length $2j$. It completes the estimates of our main theorem in all cases, except for s near half integers where we lack uniformity in the estimate. There we need to further take advantage of the fact that the conserved momenta H_{2j+1} are real.

Proof. For the pointwise bound we need to show that

$$I = \tau^{-|\alpha|} \prod \|\partial^{\alpha_l} u\|_{l_\tau^{2j} DU^2} \lesssim \tau^{-2s+1} \|u\|^2_{H^s} \|u\|^{2j-2}_{l^2 DU^2}$$

if s is in the range

$$j - 1 + \frac{k}{2} \le s \le j - 1 + \frac{k+1}{2}$$

and

$$k + 1 \le \sum \alpha_j \le 2j - 1 + k, \qquad \max \alpha_j \le \left[\frac{k}{2}\right] + 1.$$

As for Proposition 6.2 we take a Littlewood-Paley decomposition and expand. Again we order the terms so that the first two terms are of highest frequency. The worst case is when all derivatives fall on the first two terms, more precisely we may replace the condition above by

$$\alpha_1 \le \left[\frac{k}{2}\right], k+1 \le \alpha_1+\alpha_2 \le 2j-1+k, \qquad \alpha_2 \le 2j-1+\left[\frac{k+1}{2}\right] \tag{6.15}$$

and $\alpha_j = 0$ for $j \ge 3$. We assume (6.15) in the following. Repeating the argument of Proposition 6.2 - but adjusting for the derivatives - we have to multiply the constants in (6.5) by $\left(\frac{\lambda_1}{\tau}\right)^{\alpha_1}\left(\frac{\lambda_2}{\tau}\right)^{\alpha_2}$,

$$C(\tau,\lambda_1,\lambda_2,\alpha_1,\alpha_2) = \begin{cases} \left(\dfrac{\lambda_1}{\tau}\right)^{2(-s+j-1)+\alpha_1+\alpha_2} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{-s+2j-\frac{5}{2}+\frac{1}{2j}+\alpha_2} & \lambda_2 \le \lambda_1 \le \tau \\ \left(\dfrac{\lambda_1}{\tau}\right)^{-(s+\frac{1}{2})+\alpha_1} \left(\dfrac{\lambda_2}{\tau}\right)^{-s+2j-\frac{5}{2}+\frac{1}{2j}+\alpha_2} & \lambda_2 \le \tau \le \lambda_1 \\ \left(\dfrac{\lambda_1}{\tau}\right)^{-(s+\frac{1}{2})+\alpha_1} \left(\dfrac{\lambda_2}{\tau}\right)^{-(s+\frac{1}{2})+\alpha_2} & \tau \le \lambda_2 \le \lambda_1. \end{cases} \tag{6.16}$$

As above this implies the pointwise bound

$$\tau^{2s+1-|\alpha|} \prod \|\partial^{\alpha_l} u\|_{l^{2j}_\tau DU^2} \le \sqrt{A_p^{\alpha_1,\alpha_2} B_p^{\alpha_1,\alpha_2}} \|u\|^2_{H^s} \|u\|^{2j-2}_{l^2_1 DU^2}$$

and the integrated bound

$$\int_1^\infty \tau^{2s+1-|\alpha|} \prod \|\partial^{\alpha_l} u\|_{l^{2j}_\tau DU^2} d\tau \le A_i^{\alpha_1,\alpha_2} \|u\|^2_{H^s} \|u\|^{2j-2}_{l^2_1 DU^2}$$

where the constants are determined by geometric sums. Again we give the result for the integrated bounds A_i up to multiplicative constants.

$$1+\frac{1}{(2(j-1-s)+\alpha_1+\alpha_2)(2j+\alpha_2-s-\frac{5}{2}+\frac{1}{2j})} \le 1+\frac{1}{2(j-1+\frac{k+1}{2}-s)(j-1+\frac{k+1}{2}-s+\frac{1}{4})} \tag{6.17}$$

in the range $\lambda_2 \le \lambda_1 \le \tau$,

$$1+\frac{1}{(s+\frac{1}{2}-\alpha_1)(2j+\alpha_2-s-\frac{5}{2}+\frac{1}{2j})} \le 1+\frac{1}{(2(j-1+\frac{k+1}{2})-s)(j-1+\frac{k+1}{2}-s+\frac{1}{4})} \tag{6.18}$$

in the range $\lambda_1 \le \tau \le \lambda_2$ and finally in the range $\tau \le \lambda_2 \le \lambda_1$

$$1+\frac{1}{(s+\frac{1}{2}-\frac{k}{2})(2s-2j-k)} \tag{6.19}$$

We arrive at (6.12) and (6.13) and note that estimate (6.14) follows directly from (6.13) unless s is close to a half-integer. It remains to prove (6.14) near half integers. Hence we choose $s_0 = j+k-\frac{1}{2}$ and consider s close to s_0. To shorten the notation we define

$$\mathbb{T}^N_{2j}(i\tau)(u) = \tilde{T}_{2j}(i\tau) + \frac{1}{2\pi i}\sum_{l=0}^N H_{2j,l}(i\tau)^{1-2j-l}.$$

The crucial observation is that most partial estimates above extend to a larger range of s. The claim will follow from the inequalities

$$\int_1^\infty (\tau^2-1)^s \left|\mathbb{T}^{2k}_{2j}(i\tau)(u) - \mathbb{T}^{2k}_{2j}(i\tau)(u_{>\tau})\right| \lesssim \|u\|^2_{H^s}\|u\|^{2j-2}_{l^2_1 DU^2} \tag{6.20}$$

and

$$\int_1^\infty (\tau^2-1)^s \left|\mathbb{T}_{2j}^{2k+1}(i\tau)(u) - \mathbb{T}_{2j}^{2k+1}(i\tau)(u_{<\tau})\right| \lesssim \|u\|_{H^s}^2 \|u\|_{l_1^2 DU^2}^{2j-2} \tag{6.21}$$

for $|s_0 - s| \le \frac{1}{8}$, and the similar estimates (6.23) and (6.22) below. Each term in the Littlewood-Paley expansion of the factors in $\mathbb{T}_{2j}^{2k+1}(u) - \mathbb{T}_{2j}^{2k+1}(u_{<\tau})$ contains at least one factor with frequency $\ge \tau$ and hence $\lambda_1 \ge \tau$. Thus we only need to consider the cases $\lambda_2 \le \tau \le \lambda_1$ and $\tau \le \lambda_2 \le \lambda_1$. In these cases the estimates above extend to $|s - s_0| \le \frac{1}{8}$, see (6.18) and (6.19).

Similarly each term of the Littlewood-Paley expansion of $\mathbb{T}_{2j}^{2k}(u) - \mathbb{T}^{2k}(u_{>\tau})$ contains at least factor with one frequency below τ. If $\lambda_2 \le \tau$ then we are in the middle regime which is bounded for $|s - s_0| \le \frac{1}{8}$. If $\lambda_2 \ge \tau$ we order the expansion so that $\lambda_3 \le \tau$, which we estimate by

$$\|u_{\lambda_3}\|_{l_\tau^{2j} DU^2} \le c \left(\frac{\lambda_3}{\tau}\right)^{\frac{1}{2}-\frac{1}{2j}} \|_{l_1^2 DU^2}.$$

Then

$$\int_1^\infty \sum_{\lambda_3 \le \tau \le \lambda_2 \le \lambda_1} \tau^{2s} \prod_{j=1}^{3} \|u_{\lambda_j}\|_{l_\tau^{2j} Du^2} \|u_{\le \lambda_2}\|_{l_\tau^{2j} DU^2}^{2j-3} d\tau \lesssim \|u\|_{H^s}^2 \|u\|_{l_1^2 DU^2}^{2j-2}.$$

The same argument shows that

$$\int_1^\infty (\tau^2-1)^s \left|\mathbb{T}^{2k}(u_{>\tau})\right| d\tau \lesssim \|u\|_{H^s}^2 \|u\|_{l_1^2 DU^2}^{2j-2} \tag{6.22}$$

and

$$\int_1^\infty (\tau^2-1)^s \left|\mathbb{T}^{2k+1}(u_{<\tau})\right| d\tau \lesssim \|u\|_{H^s}^2 \|u\|_{l_1^2 DU^2}^{2j-2}. \tag{6.23}$$

Since $\operatorname{Im} \mathbb{T}^{2k}(u_{<\tau}) = \operatorname{Im} \mathbb{T}^{2k+1}(u_{<\tau})$ we obtain the uniform estimate below s_0 by the triangle inequality. For $s_0 \le s \le s_0 + \frac{1}{8}$ $\operatorname{Im} \mathbb{T}^k(u_{>\tau}) = \operatorname{Im} \mathbb{T}^{k+1}(u_{>\tau})$ and again the uniform estimate follows by an application of the triangle inequality. □

7. Proof of the main theorems and variants

In this section we combine the bounds in the previous sections in order to prove our main results in Theorem 1.1 and Theorem 2.4. We also discuss some further developments of our ideas and results.

7.1. **Proof of Theorem 1.1.** This is done in two steps. First we show that the energies E_s defined by the right hand side of formulas (2.11), (2.12) are smooth as functions of $u \in H^s$, and satisfy the bounds in part (2) of the theorem. Secondly, we show that they are conserved along the flow.

For this we begin with the multilinear expansion for $\ln T$, namely

$$\ln T = \sum_{j=1}^\infty \tilde{T}_{2j}$$

which, by Corollary 5.11, converges on the half-line $i[1,\infty)$ provided that

$$\|u\|_{l^2 DU^2} \ll 1.$$

Correspondingly, this yields a multilinear expansion for the energies E_s,

$$E_s = \sum_{j=1}^\infty E_{s,2j}$$

We will estimate separately each of the terms in the series, in several steps:

7.1.1. *The leading term* $j = 1$. As proved in Proposition 4.4, the first term in this expansion is exactly the H^s norm,

$$E_{s,2} = \|u\|_{H^s}^2$$

It remains to consider the rest of the terms in the series, and show that we can bound them as error terms.

7.1.2. *Large* j, $j > 2s+1$. Here we discuss the tail of the series, for which we can directly use Proposition 5.13; this contains bounds for $\tilde{T}_{2j}$ which can be obtained directly from similar T_{2j} bounds. Precisely, Proposition 5.13; shows that for large enough j we have a favorable bound for $E_{s,2j}$, namely

$$|E_{s,2j}| \leq \|u\|_{H^s}^2 (c\|u\|_{l^2DU^2})^{2j-2}, \qquad j > 2s+1$$

This suffices if $s < \frac{1}{2}$. For larger s, however, we need stronger bounds if j is small. We remind the reader that the issue here is that the bounds for $\tilde{T}_{2j}$ in Proposition 5.13 are derived from similar bounds for T_{2j}; on the other hand, for smaller s we expect such bounds to hold for $\tilde{T}_{2j}$ but not for T_{2j}. Thus, we need to take advantage of the fact that $\tilde{T}_{2j}$ has a better structure than T_{2j} (namely the fact that, unlike T_{2j}, $\tilde{T}_{2j}$ contains only connected integrals).

7.1.3. *Medium* j, $j > s+1$. Here we use the bounds for $\tilde{T}_{2j}$ in Proposition 6.2, which yield

$$|E_{s,2j}| \leq \|u\|_{H^s}^2 (c\|u\|_{l^2DU^2})^{2j-2}, \qquad j > s+1$$

We remark that, due to the previous step, for fixed s we need this bound only for a finite set of j's, therefore the uniformity of the bounds with respect to j is no longer an issue.

7.1.4. *Small* j, $j \leq s+1$. In this range we face the additional difficulty that in the $2j$ component of the formula (2.12) we need to also take into account the energy corrections T^l_{2j}. This is addressed in Section 6. At least for s away from $\mathbb{Z}/2$, we can directly use the bound (6.13) to conclude that we have the bound

$$|E_{s,2j}| \leq \|u\|_{H^s}^2 (c\|u\|_{l^2DU^2})^{2j-2}, \qquad j \geq 2 \tag{7.1}$$

The only remaining issue is to show that this bound is uniform as well for s near integers and half-integers. These two situations are somewhat different, and are considered separately.

7.1.5. *s near integers.* To set the notations, we fix an integer $s_0 \geq 1$ and consider s near s_0. On one hand, in the bound (6.13) we have $(s-s_0)^{-1}$ growth. On the other hand, in the integral (2.12) we have an additional $\sin(\pi s)$ factor. These two expressions cancel and we obtain again the bound (7.1), uniformly for s near s_0.

7.1.6. *s near half- integers.* Here we fix $s_0 \in \mathbb{N} + \frac{1}{2}$ and again consider s near s_0. The integral bound for $\tilde{T}_{2j}(z)$ fails to be uniform as s approaches s_0. Nevertheless, we are saved by the fact that only $\operatorname{Re} \tilde{T}_{2j}(z)$ is needed. But this satisfies a favorable bound by Proposition 6.5.

7.1.7. *Dependence on s.* If $u \in H^{s_0}$ with $s_0 > -\frac{1}{2}$ then, with $N = [s_0]$, analyticity with respect to $s \in (-\frac{1}{2}, s_0)$ is a consequence of the integral formula

$$E_s(u) = 4\sin(\pi s) \int_1^\infty (t^2-1)^s \Big(\operatorname{Re} \ln T(it/2) + \frac{1}{2\pi} \sum_{j=0}^{N} (-1)^N H_{2j}\, t^{-2j-1} \Big) dt + \sum_{j=0}^{N} \binom{s}{j} H_{2j}, \tag{7.2}$$

It allows for an extension of E_s in a complex neighborhood of $(-\frac{1}{2}, s_0)$ to a holomorphic function.

Continuity at s_0 from the left follows from the above analiticity property and the uniform Lipschitz bounds for $E_s(u)$, $s \leq s_0$, in terms of $u \in H^{s_0}$, simply by approximating u with more regular functions.

Finally, we remark that if s is an integer then the coefficient of the integral vanishes and the sum gives the desired identification (4).

7.1.8. *The conservation of E_s.* Here we show that E_s is conserved along the NLS and mKdV flow. It suffices to do this for Schwartz initial data, in which case T is constant along the flow. Solutions with Schwartz initial data remain Schwartz functions, and for those the formal analysis of the Jost solutions becomes rigorous.

7.2. **Proof of Theorem 2.4.** Here the goal is to extend the considerations in the previous proof to data u which is no longer small in l^2DU^2. We will do this in two steps. First we show that the energies given by (2.11), (2.12) are well defined for all u in H^s, and have the appropriate regularity properties. Then we will prove that the trace formulas hold.

7.2.1. *The energies E_s for large data.* Given $u \in H^s$, we can choose some large τ_0 depending on $\|u\|_{H^s}$ so that

$$\|u\|_{l^2_{\tau_0}DU^2} \leq \delta$$

Rescaling the bounds for $\ln|T|$ in the proof of Theorem 1.1, it follows that the formal series for $\ln|T|$ converges on $i[\tau_0,\infty)$ with favourable bounds. This shows that the part of the integrals (2.11), (2.12) for $t \in [\tau_0,\infty)$ converges and satisfies the desired H^s bounds.

It remains to consider the part of the integrals in the interval $i[1,\tau_0]$. In the defocusing case we know that in this interval $\ln|T|$ depends analytically on $u \in l^2DU^2$, which is more than enough to close the argument.

However, in the focusing case we only know that T^{-1} depends analytically on $u \in l^2DU^2$. Thus if T has poles in the interval $i[1,\tau_0]$ then $\ln|T|$ has logarithmic singularities there. This still suffices for the convergence of the integrals (2.11), (2.12) for $t \in [\tau_0,\infty)$, but we need an additional argument to establish the regularity properties of the integral with respect to u. Precisely, we need to account for the poles of T which cross the interval $i[1,\tau_0]$. Changing the contour of integration to one without poles we get the sum of residues

$$\sum_k m_k \Xi(2z_k)$$

over the poles near the line $i[1,\tau_0]$. Since Ξ is real analytic outside $z = i$, it follows that the contributions of the poles away from $i/2$ is analytic in u. At $z = i/2$ the function Ξ is of class C^{s+1}, so continuity of E_s follows.

7.2.2. *The trace formula for small data.* Here we show that the middle terms of (2.14) and (2.15) are defined and coincide with the right hand side. We first consider small data $\|u\|_{l^2DU^2} \ll 1$.

a) $s < 0$. In the defocusing case $-\ln|T|$ is a nonnegative harmonic function in the upper half-space, so the conclusion (2.14) follows directly from Lemma 4.3. In the focusing case $-\ln|T|$ is a nonnegative subharmonic function in the upper half-space. Further, denoting the poles of T in the upper half-space by z_k and their corresponding multiplicities by m_k, we have

$$-\Delta \ln|T| = \sum_k m_k \delta_{z_k}$$

By Lemma 5.9 and Lemma 5.6, these are located in $\{0 < \Im z \ll (1+|\Re z|)^s\}$. Hence the trace formula (2.15) is again a consequence of Lemma 4.3.

b) $s = N \geq 0$, integer. In the defocusing case we can directly apply Lemma 4.2. In the focusing case the condition $u \in L^2$ guarantees that all poles of T are contained in a strip along the real axis. Then we can apply Lemma 4.2.

c) $N < s < N+1$, noninteger. Then we apply first step (b) above for $s = N$. This places us in a context where we can use Lemma 4.3.

7.2.3. *The trace formula for large data.* Here we take arbitrary $u \in H^s$, and we first claim that we still have the bounds

$$\int_{\mathbb{R}} (1+\xi^2)^s d\mu + \int_U \operatorname{Im} z(1+|z|^2)^s d\nu < \infty \tag{7.3}$$

To see that we first observe that for large enough τ we have

$$\|u\|_{l^2_{\tau_0} DU^2} \ll 1$$

Hence the rescaled function

$$u_\tau(x) = \tau_0 u(\tau_0 x)$$

has small ${}^2DU^2$ norm, and then we can apply the small data trace regularity. This proves our claim.

Next we consider the trace formulas for u. In view of the trace regularity property (7.3), this follows directly as in Lemma 4.3.

7.2.4. *The global regularity of the energies E_s.* In the defocusing case, for arbitrary $u \in H^s$ we consider a large frequency scale τ as above. For $\tau \gg \tau_0$ we can take advantage of the rescaling to conclude summability of the asymptotic expansion for $\log |T|$ with good H^h bounds, and analyticity follows. For $1 \leq \tau \lesssim \tau_0$ we instead use directly the analyticity of $\log |T|$ as a function of $u \in l^2 DU^2$.

In the focusing case we argue in the same manner provided that there are no poles for T in $i[1, \tau_0)$. Further, the trace formulas show that the poles away from i are in effect harmless, as the function Ξ_s is analytic there.

7.3. **More conserved quantities.** We denote

$$\mathbb{C}_+ = \Big\{ z \in \mathbb{C}, \operatorname{Im} z > 0, z \notin i[1,\infty) \Big\}.$$

Definition 7.1. *For $s > -\frac{1}{2}$ we define the class of weight functions $\mathcal{E}_s$ which are holomorphic in $\mathbb{C}_+$, continuous up to the boundary (with possible different limits from the left and the right), real on the real axis and which satisfy*

$$|\mu(z)| \leq c < z >^{2s} .$$

By the Schwarz reflection principle they are holomorphic functions in a domain symmetric around the real axis. We define Ξ by

$$\frac{d}{dz}\Xi = \mu, \quad \Xi(0) = 0$$

for $\xi \in \mathcal{E}_s$. Let

$$[\mu](t) = \mu(it+0) + \mu(it-0)$$

for $t > 1$. If μ is even on $\mathbb{R}$ then $\mu(-\bar z) = \overline{\mu(z)}$ and if μ is odd then $\mu(-\bar z) = -\overline{\mu(z)}$. Thus $[\mu]$ is purely imaginary if ξ is even and real if μ is odd on $\mathbb{R}$.

Theorem 7.2. *There exists $\delta > 0$ so that the following is true: If $s > -\frac{1}{2}$, $s \notin \mathbb{Z}/2$, $\xi \in \mathcal{E}_s$ then there is a unique conserved quantity E_μ on*

$$\{u \in H^s, \|u\|_{l^2_1 DU^2} \leq \delta\}$$

which satisfies

$$\left| E_\mu(u) - \int \mu(\xi) |\hat u(\xi)|^2 d\xi \right| \leq c \|u\|_{H^s}^2 \|u\|_{l^2_1 DU^2}^2$$

In particular the quadratic part is given by the Fourier multiplier μ. Moreover in the defocusing case E_μ is defined on whole of H^s and

$$
\begin{aligned}
E_\mu(u) &= -\int \mu(\xi)\ln|T(-\xi/2)|d\xi \\
&= -\operatorname{Re}\int_1^\infty [\mu](t)\Big(\ln(T) + \frac{1}{2\pi i}\sum_{j=1}^N \mu^{(j)}(0)H_{j-1}(it)^{-j}\Big)dt + \sum_{j=1}^N \frac{1}{j!}\mu^{(j)}(0)H_{j-1}
\end{aligned}
\tag{7.4}
$$

and in the focusing case

$$
\begin{aligned}
E_\xi(u) &= \int \mu(\xi)\ln|T(-k/2)|d\xi + \sum m_j \operatorname{Im}\Xi(z_j) \\
&= -\operatorname{Re}\int_1^\infty [\mu](t)(\ln(T) - \frac{1}{2\pi i}\sum_{j=1}^N \mu^{(j)}(0)H_{j-1}(i\tau)^{-j}d\tau + \sum_{j=1}^N \frac{1}{j!}\mu^{(j)}(0)H_{j-1}
\end{aligned}
\tag{7.5}
$$

where $N = [2s]$, and the sum runs over the eigenvalues z_j, with algebraic multiplicities m_j. It is continuous in μ and uniformly analytic in u. If $u \in H^{s_0}$ for some $s_0 > s$ then the map

$$\mathcal{E}_s \ni \mu \to E_\mu(u)$$

is analytic in μ. The quartic term is given by

$$\int_{\xi_1+\xi_2=\eta_1+\eta_2} \frac{\mu(\xi_1)+\mu(\xi_2)-\mu(\eta_1)-\mu(\eta_2)}{(\xi_1-\eta_1)(\xi_1-\eta_2)}\hat{u}(\xi_1)\hat{u}(\xi_1)\overline{\hat{u}(\eta_1)\hat{u}(\eta_2)}.$$

Proof. In the defocusing case the middle term of (7.4) is well-defined for $u \in H^s(\mathbb{R})$ if $s > -\frac{1}{2}$ by trace formula of Theorem 2.4. In the focusing case smallness ensures that there is no eigenvalue on $i[1,\infty)$ and hence the evaluation of Ξ is uniquely defined. Again the middle integral is well defined by (2.16) after Theorem 2.4. If s is not a half integer, and if we consider either the defocusing case or the focusing case under the smallness assumption then the integrals on the right hand side of (7.4) and (7.5) are well defined. The arguments of Section 4 combined with dominated convergence imply the second equality in (7.4) and (7.5). Analyticity with respect to μ is obvious for the term in the middle, whereas analyticity with respect to u is obvious for the term on the right hand side of (7.4) resp. (7.5). The identification of the quadratic term follows as for E_s in Section 4. The quartic term is identified as in Proposition 8.2. $\square$

7.4. **Frequency envelopes.** One consequence of our main theorem is that if the initial data is small in H^s then the solution remains in H^s at all times, with a comparable uniform bound[2]. A natural follow-up question is whether the H^s energy can migrate arbitrarily between frequencies. To provide some insight into this it is convenient to use weighted norms which are closely related to frequency envelopes. Precisely, we consider sequences $\{a_k\}_{k\geq 0}$ for which the following properties hold:

1. The following bound from above holds: $a_k \lesssim 1$.
2. The sequence a_k is slowly varying, $a_k/a_j \leq 2^{c|j-k|}$ Here c is a sufficiently small constant, whose choice may depend on the problem at hand.

For such sequences we define the weighted norm

$$\|u\|_{H^a}^2 = \sum a_j^2 2^{2js}\|u_{2^j}\|_{L^2}^2.$$

Then one way of saying that the H^s energy of a solution u does not travel much between frequencies is if not only the H^s norm stays uniformly bounded, but for any sequence $\{a_k\}$ as above, $\|u(t)\|_{H^a}$

[2]We also recall that by rescaling a similar result holds for large data, but there the uniform bound is no longer universal, instead it will depend also on the data size.

is uniformly bounded in terms of $\|u(0)\|_{H^a}$. This is not exactly the standard frequency envelope formulation, but the two are equivalent.

A small variation of the proof of Theorem (7.2) shows indeed that this is the case:

Theorem 7.3. *For all $s > 0$ there exists $\varepsilon > 0$ so that the following is true. Let (a_j) be a sequence satisfying*

(1) $2^{\varepsilon - \frac{1}{2}j} \le a_j \le 2^s$
(2) *If $j_1 < j_2 < j_3$ then*

$$\frac{\ln a_{j_2} - \ln a_{j_1}}{j_2 - j_1} \le \frac{\ln a_{j_3} - \ln a_{j_2}}{j_3 - j_2} + \varepsilon.$$

Then there exists $\mu \in \mathcal{E}_s$ so that

$$\sum_{j=1}^{\infty} a_j \|u_{2^j}\|_{L^2}^2 \approx \int \mu(\xi) |\hat{u}|^2 d\xi$$

and

$$|E_\mu(u) - \int \mu(\xi)) |\hat{u}(\xi)|^2 | \le c \|u\|_{H^a}^2 \|u\|_{l^2_1 DU^2}^2 \tag{7.6}$$

provided $\|u\|_{l^2_1 DU^2} \le \delta$. In particular if $s > -\frac{1}{2}$, and u_0 small in $l^2_1 DU^2$ then

$$\|u(t)\|_{H^a} \lesssim \|u_0\|_{H^a}.$$

Proof. The first part implies the second one. By a superposition argument it suffices to prove the theorem only for a sequence

$$a_j = \begin{cases} 2^{s_0 j} & \text{for } j \le j_0 \\ 2^{s_1 j + (s_0 - s_1) j_0} & \text{for } j \ge j_0. \end{cases}$$

with $-\frac{1}{2} < s_1 < s_0 < s_1 + \varepsilon$. Then

$$\|u\|_{H^s_a} \approx \|u_{<2^{j_0}}\|_{H^{s_0}} + 2^{(s_0 - s_1) j_0} \|u_{>2^{j_0}}\|_{H^{s_1}}$$

We choose

$$\mu_{s_0, s_1, j_0}(\xi) = (1 + \xi^2)^{s_0} (1 + (2^{-j_0} \xi)^2)^{s_1 - s_0}$$

so that the quadratic part of the energy E_μ is given by this Fourier multiplier by Theorem 7.2 and we have to prove the bound (7.6). This follows by an adaptation of the proof of Theorem 1.1 resp. an application of various estimates proven there.

(1) If $-\frac{1}{2} < s_1 < s_0 < 1$ then the estimate follows as in the proof of Prosition 6.2. The geometric series allow this range of exponents.
(2) If $j \le s_1 < s_0 < j + 1/2$ or $j + 1/2 < s_1 < s_2 < j + 1$ then the bound follows similarly from the integrated bounds (6.13) resp. their proof. This again deteriorates as we approach $j + \frac{1}{2}$ but not at the other endpoints.
(3) If $2s_0$ and $2s_1$ are close to $2j + 1$ then we want to argue as for (6.14). We check that for $t \ge 1$

$$[\mu](t) = 2 \begin{cases} -2 \sin \pi s_0 (t^2 - 1)^{s_0} (1 - t^2 2^{-2j_0})^{s_1 - s_0} & \text{for } 1 \le t \le 2^{j_0} \\ -2 \sin \pi s_1 (t^2 - 1)^{s_0} (t^2 2^{-2j_0} - 1)^{s_1 - s_0} & \text{for } 2^{j_0} \le t \end{cases}$$

Again a direct adaption of the proof gives the result.

(4) If $s_1 < j \leq s_0$ we need a more careful correction. In this case we try to bound

$$\mathrm{Re} \int_1^\infty [\mu](t) \Big(\ln T(it/2) + \frac{1}{2\pi i} \sum_{l=0}^{j-1} E_{2l}(u)(it)^{-2l-1} - E_{2j}(u_{<2^{-j_0}})(it)^{-2j-1} dt + \dots$$

This follows in the same fashion as part of the proof of (6.14).

□

We also remark on an interesting but more straightforward case of the above result. If we confine ourselves to negative Sobolev norms, then the above theorem holds for any frequency envelope satisfying

$$a_k \geq a_{k+1} \geq 2^{-(\frac{1}{2}+\delta)} a_k$$

The simplification here is that no new energies are needed; instead it suffices to work with $E_{-\frac{1}{2}+\delta}$ and its rescaled versions.

7.5. **Generalized momentae.** As noted earlier in the paper, the conserved quantities E_s in our main result in Theorem 1.1 are positive definite, and can be viewed as inhomogeneous extensions of the even conserved Hamiltonians H_{2k}. In particular they are not connected at all with the odd conserved energies, i.e. the generalized momenta.

A natural question would be whether one can similarly define a continuous family of momenta. This is indeed the case, as one can replicate as follows the construction of the conserved energies E_s. Starting from the observation that, at least in the defocusing case, the odd Hamiltonians can be expressed in the form

$$H_{2k+1} = \int_{\mathbb{R}} \xi^{2k+1} \ln |T| d\xi$$

it is natural to seek to define the generalized momenta as

$$P_s = \int_{\mathbb{R}} \xi (1+\xi^2)^{s-\frac{1}{2}} \ln |T| d\xi, \qquad s > -\frac{1}{2},$$

so that if s is a half-integer we recover the (linear combinations of) odd Hamiltonians, i.e. the momenta.

To render this definition useful for non-decaying data, as well as in the focusing case, we switch this as in the definition of the energies E_s to a contour integral over the double half-line $i[1,\infty)$. For small s this is done directly,

$$P_s = -\frac{\sin(\pi(s-\frac{1}{2}))}{2\pi} \int_1^\infty t(t^2-1)^{s-\frac{1}{2}} \,\mathrm{Re} \ln T(it)\, dt, \qquad -\frac{1}{2} < s < \frac{1}{2} \tag{7.7}$$

For larger s we need to remove a number of terms in the formal expansion of $T(it)$ in order for the above integral to converge. Precisely, for s in the range

$$N + \frac{1}{2} \leq s < N + \frac{3}{2}, \qquad N \geq 0$$

we set

(7.8)

$$P_s = -\frac{\sin(\pi(s-\frac{1}{2}))}{2\pi} \int_1^\infty t(t^2-1)^{s-\frac{1}{2}} \left(\mathrm{Re} \ln T(it) + \sum_{j=0}^{N} (-1)^j H_{2j+1} t^{-2j-2} \right) dt + \frac{1}{2} \sum_{j=0}^{N} \binom{s-\frac{1}{2}}{j} H_{2j+1}$$

We obtain the following result, which for $s \in (-\frac{1}{2}, \infty) \backslash (\mathbb{Z}/2)$ is a special case of Theorem 7.2:

Theorem 7.4. *For each $s > -\frac{1}{2}$ and $\delta > 0$ and both for the focusing and defocusing case the functional*

$$P_s : \{u \in H^s | \|u\|_{H^{-\frac{1}{2}+\delta}} \ll 1\} \to \mathbb{R}^+$$

is conserved along the NLS and mKdV flow. Further, we have

$$\left| P_s(u) - \langle u, i\partial_x u\rangle^2_{H^{s-\frac{1}{2}}} \right| \lesssim \|u\|^2_{l^2_1 DU^2} \|u\|^2_{H^s}. \tag{7.9}$$

The case $s \in \mathbb{Z}/2$ follows as for Theorem 2.4, but with the role of integers and half integers reversed.

8. The terms of homogeneity 4 and 6 in u

In this section we will derive a simple expression for T_4, T_6, $E_{s,4}$ and $E_{k,6}$ for integers k.

8.1. **The quartic term in $\ln T(z)$ and $E_s(u)$.** Here we discuss in more detail the expression $\tilde{T}_4(z)$, as well as the corresponding quartic term in a homogeneous expression $E_s(u)$. By the same techniques one can obtain similar results for $\tilde{T}_{2j}(z)$. We begin by computing the symbol for $\tilde{T}_4(z)$ viewed as a translation invariant quadrilinear form. We recall that $\tilde{T}_4(z)$ is given by (using Fourier inversion in the second step)

$$\begin{aligned}
\tilde{T}_4(z) = & -2 \int_{x_1<x_2<y_1<y_2} e^{2iz(y_1+y_2-x_1-x_2)} u(y_1)u(y_2)\overline{u(x_1)u(x_2)} dxdy \\
= & -\frac{1}{2\pi^2} \int_{\{x_1<x_2<y_1<y_2\}\times\mathbb{R}^4} e^{2iz(y_1+y_2-x_1-x_2)-i(x_1\xi_1+x_2\xi_2-y_1\eta_1-y_2\eta_2)} \overline{\hat{u}(\xi_1)\hat{u}(\xi_2)} \hat{u}(\eta_1)\hat{u}(\eta_2) \\
& \times dx_1dx_2\, dy_1dy_2\, d\xi_1 d\xi_2\, d\eta_1 d\eta_2.
\end{aligned}$$

To compute the kernel we first symmetrize with respect to the x variables and integrate exponentials succesively:

$$\begin{aligned}
K_z(\xi,\eta) = & -\frac{1}{4\pi^2} \int_{x_1<x_2<y_1,y_2} e^{2iz(y_1+y_2-x_1-x_2)-i(x_1\xi_1+x_2\xi_2-y_1\eta_1-y_2\eta_2)} dxdy \\
= & -\frac{1}{4\pi^2 i} \frac{1}{2z+\xi_1} \frac{1}{2z+\xi_2} \frac{1}{2z+\eta_1} \int e^{-iy_1(\xi_1+\xi_2-\eta_1-\eta_2))} dy_1 \\
= & -\frac{1}{4\pi^2 i} \frac{1}{2z+\xi_1} \frac{1}{2z+\xi_2} \frac{1}{2z+\eta_2} \delta_{\xi_1+\xi_2-\eta_1-\eta_2}.
\end{aligned}$$

Replacing the above kernel with the symmetrization with respect to η we obtain

$$\tilde{T}_4(z) = -\frac{1}{8\pi^2 i} \int_{\mathbb{R}^3} \frac{4z+\xi_1+\xi_1}{(2z+\xi_1)(2z+\xi_2)(2z+\eta_1)(2z+\eta_2)} \overline{\hat{u}(\eta_1+\eta_2-\xi_1)\hat{u}(\xi_1)} \hat{u}(\eta_1)\hat{u}(\eta_2) d\xi_1 d\eta_1 d\eta_2. \tag{8.1}$$

Of course we can do the same calculation with the roles of ξ and η interchanged. This allows to replace the product of the Fourier transforms by the real part of the Fourier transforms. By an abuse of notation we write the last integral as $\int_{\xi_1+\xi_2=\eta_1+\eta_2}$. Undoing the symmetrization we arrive at

$$\tilde{T}_4(z) = -\frac{1}{4\pi^2 i} \int_{\xi_1+\xi_2=\eta_1+\eta_2} \frac{1}{(2z+\xi_1)(2z+\eta_1)(2z+\eta_2)} \operatorname{Re}\{\overline{\hat{u}(\xi_1)\hat{u}(\xi_1)} \hat{u}(\eta_1)\hat{u}(\eta_2)\} \tag{8.2}$$

From (8.1) we deduce the Laurent expansion of the quartic term at infinity.

Lemma 8.1. *Suppose that u is a Schwartz function. Then we obtain the asymptotic series*

$$\tilde{T}^4(z) \sim -\frac{1}{2\pi i} \sum_{j=3}^{\infty} H_{j-1,4} (2z)^{-j} \tag{8.3}$$

where

$$H_{j,4} = -\operatorname{Re} i^j \sum_{\alpha_1+\alpha_2+\alpha_3=j-2} (-1)^{\alpha_2+\alpha_3} \int u^{(\alpha_1)} u^{(\alpha_2)} \overline{u^{(\alpha_3)} u} dx. \tag{8.4}$$

Proof. We expand (8.2) in negative powers of z. Then

$$\begin{aligned} H_{j,4} =& \operatorname{Re} \frac{1}{2\pi} \sum_{\alpha_1+\alpha_2+\alpha_3=j-2} \int_{\xi_1+\xi_2=\eta_1+\eta_2} \xi_1^{\alpha_1} \eta_1^{\alpha_2} \eta_2^{\alpha_3} \overline{\hat{u}(\xi_1)\hat{u}(\xi_2)} \hat{u}(\eta_1)\hat{u}(\eta_2) \\ =& -\operatorname{Re} \Big[i^j \sum_{\alpha_1+\alpha_2+\alpha_3=j-2} (-1)^{\alpha_2+\alpha_3} \int u^{(\alpha_1)} u^{(\alpha_2)} \overline{u^{(\alpha_3)} u} \Big] dx. \end{aligned}$$

□

Next we turn our attention to the quartic term in the energies E_s. The 4-linear component of E_s^4 is given by the contour integral (2.11) or (2.12) over the half-line $i[1,\infty)$ with $\ln T$ replaced by $\tilde{T}_4$. However we can change the contour of integration back to the real line to obtain the representation that corresponds to (2.10) and

$$E_4^s = \frac{1}{4\pi^2} \int_{\xi_1+\xi_2=\eta_1+\eta_2} K_s^4(\xi,\eta) \operatorname{Re}(\overline{u(\xi_1)u(\xi_2)} \hat{u}(\eta_1)\hat{u}(\eta_2)) d\xi$$

where

$$\begin{aligned} K_s^4(\xi,\eta) =& \int_{\mathbb{R}} (1+z^2)^s \operatorname{Im} \left(\frac{2z-\xi_1-\xi_2}{(z+i0-\xi_1)(z+i0-\xi_2)(z+i0-\eta_1)(z+i0-\eta_2)} \right) dz \\ =& \int_{\mathbb{R}} (1+\xi^2)^s \frac{1}{(\xi_1-\eta_1)(\xi_1-\eta_2)} \left(\delta_{\xi-\xi_1} + \delta_{\xi-\xi_2} - \delta_{\xi-\eta_1} - \delta_{\xi-\eta_2}\right) d\xi. \end{aligned}$$

Thus we immediately obtain the following:

Proposition 8.2. *The quartic part of E_s is given by*

$$E_{s,4} = \frac{1}{4\pi} \int_{\xi_1+\xi_2=\eta_1+\eta_2} \frac{(1+\xi_1^2)^s + (1+\xi_2^2)^s - (1+\eta_1^2)^s - (1+\eta_2^2)^s}{(\xi_1-\eta_1)(\xi_1-\eta_2)} \overline{\hat{u}(\xi_1)\hat{u}(\xi_2)} \hat{u}(\eta_1)\hat{u}(\eta_2) d\xi_1 d\eta$$

We observe that this coincides with $H_{k,4}$ for $s = k \in \mathbb{N}$. Of course, this is no surprise, as this expression is exactly the quartic I-method energy correction term, see [14], [13] and [4].

8.2. **The term $\tilde{T}_6(z)$.** For completeness we also provide a brief computation for the expansion of the $\tilde{T}_6$ term, which gives the sixth-linear terms in the conserved energies.

Lemma 8.3. *The following identity holds:*

(8.5)

$$\begin{aligned} \tilde{T}_6(z) =& -\frac{1}{(2\pi)^3 i} \int_{\xi_1+\xi_2+\xi_3=\eta_1+\eta_2+\eta_3} \frac{1}{2z+\xi_1} \frac{1}{2z+\xi_2} \frac{1}{2z+\eta_2} \frac{1}{2z+\eta_3} \left(\frac{1}{2z+\eta_1} + \frac{1}{2z+(\eta_2+\eta_3-\xi_3)} \right) \\ & \times \hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3) \overline{\hat{u}(\eta_1)\hat{u}(\eta_2)\hat{u}(\eta_3)} d\xi d\eta \sim -\frac{1}{2\pi i} \sum_{j=5}^{\infty} H_{j-1,6} (2z)^{-j} \end{aligned}$$

where

(8.6)

$$H_{j,6} = \operatorname{Re} \Big[i^j \sum_{|\alpha|=j-4} (-1)^{\alpha_1+\alpha_2} \int u^{(\alpha_1)} u^{(\alpha_2)} \overline{u u^{(\alpha_3)} u^{(\alpha_4)} u^{(\alpha_5)}} + u^{(\alpha_1)} u^{(\alpha_2)} \bar{u} \frac{d^{\alpha_3}}{dx^{\alpha_3}} \Big(\overline{u u^{(\alpha_4)} u^{(\alpha_5)}} \Big) dx \Big].$$

Proof. By Proposition A.2 below we have $\tilde{T}_6(z) = 12$  $+ 4$ . As above we compute (with $v = \bar{u}$)

12 
$$
\begin{aligned}
&= 12 \int_{x_1<x_2<x_3<y_1<y_2<y_2} \prod_{j=1}^{3} e^{2iz(y_j-x_j)} u(y_j)\overline{u(x_j)} dxdy \\
&= \int_{x_1,x_2<x_3<y_1,y_2,y_3} e^{2iz(y_j-x_j)} u(y_j)\overline{u(x_j)} dxdy \\
&= \frac{1}{(2\pi)^3} \int \int_{x_1,x_2<x_3<y_1,y_2,y_2} e^{2iz(y_1+y_2+y_3-x_1-x_2-x_3)-i\sum_{j=1}^{3}(\xi_j x_j - \eta_j y_j)} \prod_{j=1}^{3} \hat{u}(\eta_j)\overline{\hat{u}(\xi_j)} dxdyd\xi d\eta \\
&= \frac{-1}{(2\pi)^3 i} \int \int \frac{e^{-ix_3(\xi_1+\xi_2+\xi_3-\eta_1-\eta_2-\eta_3)}}{(2z+\xi_1)(2z+\xi_2)(2z+\eta_1)(2z+\eta_2)(2z+\eta_3)} dx_3 \prod_{j=1}^{3} \hat{u}(\eta_j)\overline{u(\xi_j)} d\xi d\eta \\
&= \frac{-1}{(2\pi)^3 i} \int_{\mathbb{R}^5} \frac{1}{(2z+\xi_1)(2z+\xi_2)(2z+\eta_1)(2z+\eta_2)(2z+\eta_3)} \prod_{j=1}^{3} \hat{u}(\eta_j)\overline{u(\xi_j)} d\xi_1 d\xi_2 d\eta.
\end{aligned}
$$

The next calculation is similar so we only point out the differences. In the first step we only symmetrize with respect to ξ_1 and ξ_2, and with respect to η_1 and η_2. This leads to a factor $\frac{1}{4}$ and an inner integral

$$
\int_{y_1<x_3} e^{2iz(x_3-y_1)+i(x_3(\xi_3-\eta_2-\eta_3)-y_1(\eta_1-\xi_1-\xi_2))} dy_1 dx_3 = \frac{i}{2z+(\eta_2+\eta_3-\xi_3)} \delta_{\xi_1+\xi_2+\xi_3-\eta_1-\eta_2-\eta_3}.
$$

This gives

4 
$$
= \frac{-1}{(2\pi)^3 i} \int_{\mathbb{R}^5} \frac{\hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)\overline{\hat{u}(\eta_1)\hat{u}(\eta_2)\hat{u}(\eta_3)}}{(2z-\xi_1)(2z-\xi_2)(2z-\eta_2)(2z-\eta_3)(2z-(\eta_2+\eta_3-\xi_3))} d\xi_1 d\xi_2 d\eta.
$$

The claimed formula for $H_{j,6}$ follows by an expansion in inverse powers of $2z$ and Plancherel. ☐

9. The KdV equation

The arguments in the proof of our results for NLS and mKdV carry over easily to the KdV equation. In effect the algebraic part of the analysis is virtually identical. For this reason we outline here the corresponding results for the KdV equation

$$
u_t + u_{xxx} - 6uu_x = 0 \tag{9.1}
$$

where we consider real solutions on the real line. The scaling for this equation is

$$
u(x,t) \to \lambda^2 u(\lambda x, \lambda^3 t)
$$

which corresponds to the critical Sobolev space $\dot{H}^{-\frac{3}{2}}$.

This is also a completely integrable flow, and admits an infinite number of conservation laws, of which the first several are as follows:

$$
\begin{aligned}
E_0 &= \int u^2 dx \\
E_1 &= \int u_x^2 + 2u^3 dx \\
E_2 &= \int u_{xx}^2 - 10uu_x^2 + 15u^4 dx
\end{aligned}
$$

In a Hamiltonian interpretation, these energies generate commuting Hamiltonian flows with the Poisson structure defined by

$$\omega(u,v)=\int uv_x dx.$$

The first of these flows is the group of translations, the second is the KdV, etc. As in the NLS/mKdV case we first state a simplified version of the main result, which does not require knowledge of the scattering transform. Later we will return with a more complete version, see Theorem 9.3, together with the appropriate trace formulas 9.2. We have

Theorem 9.1. *There exists $\delta>0$ so that for each $s\geq -1$ there exists an energy functional*

$$E_s : H^s \cap \{\|u\|_{H^{-1}} \leq \delta\} \to \mathbb{R}^+$$

with the following properties:

(1) *E_s is conserved along the KdV flow.*
(2) *We have*

$$\left|E_s(u)-\|u\|^2_{H^s}\right| \lesssim \|u\|_{H^{-1}}\|u\|^2_{H^{\max\{-1,s-\frac{1}{4}+\varepsilon\}}}. \tag{9.2}$$

(3) *E_s is continuous in s and $u\in H^\sigma$ for $-1\leq s\leq\sigma$, analytic in u, and analytic in both variables if in addition $s<\sigma$.*

With a little more work it is possible to choose $\varepsilon=0$ and we will indicate how this is done below.

As in the case of NLS and mKdV, we establish the energy conservation result 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 $s\geq-\frac{3}{4}$. If s is below this threshold, then the energy conservation property holds for all data at the threshold. It is not known whether the KdV equation is well-posed in the range $-1\leq s<-\frac{3}{4}$; however, it is known that local uniformly continuous dependence fails in this range, see [4], and also that the problem is ill-posed in H^s for $s<-1$, see [18].

Again, one consequence of our result is that if the initial data is in H^s then the solutions remain bounded in H^s globally in time. This is immediate if $\|u_0\|_{H^{-1}}\ll 1$, but also follows by scaling for larger data. This has been done by making use of the Miura map by [3] for $s=-1$ and independently making clever use of Fredholm determinants by [12] for $-1\leq s\leq 1$.

For the remainder of this section we outline the proof of this result, and connect it to the proof of our NLS and KdV results.

9.1. **The scattering transform for KdV.**

Here we recall some basic facts about the inverse scattering transform KdV. The Lax pair for KdV is given by the pair of operators $(\mathcal{L},\mathcal{P})$ defined as follows

$$\mathcal{L}=-\partial_x^2+u, \qquad \mathcal{P}=-4\partial_x^3+3(u\partial_x+\partial_x u)$$

The scattering transform associated to KdV is defined via the spectral problem for the operator $\mathcal{L}$, namely

$$\mathcal{L}\psi=z^2\psi \tag{9.3}$$

As before, we use this equation first for z on the real line, and then for z in the upper half-space. As u is real, it follows that the operator $\mathcal{L}$ is self-adjoint. Its continuous spectrum is the positive real line $\mathbb{R}^+$ and it may have isolated (but possibly infinite) negative eigenvalues, as well as a resonance at frequency 0, but no eigenvalues inside the continuous spectrum if u is a Schwartz potential.

To relate this to the corresponding NLS spectral problem we also rewrite it as a linear system for $(\psi_1, \psi_2) = (\psi, \psi_x + iz\psi)$, namely

(9.4)
$$\begin{cases} \dfrac{d\psi_1}{dx} = - iz\psi_1 + \psi_2 \\ \dfrac{d\psi_2}{dx} = iz\psi_2 + u\psi_1 \end{cases}$$

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

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

respectively

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

These are viewed as initial value problems with data at $-\infty$, respectively $+\infty$. Since the Wronskian $\det(\psi_l, \psi_r)$ is constant we can evaluate it at both sides and see that both the transmission coefficient and the reflection coefficients are the same on both sides.

Let $\phi(\xi)$ and $\phi(-\xi)$ denote the first component of $\psi_l(\pm\xi)$. They both satisfy

$$-\phi'' + u\phi = \xi^2\phi,$$

and hence their Wronskian

$$\det \begin{pmatrix} \phi(\xi) & \phi'(-\xi) \\ \phi(\xi) & \phi'(-\xi) \end{pmatrix}$$

is constant. We evaluate is on both sides. At $-\infty$ we obtain $2i\xi$ and at $+\infty$

$$\det \begin{pmatrix} a_+e^{-i\xi x} + b_+e^{i\xi x} & a_-e^{i\xi x} + b_-e^{-i\xi x} \\ -i\xi a_+e^{-i\xi x} + i\xi b_+e^{i\xi x} & i\xi a_-e^{i\xi x} - i\xi b_-e^{-i\xi x} \end{pmatrix} = 2i\xi(a_+a_- - b_+b_-)$$

and hence

$$T(\xi)T(-\xi) = 1 + R(\xi)R(-\xi).$$

Since

$$\psi_l(-\xi, x, t) = \overline{\psi_l(\xi, x, t)},$$

and hence $T(-\xi) = \overline{T(\xi)}$ and $R(-\xi) = \overline{R(\xi)}$, we arrive at

(9.5)
$$|T(\xi)|^2 = 1 + |R(\xi)|^2.$$

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. Thus for T we obtain a meromorphic extension, with poles only at those z on the positive imaginary axis for which z^2 is an eigenvalue for $\mathcal{L}$. We also note the symmetry

(9.6)
$$T(-\bar{z}) = \bar{T}(z).$$

As u evolves along the (9.1) flow, the functions L, R, T evolve according to

$$T_t = 0, \qquad L_t = -8i\xi^3 L, \qquad R_t = 8i\xi^3 R.$$

The scattering map for u is given by

$$u \to R,$$

and the map $u \to R$ conjugates the KdV flow (9.1) to (the Fourier transform of) the linear Airy flow. Reconstructing u from R requires solving a Riemann-Hilbert problem, see [6] for this approach for the Korteweg-de Vries equation.

One difference between the case of NLS-mKdV and that of KdV is that for the latter we need[3] $u \in L^1$ in order to define $T(\xi)$ in the upper half-plane. This is due to the linear term in T, which after one iteration is seen to have the form

$$T(z,u) = 1 - \frac{1}{2iz}\int_{\mathbb{R}} u(x)dx + \text{quadratic and higher terms}$$

However the linear effects are best understood at the level of $\log T$, for which not only we have the similar relation

$$\log T(z,u) = -\frac{1}{2iz}\int_{\mathbb{R}} u(x)dx + \text{quadratic and higher terms}$$

but the remainder term in this approximation can be defined only in terms of L^2 type norms of u, and no longer requires integrability for u. This can be viewed as a renormalization of $\log T$, and is discussed in detail in the next subsection.

For now we note one consequence of this, namely the bound

$$\operatorname{Re}(z^2 \log T) - \frac{1}{2}\operatorname{Im} z \int_{\mathbb{R}} u(x)\ dx \geq 0 \tag{9.7}$$

This is easily seen for Schwartz functions u, where the above expression is subharmonic in the upper half-space, nonnegative on the real axis and vanishing at infinity. Then by density it holds for all $u \in H^{-1}$.

9.2. **The transmission coefficient in the upper half-plane and conservation laws.** As in the NLS-mKdV case, the conserved energies are constructed using the transmission coefficient T in the upper half-space. Assuming in a first approximation that u is a Schwartz function, the transmission coefficient T will be a meromorphic function in the upper half-space, with poles in a compact subset of the positive imaginary axis. Further, $\ln|T|$ is a Schwartz function on $\mathbb{R} \setminus \{0\}$, and $\ln T$ has a formal series expansion as $z \to \infty$ in the upper half-space,

$$\ln T \approx -\frac{1}{2\pi i}\sum_{j=-1}^{\infty} E_j (2z)^{-2j-3}$$

where E_j are the conserved energies for $j \geq 0$, and

$$E_{-1} = 2\pi \int_{\mathbb{R}} u\ dx.$$

Here in view of the symmetry (9.6) we have only odd terms in the expansion.

In particular if there are no poles for T in the upper half-plane, then the E_j's can be expressed by Cauchy's theorem as

$$E_j = \int \xi^{2j+2} \ln|T| d\xi \tag{9.8}$$

We would like to define our conserved energies as

$$E_s = \int \xi^2 (1+\xi^2)^s \ln|T| d\xi, \qquad s \geq -1$$

[3] Or some other setting where $\int u$ is defined.

except that this formula would have a very restrictive range of validity, namely for Schwartz u for which $\mathcal{L}$ has no negative eigenvalues. To avoid this difficulty, we again switch the contour of integration to the positive imaginary line.

Precisely, if u satisfies the smallness condition $\|u\|_{H^{-1}} \ll 1$, the transmission coefficient T will be a holomorphic function in $H \setminus i[0,1/4]$. Then, as in the NLS-mKdV case, we define the energies E_s in the range

$$N \leq s < N+1, \quad N \geq -1$$

by

$$E_s(u) = -4\sin(\pi s) \int_1^\infty t^2(t^2-1)^s \left(\operatorname{Re} \ln T(it/2) + \sum_{j=-1}^{N} (-1)^j E_j t^{-2j-3} \right) dt + \sum_{j=0}^{N} \binom{s}{j} E_j, \tag{9.9}$$

where the last term accounts for the half-residues at zero of the corrections. Again, if u is Schwartz and T has no poles in the upper half-space then this agrees with the previous formula. However, if $\|u\|_{H^{-1}} \ll 1$ then this last expression is defined as an absolutely convergent integral for all $u \in H^s$.

We remark here that the E_{-1} contribution is subtracted for all values of s, which corresponds to the renormalization of $\log T$ previously alluded to.

Just as in the NLS/mKdV case, here we also have trace formulas relating the above energies to the corresponding integrals on the real line:

Proposition 9.2 (Trace formula)**.** *a) Let $s > -1$ and $N \geq [s]$. Define*

$$\Xi_s(t) = \int_0^t \zeta^2(1-\zeta^2)^s d\zeta, \qquad s > -1.$$

Then for $u \in \mathcal{S}$ we have

(9.10)

$$\begin{aligned} E_s &= \int_{\mathbb{R}} \xi^2(1+\xi^2)^s(\operatorname{Re}\ln T(\xi/2))d\xi + 2\sum_j \Xi(\kappa_j/2) \\ &= 4\sin(\pi s)\int_1^\infty \tau^2(\tau^2-1)^s\Big[\operatorname{Re}\ln T(i\tau/2) - \frac{1}{2\tau}\int u\,dx - \frac{1}{2\pi}\sum_{j=0}^{N}(-1)^j H_j \tau^{-2j-3}\Big]d\tau + \sum_{j=0}^{N}\binom{s}{j} H_{2j} \end{aligned}$$

where the sum runs over the eigenvalues $i\kappa_j$ of $\mathcal{L}$ on the positive imaginary axis.

b) In the case $s = -1$ we have, provided that $-1/4$ is not an eigenvalue of the Lax operator

$$\begin{aligned} E_{-1} &= \int_{\mathbb{R}} \frac{\xi^2}{1+\xi^2} \operatorname{Re}\ln T(\xi/2)d\xi + \sum(\ln(1+2\kappa_j) - \ln(1-2\kappa_j)) - 4\kappa_j \\ &= 8\left(\operatorname{Re}\ln T(i/2) - \int u dx\right). \end{aligned} \tag{9.11}$$

The proof of the trace formula is similar to the corresponding proof in the NLS/mKdV case, using Lemmas 4.2,4.3, with the notable difference that these are now applied to the nonnegative subharmonic function

$$G(z) = \operatorname{Re}\left(z^2 \ln T(z) - \frac{iz}{2}\int u\ dx\right). \tag{9.12}$$

We note that if $u \notin L^1$ then in the first integral in (9.10) one should replace $\xi^2 \operatorname{Re}\ln T(\xi/2)d\xi$ with the trace of G on the real line, which is defined as a bounded measure.

Now we are ready to state the complete form of our main result:

Theorem 9.3. *For each $s > -1$ and if $s = -1$ and $-1/4$ is not in the spectrum of the Lax operator $\mathcal{L}$, both sides of* (9.10) *resp.*(9.11) *are well-defined. They define a continuous map*

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

and

$$E_{-1} : H^{-1} \backslash \{u : -1/4 \text{ is a } \mathcal{L} \text{ eigenvalue } \} \to \mathbb{R}^+$$

with the following properties:

(1) *E_s is conserved along the KdV flow.*
(2) *The trace of the function G defined by* (9.12) *on the real line is defined as a locally finite measure on the real line, and the trace formulas in* (9.10) *and* (9.11) *hold.*
(3) *If $\|u\|_{H^s} \leq \delta$ then we have*

$$\left| E_s(u) - \|u\|_{H^s}^2 \right| \lesssim \|u\|_{H^{-1}} \|u\|^2_{H^{\max\{-1, s-\frac{1}{8}\}}}. \tag{9.13}$$

(4) *E_s is continuous in s and $u \in H^\sigma$ for $-1 \leq s \leq \sigma$, analytic in u if $-1/4$ is not an eigenvalue of the Schrödinger operator, and analytic in both variables if in addition $s < \sigma$.*
(5) *There exists $\delta > 0$ so that all eigenvalues of $\mathcal{L}$ are above $-\frac{1}{8}$ if $\|u\|_{H^s} \leq \delta$.*

The proof of the theorem follows the same outline as in the NLS/KdV case. The transition from the small data case $\|u\|_{H^{-1}} \ll 1$ and the large data case is based again on the trace formulas in the preceding proposition. Hence for the remainder of this section we discuss the small data case, where the bulk of the analysis is devoted to the estimates for the terms in the asymptotic expansion of $\log T$.

9.3. **Estimates for the transmission coefficient.** We begin with the formal homogeneous expansion of the transmission coefficient T, namely

$$T(z) = 1 + \sum_{j=1}^{\infty} T_{2j}(z)$$

where $T_{2j}(z)$ are multilinear integral forms, homogeneous of degree j in u. There is a similar though less explicit expansion for $\ln T$,

$$\ln T(z) = \sum_{j=1}^{\infty} \tilde{T}_{2j}(z)$$

We have $\tilde{T}_2 = T_2$, while $\tilde{T}_{2j}$ are still multilinear integral forms of degree j in u.

Here we take advantage of the similarity between the systems (9.4) and (2.5). Precisely, the multilinear forms T_{2j} and $\tilde{T}_{2j}$ are the same as before, with the only difference that these forms apply to the pair of functions $(1, u)$ rather than $(u, \pm\bar{u})$. In particular, we can still take advantage of the improved structure of $\tilde{T}_{2j}$.

We now discuss the successive terms in the $\ln T$ series.

9.3.1. *The role of the T_2 term.* This is linear in u, and equal

$$T_2(z) = -\frac{1}{2\pi i} E_{-1}(2z)^{-1} = \frac{i}{2z} \int u(x)dx$$

We note that such a term did not arise for NLS-mKdV, and it cannot be bounded by $\|u\|_{H^s}$. Thus the correct strategy is to view this as a renormalization term. Precisely, while $T_2(z)$ is not generally defined for $u \in H^s$, the expression $\ln T(z) - T_2(z)$ restricted to $i[1, \infty)$ will extend smoothly to all u which are small in H^{-1}. Indeed, we have

Lemma 9.4. *The map $u \to e^{-T_2(z)} T(z)$ is analytic for $\|u\|_{H^{-1}} \ll 1$ and $z \in i[1, \infty)$.*

Proof. We seek to solve the system

$$\begin{cases} \dfrac{d\psi_1}{dx} = - iz\psi_1 + \psi_2 \\ \dfrac{d\psi_2}{dx} = iz\psi_2 + u\psi_1 \end{cases} \tag{9.14}$$

and prove Lemma 9.4. If $u \in L^1 \cap H^{-1}$ then a direct iterative scheme yields the existence of a unique solution

$$(\psi_1, \psi_2) \in (\dot{W}^{1,1} \cap \dot{W}^{1,2}) \times (W^{1,1} \cap L^2)$$

with the property that

$$\lim_{x \to -\infty} (\psi_1, \psi_2) = (1, 0), \qquad \lim_{x \to -\infty} (\psi_1, \psi_2) = (T^{-1}(iz), 0).$$

Our goal here is to renormalize $T(i\tau)$, and show that the expression

$$S(i\tau) = T(i\tau)e^{-\frac{1}{\tau}\int u}$$

depends analytically only on u in H^{-1}. Setting $z = i\tau$, we will rewrite the equation in a more favorable form. First we multiply ψ_1 and ψ_2 by $e^{\tau x}$. By an abuse of notation we keep the notation for ψ_1 and ψ_2, which now satisfy

$$\begin{aligned} \psi_1' &= \psi_2 \\ \psi_2' &= - 2\tau\psi_2 + u\psi_1. \end{aligned} \tag{9.15}$$

Here we start with initial data $(1, 0)$ at $-\infty$, and $\psi_2(\infty) = 0$ while $\psi_1(\infty) = T(i\tau)^{-1}$.

To peel off the low regularity part of ψ_2 we split u into low and high frequencies and define

$$U = \int_{-\infty}^{x} e^{-2\tau(x-y)} u(y) dy, \qquad u_{lo} = \tau \int_{-\infty}^{x} e^{-(x-y)} u(y) dy.$$

Then $u = U_x + u_{lo}$ and

$$\hat{u}_{lo}(\xi) = \frac{1}{1 + i\xi/\tau} \hat{u}$$

so that we have

$$\|u\|_{H^{-1}} \leq \tau \|u_{lo}\|_{L^2} + \|U\|_{L^2} \leq 2\|u\|_{H^{-1}} \tag{9.16}$$

We now define renormalized variables (w_1, w_2) by

$$w_2 = e^{-\frac{1}{\tau}\int_{-\infty}^{x} u_{lo}} (\psi_2 - U\psi_1), \qquad w_1 = e^{-\frac{1}{\tau}\int_{-\infty}^{x} u_{lo}} (\psi_1 + \psi_2)$$

and rewrite the system in terms of w_1 and w_2 as

$$\begin{cases} w_1' = - (u_{lo} + U)w_2 - U^2(w_1 - w_2) \\ w_2' = - 2\tau w_2 - Uw_1 + u_{lo}(w_1 - 2w_2) - U^2(w_1 - w_2) \end{cases} \tag{9.17}$$

Here (w_1, w_2) start with the same data $(1, 0)$ at $-\infty$, but now $w_2(\infty) = 0$ while $w_1(\infty) = S(i)$. Hence it suffices to show that we can solve the equation (9.17) iteratively provided that $\|u\|_{H^{-1}} \ll 1$. Notably, we are no longer assuming that $u \in L^1$.

Let C_c be the space of continuous functions with limits at $\pm\infty$. We define

$$X = C_c \times l^2_\tau L^\infty \tag{9.18}$$

and the linear map $L : X \to X$ where $(v_1, v_2) = L(w_1, w_2)$ if (v_1, v_2) solve the linear system

$$\begin{aligned} v_1' &= - (u_{lo} + U)w_2 + U^2(w_1 - w_2) \\ v_2' &= - 2\tau v_2 + u_{lo}(w_2 - w_1) + Uw_1 + U^2(w_1 - w_2) \end{aligned}$$

with zero Cauchy data at $-\infty$. We can now recast the system (9.17) in the form

$$(w_1, w_2) = (1, 0) + L(w_1, w_2) \tag{9.19}$$

In order to solve this iteratively we need the following

Lemma 9.5. *We can decompose $L = L_1 + L_2$ where*
(i) L_1 is linear in u and (w_1, w_2) and satisfies

$$\|L_1\|_{X \to X} \lesssim \|u\|_{H^{-1}}.$$

(ii) The map L_2 is linear in (w_1, w_2) and quadratic in U, and satisfies

$$\|L_2\|_{X \to X} \le c\|U\|_{L^2}^2$$

Proof. By the first equation

$$\|v_2\|_{l^2 L^\infty} \le \Big(\|u_{lo}\|_{L^2} + \|U\|_{L^2} + \|U\|_{l^2_\tau L^2}^2\Big)\Big(\|w_2\|_{L^\infty} + \|w_1\|_{L^\infty}\Big)$$

and by the second

$$\|v_1\|_{L^\infty} \le \|u_{lo} + U\|_{L^2}\|w_2\|_{L^2} + \|U\|_{L^2}^2\|w_1\|_{L^\infty} + \|w_2\|_{L^\infty}.$$

□

Hence, if $\|u\|_{H^{-1}} \ll 1$ then $\|L\|_{X \to X} < 1$, which allows us to solve the equation (9.17) by

$$(w_1, w_2) = \sum_{j=0}^{\infty} L^j(1, 0)$$

Finally, we have

$$(S(i), 0) = \lim_{x \to \infty} L^j(1, 0)(x)$$

and the conclusion of Lemma 9.4 follows. □

Lemma 9.4 together with a rescaling implies that there is a constant C so that with

$$h(t) = 1 + t^2 e^t$$

$$e^{-T_2(i\tau)} T(i\tau) \preceq h(C\tau^{-1/2}\|u\|_{H_\tau^{-1}}).$$

where terms with like homogeneity are separately compared, and

$$\|u\|_{H_\tau^{-1}}^2 = \int (\tau^2 + \xi^2)^{-1} |\hat{u}(\xi)|^2 d\xi$$

Since $\ln(1 + t) \preceq \dfrac{t}{1 - t}$, this gives

$$\ln(e^{-T_2(i\tau)} T(i\tau)) \preceq \frac{h(C(\tau^{-1/2})\|u\|_{H_\tau^{-1}})}{1 - h(C(\tau^{-1/2})\|u\|_{H_\tau^{-1}})}.$$

Hence there exists $C > 0$ so that , for $j \ge 2$,

$$|\tilde{T}_{2j}| \le C(\tau^{-1/2}\|u\|_{H_\tau^1})^j. \tag{9.20}$$

9.3.2. *The $\tilde{T}_4$ term.* This is quadratic in u, and plays exactly the same role played by T_2 before. We have

$$\begin{aligned}\tilde{T}_4(z) =& -2\int_{x_1<x_2<y_1<y_2} e^{2iz(y_1+y_2-x_1-x_2)}u(x_1)u(x_2)dxdy \\ =& \frac{1}{4z^2}\int_{x_1<x_2} e^{2iz(x_2-x_1)}u(x_1)u(x_2)dx_1dx_2\end{aligned}$$

and hence, as for the NLS case,

$$E_{s,4} = \|u\|_{H^s}^2 \tag{9.21}$$

9.3.3. *Low regularity bounds for the $\tilde{T}_{2j}$ term, $j \geq 3$.* We note that unlike the case of NLS-mKdV, no such bound holds for $T_{2j}(i\tau)$, which contains contributions from the forbidden energy E_{-1}. From Lemma 9.4 we obtain

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-\frac{j}{2}}\|u\|^j_{H^{-1}_\tau} \leq \tau^{-\frac{3j}{2}}\|u\|^j_{L^2} \lesssim \tau^{-\frac{3j}{2}}\|u\|^2_{H^{\frac{j-2}{2}}}\Big(c\|u\|_{H^{-1}}\Big)^{j-2}. \tag{9.22}$$

The trivial bound

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-\frac{j}{2}}(c\|u\|_{H^{-1}_\tau})^j \leq \tau^{-\frac{j}{2}}(c\|u\|_{H^{-1}})^j$$

suffices for the pointwise estimate up to $s = \frac{j}{4} - \frac{3}{2}$. For $-\frac{3}{4} \leq s \leq \frac{3(j-2)}{4}$ we interpolate the two estimates to arrive at

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-3-2s}\|u\|^2_{H^{s-1/4}}\|u\|^{j-2}_{H^{-1}},$$

with the case $j = 3$ be the worst one.

The benefit in having such a bound is two-fold:

1. (1) It fully covers all $s \leq 0$; to account for larger s we will no longer have to deal with H^{-1} norms.
2. (2) For every $s > 0$ it accounts for all large enough j, so in the sequel we only need to obtain bounds for fixed j.

9.3.4. *Higher regularity bounds for the $\tilde{T}_{2j}$ term, $j \geq 3$.* Here we take advantage of the structure properties of $\tilde{T}_{2j}$, namely that it contains only connected integrals. Because of this we can use an unbalanced version of the bound (6.1), namely

$$|\tilde{T}_{2j}(i\tau)(u,v)| \lesssim \|u\|^j_{l^j_\tau DU^2}\|v\|^j_{l^\infty_\tau DU^2}. \tag{9.23}$$

We apply this with $v = 1$, which satisfies

$$\|1\|_{l^\infty_\tau DU^2} \lesssim \tau^{-1},$$

to obtain

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-j}\|u\|^j_{l^j_\tau DU^2}. \tag{9.24}$$

Using the Sobolev embedding (5.7), this allows us to extend the bound (9.22) to the range $-1 \leq s \leq \frac{1}{2} - \frac{1}{j}$. In particular we have the estimate

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-(2j-1)}\|u\|^j_{H^{\frac{1}{2}-\frac{1}{j}}}. \tag{9.25}$$

Then by interpolation we get the uniform bounds

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-(2j-1)}\|u\|^2_{H^{\frac{3(j-2)}{4}}}\|u\|^{j-2}_{H^{-1}_\tau} \tag{9.26}$$

and

$$|\tilde{T}_{2j}(i\tau)| \lesssim \tau^{-3-2s}\|u\|^2_{H^{s-\frac14}}\|u\|^{j-2}_{H^{-1}_\tau} \tag{9.27}$$

for $-1 \leq s \leq j-2$. In particular, for such s we obtain for all $\varepsilon > 0$

$$E_{j,s} \lesssim \|u\|^2_{H^{s-\frac14+\varepsilon}}\|u\|^{j-2}_{H^{-1}}. \tag{9.28}$$

Thus it remains to consider the case $s > j-2$.

9.3.5. *The expansion for the $\tilde{T}_{2j}$ term, $j \geq 3$.* Here the proof of Proposition 6.4 applies with minor but important changes to give

$$\left|\tilde{T}_{2j}(i\tau) - \sum_{l=0}^{N}(-1)^{j+l}T_{2j,l}\tau^{-2j-2l-1}\right| \lesssim \sum_{2N+1\leq|\alpha|\leq j-1+2N}^{\max\alpha_k\leq N+1} \tau^{-2j-1-|\alpha|}\|\partial^{\alpha_k}u\|_{l^j_\tau DU^2}. \tag{9.29}$$

This we seek to use in the range $j-2+N \leq s < j-1+N$. Indeed, for $j-2+N+\frac14 \leq s \leq j-1+N+\frac14$ the above bound implies that

$$|\tilde{T}_{2j}(i\tau) - \sum_{l=0}^{N}(-1)^{j+l}T_{2j,l}\tau^{-2j-2l-1}| \lesssim \tau^{-3-2s}\|u\|^2_{H^{s-\frac14}}\|u\|^{j-1}_{H^{-1}}. \tag{9.30}$$

Due to the room of $\frac14$ we easily obtain the integrated bound

$$\int_1^\infty \tau^{2s+2}|\tilde{T}_{2j}(i\tau) - \sum_{l=0}^{N}(-1)^{j+l}T_{2j,l}\tau^{-2j-2l-1}|d\tau \leq c\|u\|^2_{H^{s-\frac14+}}\|u\|_{H^{-1}}.$$

One can further refine the argument using a Littlewood-Paley decomposition, as in the proof of Proposition 6.4, to reach the $H^{s-\frac14}$ bound in the last estimate and in (9.28).

9.4. **The cubic term in the KdV energies.** For completeness we also provide the Fourier expression for the cubic term in $\ln T(z)$, namely $\tilde{T}_6(u)$, as well as the cubic term in our energy functionals E_s^3. These are given by specializing (8.3) and symmetrizing in ξ:

$$\tilde{T}_6(z) = \frac{2i}{\sqrt{2\pi}}\int_{\xi_1+\xi_2+\xi_3=0}\frac{1}{(2z)^2(2z-\xi_1)(2z-\xi_2)}\left(\frac{1}{2z}+\frac{1}{2z+\xi_3}\right)\hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)d\xi_1 d\xi_2 \tag{9.31}$$

and

$$E_{k,3} = 2^{-k-3}\sum_{\alpha_1+\alpha_2=2k-2}\int u^{(\alpha_1)}u^{(\alpha_2)}u\,dx + \sum_{\alpha_1+\alpha_2+2\alpha_3=2k-2}(-1)^{\alpha_3}\int u^{(\alpha_1)}u^{(\alpha_2)}u^{(\alpha_3)}dx. \tag{9.32}$$

To compute the cubic component E_s^3 of the energy it is useful to symmetrize in ξ

$$\tilde{T}_6(z) = \frac{4i}{3\sqrt{2\pi}}\int_{\xi_1+\xi_2+\xi_3=0}\frac{24z^3 - z(\xi_1^2+\xi_2^2+\xi_3^2)}{(2z)^2((2z)^2-\xi_1^2)((2z)^2-\xi_2^2)((2z)^2-\xi_3^2)}\hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)d\xi_1 d\xi_2 \tag{9.33}$$

Then we have

$$E_s^3 = \mathrm{Re}\,\frac{i}{3\sqrt{2\pi}}\int_{\mathbb{R}+i0}(1+z^2)^s\int_{\xi_1+\xi_2+\xi_3=0}\frac{24z^3 - z(\xi_1^2+\xi_2^2+\xi_3^2)}{((2z)^2-\xi_1^2)((2z)^2-\xi_2^2)((2z)^2-\xi_3^2)}\hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)d\xi_1 d\xi_2.$$

To express the complex conjugate of the above integral in a similar manner we use the fact that u is real, which implies that we have $\hat{u}(-\xi) = \bar{\hat{u}}(\xi_1)$. Then a straightforward computation yields

$$E_s^3 = -\frac{1}{3\sqrt{2\pi}}\int_{\mathbb{R}+i0}(1+z^2)^s\int_{\xi_1+\xi_2+\xi_3=0}\mathrm{Im}\,\frac{24z^3 - z(\xi_1^2+\xi_2^2+\xi_3^2)}{((2z)^2-\xi_1^2)((2z)^2-\xi_2^2)((2z)^2-\xi_3^2)}\hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)d\xi_1 d\xi_2$$

For the integrand above we have

$$K_\xi(z) = \operatorname{Im} \frac{24z^3 - z(\xi_1^2 + \xi_2^2 + \xi_3^2)}{((2z+i0)^2 - \xi_1^2)((2z+i0)^2 - \xi_2^2)((2z+i0)^2 - \xi_3^2)}$$
$$= \frac{\pi}{\xi_1\xi_2\xi_3} \left(\xi_1(\delta_{\xi_1/2} + \delta_{-\xi_1/2}) + \xi_1(\delta_{\xi_2/2} + \delta_{-\xi_2/2}) + \xi_3(\delta_{\xi_3/2} + \delta_{-\xi_3/2})\right)$$

Hence we obtain

$$(9.34) \quad E_3^s = \frac{2^{1-2s}\sqrt{2\pi}}{3} \int_{\xi_1+\xi_2+\xi_3=0} \frac{(4+\xi_1^2)^s\xi_1 + (4+\xi_2^2)^s\xi_2 + (4+\xi_3)^s\xi_3}{\xi_1\xi_2\xi_3} \hat{u}(\xi_1)\hat{u}(\xi_2)\hat{u}(\xi_3)d\xi_1 d\xi_2$$

where the latter formula agrees again with the I-method prediction, see [5, 16].

Appendix A. A Hopf algebra

The aim of this section is to frame the algebra of iterated integrals as a Hopf algebra, and then to take advantage of the Hopf algebra structures in order to prove Theorem 3.3. Our iterated integrals are linear combinations of integrals of the form

$$\int_\Sigma u(x_1)\overline{v(y_1)} \cdots u(y_j)\overline{v(y_j)} dx dy$$

where Σ represents a complete ordering of the variables x_l, y_l with the constraint that

$$\Sigma \subset \{x_l < y_l;\ l = 1, ..., j\}$$

Omitting the indices this can be represented as a word with two letters X and Y like

$$XXYY \qquad \text{for} \qquad x_1 < x_2 < y_1 < y_2.$$

Here we will restrict our attention to words which have an equal number of X's and Y's. Further, the requirement $x_l < y_l$ leads to the following additional property of our words: For any splitting of the word into two words, the left word contains at least as many X's as Y's.

To each word in this class we associate a graphical representation by nonintersecting arcs as follows. Replace X by ◜ and Y by ◝. Then there is exactly one way of connecting each ◜ arc by an ◝ arc so that the arcs do not intersect, for instance

$$XXYY \to$$ ◜◜◝◝ $\to$ ◠◠.

Thus by a slight abuse of language we call the words in H nonintersecting.

The shuffle product $\sqcup\!\sqcup$ on words in the alphabet $\{X, Y\}$ maps two words to the sum of all the words which are obtained by shuffling the two words, i.e. forming a word from the letters of the two words while respecting the ordering of both words. For example

$$X \sqcup\!\sqcup XY = 2XXY + XYX.$$

It defines a ring of formal power series with the words as unknowns and the shuffle product defining the commutative multiplication. We denote this ring by $H^{\sqcup\!\sqcup}$ and refer to Lothaire [17] for a more detailed discussion.

Here we will not work with the full ring $H^{\sqcup\!\sqcup}$. We denote the subset of $H^{\sqcup\!\sqcup}$ containing only formal series of nonintersecting words by H. The shuffle product of nonintersecting words can easily seen to be a sum of nonintersecting words. Hence H endowed with the shuffle product is a subalgebra of $H^{\sqcup\!\sqcup}$.

We define the length of a symbol to be half of the length of the nonintersecting word. A symbol of length j is identified with an integral over a subset of $\mathbb{R}^{2j}$ containing j factors u and j factors v. The length $2j$ introduces a grading on H which is compatible with the shuffle product: Words of length $2j$ have degree j. The elements of degree $> m$ form trivially an ideal I_m. The quotient $H_m = H/I_m$ is a finite dimensional algebra.

We identify nonintersecting words with integrals. For two given Schwarz functions u and v, any nonintersecting word defines an integral as above. For example

$$\overset{\frown}{\frown\frown} = \int_{x_1<x_2<y_1<x_3<y_2<y_3} \prod u(x_j)\overline{v(y_j)}dx_j dy_j.$$

For two nonintersecting words a and b of length $2n$ and $2m$ the product can written by an application of Fubini's Theorem as an integral over a domain U in $\mathbb{R}^{2(m+n)}$ given by the restrictions defined by the words a and b. We decompose U into completely ordered sets of variables, neglecting sets of measure zero. The simplest nontrivial example is

$$\left(\int_{x_1<y_1} u(x_1)\overline{v(y_1)}dx_1dy_1\right)^2 = 2\int_{x_1<y_1<x_2<y_2} u(x_1)\overline{v(y_1)}u(x_2)\overline{v(y_2)}dxdy$$
$$+ 4\int_{x_1<x_2<y_1<y_2} u(x_1)\overline{v(y_1)}u(x_2)\overline{v(y_2)}dxdy$$

At the level of words this becomes

$$XY ⧢ XY = 2XYXY + 4XXYY$$

and at the level of symbols

(A.1) $$\frown ⧢ \frown = 2\frown\frown + 4\overset{\frown}{\frown}.$$

Then the product of the integrals is the same as the sum over the integrals defined by the summands in the shuffle product of the two words. In short the shuffle product on nonintersecting words is compatible with the product of integrals. In abstract terms

Lemma A.1. *Let u and v be Schwartz functions. The evaluation of the integrals is a ring homomorphism from the finite sums in H to $\mathbb{C}$.*

The product of integrals introduces obvious relations: Linear combinations of monomials of the same homogeneity are equivalent to zero, if they vanish for every choice of functions $u, v \in \mathcal{S}$. We do not know whether this leads to nontrivial relations in H.

A fundamental subset of the set of symbols is the set of connected symbols. We call a symbol connected if there is an arc from the first to the last letter. For instance $\frown$, $\overset{\frown}{\frown}$, $\overset{\frown}{\frown\frown}$ are connected, while $\frown\frown$ is not. By extension we call an integral connected if the symbol is connected.

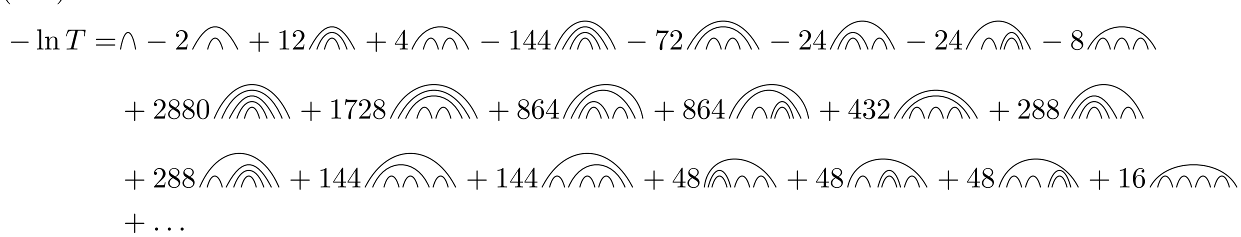
Our interest comes from the transmission coefficient T, which can be expressed as a power series (see Lemma 3.1)

$$T^{-1} = 1 + \frown + \frown\frown + \dots$$

Using the Taylor series for the log and the shuffle product, this also gives a series for $\ln T$. An involved calculation gives

Proposition A.2. *The following formula holds*

(A.2)



by which we mean that the difference lies in I_5.

The remarkable property is that the right hand side can be written as a sum of multiples of connected symbols. Indeed, we have

Theorem A.3. $\ln T$ *is a formal sum of connected integrals.*

To prove this we need more structure, namely a notion of a coproduct and a bialgebra, or even a Hopf algebra. Let $\tilde{H}$ be a graded algebra so that $\tilde{H}_m$ is always finite dimensional. A coproduct is a map

$$\Delta : \tilde{H} \to \tilde{H} \otimes \tilde{H}$$

which is coassociative and codistributive (see [21]).

An important standard example here is the shuffle algebra $H^{\sqcup\!\sqcup}$ of words in the alphabet $\{X, Y\}$ as above, see [20, 17] for the facts about the shuffle algebra used below. The shuffle product is known and easily seen to be commutative, associative and distributive. The coproduct Δ on $H^{\sqcup\!\sqcup}$ is defined on words as the sum over all splittings

$$\Delta a = \sum_{a_1 a_2 = a} a_1 \otimes a_2. \tag{A.3}$$

The coproduct of an intersecting symbol is a linear combination of tensor products of symbols with at least one intersecting factor. The coproduct of a nonintersecting symbol will contain some tensor products of symbols with nonintersecting factors, but also some tensor products of symbols with intersecting factors. In the latter case, the first factor must contain more X's, and the second more Y's. Then it is natural to define the coproduct on H as the nonintersecting part of the $H^{\sqcup\!\sqcup}$ coproduct. Thus the formula (A.3) still applies, but with a, a_1 and a_2 restricted to nonintersecting factors. As such, we note that H is *not* a subalgebra of $H^{\sqcup\!\sqcup}$.

The shuffle on the tensor product is defined in a natural fashion

$$(a \otimes b) \sqcup\!\sqcup (c \otimes d) = (a \sqcup\!\sqcup c) \otimes (b \sqcup\!\sqcup d).$$

The coproduct of $H^{\sqcup\!\sqcup}$ is coassociative: If

$$\Delta^{\sqcup\!\sqcup} a = \sum b_i \otimes c_i$$

then

$$\sum (\Delta^{\sqcup\!\sqcup} b_i) \otimes c_i = \sum b_i \otimes (\Delta^{\sqcup\!\sqcup} c_i).$$

This implies the same property for the coproduct on H. The product and the coproduct of a Hopf algebra satisfy the crucial compatibility condition

$$\Delta(a \sqcup\!\sqcup b) = \Delta a \sqcup\!\sqcup \Delta b, \tag{A.4}$$

see Reutenauer [20], Proposition 1.9.

This is again inherited by H from $H^{\sqcup\!\sqcup}$, but it is less obvious. The point is that if a and b are both nonintersecting, then nonintersecting terms in $\Delta a \sqcup\!\sqcup \Delta b$ can only arise from nonintersecting terms in Δa combined with nonintersecting terms in Δb; else, we have too many X's in the first factor and too many Y's in the second.

Now that we have the appropriate set-up, we return to our goal of proving Theorem A.3. We call an element $g \in H$ group-like if

$$\Delta g = g \otimes g.$$

Lemma A.4. *The set G of all group-like elements endowed with the shuffle product is a group.*

Proof. If $g, h \in G$ then

$$\Delta(g \sqcup\!\sqcup h) = \Delta g \sqcup\!\sqcup \Delta h = (g \otimes g) \sqcup\!\sqcup (h \otimes h) = (g \sqcup\!\sqcup h) \otimes (g \sqcup\!\sqcup h).$$

Every element of G starts with 1. The inverse can be determined recursively: If

$$g = 1 + g_1 + g_2 \dots$$

then

$$h = 1 - g_1 + g_1 \sqcup\!\sqcup\, g_1 - g_2 \dots$$

□

In particular, we note that

Lemma A.5. *The expression*

$$T = 1 + \cap + \quad + \qquad + \cdots$$

belongs to G.

Proof. Denoting by $\cap^{[k]}$ the expression obtained concatenating k $\cap$'s, we have

$$T = \sum_{n=0}^{\infty} \cap^{[n]}$$

On the other hand, we have

$$\Delta T = \sum_{n=0}^{\infty} \Delta \cap^{[n]} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \cap^{[k]} \otimes \cap^{[n-k]} = T \otimes T$$

□

The linear subspace of H where we seek to place $\ln T$ can be described as follows:

Definition A.6. *The primitive elements are*

$$P = \{p \in H : \Delta p = 1 \otimes p + p \otimes 1\}.$$

Indeed, we have

Lemma A.7. *Primitive elements are formal linear combinations of connected symbols.*

Proof. Let us call linear combinations of connected symbols indecomposable. Then trivially, indecomposible elements are primitive. Consider now a primitive element p, and show that it is idecomposable. We argue by contradiction. Let a be the homogeneous part of p of lowest grading m which is not a linear combination of connected symbols. Subtracting them, we may assume that a is a part of lowest grade in p which does not vanish. Since Δ preserves the grading (we equip the tensor product with the natural grading), we must have

$$\Delta a = 1 \otimes a + a \otimes 1.$$

It remains to deduce that a is a linear combination of connected symbols of the same degree. After subtracting multiples of connected symbols we reduce the problem to proving that if a is homogeneous, primitive and a linear combination of non connected symbols than it has to be 0. This in turn is a consequence of injectivity of Δ. □

Given the last two lemmas, Theorem A.3 is a consequence of the following more general result:

Theorem A.8. *The above subgroup G and the subspace P of H are related by*

$$G = \exp P$$

Proof. We first show that the exponential of any primitive H element must belong to G,

$$\exp P \subset G.$$

Indeed, for primitive $h \in P$ we can write

$$\exp(h) = \sum_{n=0}^{\infty} \frac{h^n}{n!}$$

and thus

$$\begin{aligned}\Delta(\exp(h)) &= \sum_{n=0}^{\infty} \Delta \frac{h^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(\Delta h)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(h \otimes 1 + 1 \otimes h)^n}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{h^k \otimes h^{n-k}}{k!(n-k)!} \\ &= \exp(h) \otimes \exp(h)\end{aligned}$$

We will show by induction that for all $n \in \mathbb{N}$ we have

$$G/I_m = (\exp P)/I_m.$$

Since

$$1 = \exp(0)$$

the formula holds in grade $m = 0$. Suppose now that $G/I_m = \exp p/I_m$. We want to prove that then $G/I_{m+1} = \exp P/I_{m+1}$. Clearly

$$\exp P/I_{m+1} \subset G/I_{m+1}$$

Let $g \in G$. By the induction hypothesis, there exists $p \in P$ so that

$$g - \exp p \in I_m.$$

Let $h = \exp(-p) ⧢ g$. It satisfies

$$h - 1 \in I_m$$

and, since it is group-like as product of group-like elements,

$$\Delta h = h \otimes h.$$

Let h_{m+1} be the part of grade $m+1$ of h. Then identifying the terms of grade $m+1$ in the last identity we get

$$\Delta h_{m+1} = 1 \otimes h_{m+1} + h_{m+1} \otimes 1$$

Thus $h_{m+1} \in P$ and

$$h - \exp(1 + h_{m+1}) \in I_{M+1}.$$

□

We are now in a position to complete the proof of Theorem A.3. By Lemma A.5 we have $T \in G$. Theorem A.8 ensures that there is a primitive element h so that $g = \exp h$. Checking the formal power series shows that h is unique. Its homogeneous parts are linear combinations of connected symbols. Thus Theorem A.3 follows.

Appendix B. The spaces U^p and V^p

The spaces U^2 and V^2, respectively DU^2 and DV^2, are crucially used in this article as substitutes for $\dot H^{\frac12}$, respectively the scale invariant space $\dot H^{-\frac12}$. The latter spaces are plagued by the failure of Sobolev embeddings and the failure of a good regularity theory when integration against those functions. Instead, the former spaces are close neighbors, and have both the same scaling and good multiplicative properties.

The aim of this section is set up the functional context for the estimates in this paper, and in particular to clarify the relation between the pointwise definition of U^p and V^p, and their interpretation as distributions. We first recall the definition of the spaces U^p and V^p.

We will consider functions u, v defined on the open interval (a,b) and set always $u(a) = 0$ and $v(b) = 0$, even if $a = -\infty$ and/or $b = \infty$. All functions in this section are ruled functions i.e. functions which have left and right limits everywhere play a central role. If X is a suitable function space we denote by X_{rc} resp. X_{lc} the subset of right (left) continuous functions with limit 0 at the left (right) endpoint.

Let $-\infty \le a < b \le \infty$. A partition τ is a finite monotone sequence

$$a < t_1 < \dots t_N = b.$$

We denote the set of partitions by $\mathcal{T}$. There is the obvious notion of refinements of partitions. Step functions are functions associated to a partition which are constant on the open intervals between points of the partition. We denote by $\mathfrak{S}$ the space of step functions, and by $\mathfrak{S}(\tau)$ the space of step functions associated to a given partition τ.

Definition B.1. *a) Let $1 < p < \infty$. We define the space $V^p = V^p(a,b)$ as the space of those functions in (a,b) for which the following norm is finite:*

$$\|v\|_{V^p} = \sup_{\tau \in \mathcal{T}} \left(\sum_j |v(t_{j+1}) - v(t_j)|^p \right)^{\frac1p}$$

where we set $v(b) = 0$.

b) A U^p atom is a step function

$$u = \sum_j \phi_j \chi_{[t_j, t_{j+1})}(x) \qquad \text{if } \sum |\phi_j|^p \le 1.$$

A $U^p(a,b)$ function is an l^1 sum of atoms. We equip U^p with the obvious norm.

The spaces U^p and V^p are invariant under continuous monotone coordinate changes and we suppress the interval in the notation unless we specifically need it. We note that by this definition the U^p functions are right continuous and have limit 0 at the left end point a, whereas V^p functions are just ruled, and vanish by definition at b. The supremum norm is bounded by the V^p norm (taking the partition $\{t, b\}$) and the U^p norm (this is checked on atoms). Moreover, if $p < q$

$$U^p \subset U^q, \qquad V^p \subset V^q$$

with norm estimates with constant 1 and $U^p \subset V^p$ with

$$\|u\|_{V^p} \le 2^{\frac1p} \|u\|_{U^p}.$$

It is not hard to see but less obvious that the V^p functions are ruled.

There is a natural pairing between ruled functions v and right continuous step functions u with $u(a) = 0$ (and the notation $t_0 = a$)

$$B(v,u) = \sum_{j=1}^{N-1} v(t_j)(u(t_j) - u(t_{j-1})) = \sum_{j=1}^{N-1} u(t_j)(v(t_j) - v(t_{j+1}))$$

where the sum runs over the points of the partition and $t_0 = a$. The two sums are equal since $u(a) = 0$ and $v(b) = 0$. By an abuse of notation we use the suggestive notation

$$B(v,u) = \int v du.$$

Suppose now that v is a left continuous step function associated to a different partition $\sigma = \{s_j\}$. Then a simple algebraic computation shows that

$$B(v,u) = \sum_{j=1}^{N-1} v(s_j)(u(s_j) - u(s_{j-1})) = \sum_{j=1}^{N-1} u(s_j)(v(s_j) - v(s_{j+1}))$$

Let $\frac{1}{p} + \frac{1}{q} = 1$ and let u be an U^q atom and $v \in V^p$. By definition

$$|B(v,u)| \leq \|v\|_{V^p}.$$

It is not surprising but not entirely obvious that B extends to a bilinear map from $V^p \times U^q \to \mathbb{R}$ with

$$|B(v,u)| \leq \|v\|_{V^p} \|u\|_{U^q}, \tag{B.1}$$

see [9]. Thus B induces a map from $V^p \to (U^q)^*$. Further, we have

Proposition B.2. *The bilinear form B induces an isometric isomorphism from $V^p \to (U^q)^*$.*

Proof. We briefly sketch the proof and refer to [9] for more details. We first observe that

$$\|v\|_{V^p} = \sup_{\|u\|_{U^q} \leq 1} B(u,v),$$

which is easily seen by restricting u to the class of U^p atoms. Therefore, B induces an isometry from $V^p \to (U^q)^*$. To see that this map is in effect an isomorphism, let $L : U^q \to \mathbb{R}$ be a linear functional. To it we associate the function

$$v(t) = L(\chi([t,b)).$$

Then

$$B(v, \chi(t,b)) = v(t) = L(\chi([t,b)))$$

and hence the same identity holds it we replace the characteristic function by right continuous step functions, which are dense. □

Remark B.3. *By symmetry, we can use the same bilinear form B to also pair left continuous U^q functions v with ruled V^p functions u with the same boundary conditions.*

The duality implies characterizations if the V^p norm:

$$\|v\|_{V^p} = \sup\{B(v,u) : \|u\|_{U^p} = 1\}$$

It is immediate that we even may restrict the supremum to step functions, and even atoms. Moreover, of the supremum over atoms on the right hand side is finite then $v \in V^p$. Similarly

$$\|u\|_{U^q} = \sup\{B(v,u) : \|v\|_{V^p} = 1\}$$

and again we may restrict the supremum on the right hand side to step functions. We will see that right continuous functions for which the right hand side is bounded when we take the supremum over step functions are indeed in U^q. But this time the proof is highly nontrivial.

Theorem B.4. *Suppose the right continuous ruled function u with $\lim_{t\to a} u(t) = 0$ satisfies*

$$\sup |B(v,u)| < \infty$$

where the supremum is taken over all step functions in V^p_{lc} (left continuous step functions in V^p) with norm at most 1 and compact support in (a,b). Then $u \in U^q$ and

$$\|u\|_{U^q} = \sup B(v,u).$$

Before we turn to the proof we study the behavior of these spaces under decomposition of the interval. Here we include atoms with $t_1 = a$ and denote those spaces by $\tilde{U}^p$.

Lemma B.5. *The following inequality holds for all functions in $\tilde{U}^p$.*

$$\|u\|^p_{\tilde{U}^p(T_0,T_1)} \le \|u\|^p_{\tilde{U}^p(T_0,t)} + \|u\|^p_{\tilde{U}^p(t,T_1)}$$

Proof. Let b and c be p atoms and $\lambda, \mu \in \mathbb{C}$. Then

$$a = \frac{\lambda}{(|\lambda|^p + |\mu|^p)^{\frac{1}{p}}} b + \frac{\mu}{(|\lambda|^p + |\mu|^p)^{\frac{1}{p}}} c$$

is a p atom on $U^p(T_1,T_2)$. Now let

$$u = \sum \lambda_j b_j, v = \sum \mu_j c_j$$

and $L = \sum |\lambda_j| \le C\|u\|_{U^p}$, $M = |\mu_j| \le C\|v\|_{U^p}$. We can find such a decomposition by definition for every $C > 1$. Then (with an extension by 0)

$$\sum_j \lambda_j a_j + \sum_l \mu_k c_l = \sum_{jl} \frac{\lambda_j |\mu_l|}{M} a_j + \frac{|\lambda_j| \mu_l}{L} c_l$$

and hence

$$\begin{aligned}
\|u+v\|_{U^p} \le & \sum_{j,l} |\lambda_j||\mu_l| \left\| \frac{\lambda_j}{|\lambda_j|} \frac{a_j}{M} + \frac{\mu_l}{|\mu_l|} \frac{c_l}{L} \right\|_{U^p} \\
\le & LM \left(M^{-p} + L^{-p} \right)^{1/9} \\
= & (L^p + M^p)^{\frac{1}{p}} \\
\le & C^{2p} \left(\|u\|^p_{U^p} + \|v\|^p_{U^p} \right)^{\frac{1}{p}}
\end{aligned}$$

This inequalities holds for all $C > 1$ and hence also for $C = 1$. □

There is an analogous statement for functions in V^p.

Lemma B.6. *Let $v \in V^p$, τ be a partition and suppose that v vanishes at the points of the partition. Then*

$$\|v\|^p_{V^p(T_0,T_1)} \le 2^{p-1} \sum \|v\|^p_{V^p(t_j,t_{j+1})}$$

Proof. Let $\tilde{\tau}$ be another partition and

$$I = \sum |v(\tilde{t}_{j+1}) - v(\tilde{t}_j)|^p$$

We may assume that $\tilde{\tau}$ contains a point in all the intervals of τ. Otherwise we omit that interval from τ. Taking the coarsest joint refinement leads to the second factor on the right hand side. It may decrease the sum at most by the factor 2^{1-p}. □

To left, respectively right continuous functions and partitions we associate step functions as follows:

Definition B.7. *Given a left continuous function v and a partition τ we define an associated left continuous step function v_τ so that*

$$v_\tau(t) = v(t)$$

for each point of the partition. Similarly, for each right continuous function v we define an associated right continuous step function v_τ.

Then the following lemma is straightforward:

Lemma B.8. *Suppose that v is a left continuous step function with partition τ. Then for all right continuous functions u we have*

$$B(v, u) = B(v, u_\tau)$$

Symmetrically, if u is a right continuous step function then for all left continuous functions B we have

$$B(v, u) = B(v_\tau, u).$$

Moreover

$$\|v_\tau\|_{V^p} \leq \|v\|_{V^p}$$

and

$$\|u_\tau\|_{U^p} \leq \|u\|_{U^p}.$$

Lemma B.9. *Let v_j be a sequence of step functions with*

$$\|v_j\|_{V^p} \leq 1 \qquad \text{and } \|v_j\|_{sup} \to 0.$$

Then there exists a subsequence v_{j_l} so that

$$\left(\Big\| \sum_{l=1}^{m} v_{j_l} \Big\|_{V^p}^p \right)^{\frac{1}{p}} \leq (2m)^{\frac{1}{p}}.$$

Proof. We construct the subsequence and a partition recursively. Suppose we have defined j_l for $l \leq m$ so so that $v^m = \sum_{l=1}^m v_{j_l}$ satisfies $\|v^m\|_{V^p} =\leq (2m)^{\frac{1}{p}}$. We search for j_{m+1}. Let τ be a partition. Then

$$\sum |(v^m+v_j)(t_{k+1})-(v^m+v_j)(t_k)|^p \leq \sum (|v^m(t_{k_{l+1}})-v^m(t_{k_l})|+2\|v_j\|_{sup})^p + \sum |v_j(t_{k+1})-v_j(t_k)|^p$$

where the first sum runs only over those indices for which v^m has a jump. This implies the statement if we choose j sufficiently large. □

Proof of Theorem B.4. Let

$$\|u\|_{\tilde U^p} = \sup\{B(v, u) : v \in \mathfrak{S}_{lc}, \|v\|_{V^p} = 1\}$$

and let $\tilde U^p$ be the Banach space of all ruled, right continuous functions for which this norm is finite. We have to show that $\tilde U^p = U^p$. By the duality property $(U^p)^* = V^q$ we know that $U^p \subset \tilde U^p$, with the norm bound

$$\|u\|_{\tilde U^p} \lesssim \|u\|_{U^p}$$

Further, we claim that equality holds. This is because step functions are dense in U^p, and for step functions it suffices to test them with other step functions. It remains to prove the converse inclusion. We begin with an elementary but crucial observation.

Lemma B.10. *Let $u \in \tilde{U}^p(a,b)$. There exists a step function $v \in V^q_{lc}$ with $\|v\|_{V^q} = 1$ with compact support in (a,b) so that*

$$B(v,u) \geq \frac{1}{4}\sup\{B(w,u) : w \text{ is a } V^p_{lc} \text{ step function with } \|w\|_{V^p} = 1\}.$$

Proof. There exists a step function $v_0 \in V^q_{lc}$ of norm 1 so that

$$B(v_0,u) \geq \frac{3}{4}\sup\{B(w,u) : w \in V^q_{lc} : \|w\|_{V^p} = 1\}. \tag{B.2}$$

Let $\tau = (a, t_1 \dots t_{N-1}, b)$ be the points of the partition associated to v_0. By the left continuity condition, v_0 already vanishes near $t = b$. We add one extra point, $t_0 \in (a,t_1)$ and define v_1 to be 0 left of t_0 and v_0 in $[t_0,b]$. Then

$$B(u, v_0 - v_1) = -\, v_0(t_1)B(u, \chi_{a,t_0}) = -v_0(t_0)u(t_0)$$

By right continuity

$$\lim_{t_0 \to a} u(t_0) = 0$$

so we can choose t_0 so that (B.2) holds for $v = v_1$. Since $\|v_1\|_{V^q} \leq 2^{\frac{1}{q}}$ this implies the result. □

Assume by contradiction that $U^p \not\subset \tilde{U}^p$. Then there exists $u \in \tilde{U}^p$ so that

$$\|u\|_{\tilde{U}^p} = 1, \qquad \operatorname{dist}(u, U^p) > \frac{1}{2}. \tag{B.3}$$

Exactly one of the following two alternatives holds for this function u:

(1) There exists $\varepsilon > 0$ and $t \in [a,b]$ so that for any interval (t_0,t_1) containing t we have

$$\|u - u(t)\|_{\tilde{U}^p(t,t_1)} + \|u - u(t_0)\|_{\tilde{U}^p(t_0,t)} \geq \varepsilon$$

with the obvious modification at the endpoints.

(2) For every $\varepsilon > 0$ and every $t \in [a,b]$ there exists a neighborhood (t_0,t_1) of t so that

$$\|u - u(t)\|_{\tilde{U}^p(t,t_1)} + \|u - u(t_0)\|_{\tilde{U}^p(t_0,t)} < \varepsilon$$

with the obvious modification at the endpoints.

We will show that neither alternative is true, thereby reaching a contradiction.

Suppose that the first alternative holds. Then we must have either

$$\|u - u(t)\|_{\tilde{U}^p(t,t_1)} \geq \varepsilon/2 \qquad \text{for all } t_1 > t \tag{B.4}$$

or

$$\|u - u(t_0)\|_{\tilde{U}^p(t_0,t)} \geq \varepsilon/2 \qquad \text{for all } t_0 < t$$

In either case we claim that there exists a sequence of step functions v_j in V^q with disjoint support so that

$$B(u, v_j) \geq \varepsilon/8, \qquad \|v_j\|_{V^p} = 1$$

Without loss of generality let us assume (B.4). By Lemma B.2 there exists $v_1 \in \mathfrak{S}_{lc}(t,t_1)$, compactly supported, so that

$$B(u - u(t_0), v_1) = B(u, v_1) \geq \varepsilon/8$$

Let t_2 be the first point of the partition of v_1 to the right of t. Then v_1 vanishes in (t,t_1). We repeat the above argument in (t,t_2) to produce a function $v_2 \in \mathfrak{S}_{lc}(t,t_1)$, etc. Let

$$v = \sum_{j=1}^{N} v_j.$$

Then

$$B(u,v) \geq N\frac{\varepsilon}{8},$$

while by Lemma B.6 we have

$$\|v\|_{V^q} \leq 2^{\frac{1}{q}} N^{\frac{1}{p}}$$

Hence

$$N\frac{\varepsilon}{8} \leq (2N)^{\frac{1}{q}}.$$

This cannot be true for N large. This is a contradiction, which shows that the first alternative cannot hold for any function $u \in \tilde{U}^p$.

It remains to disprove the second alternative. Here we will use the fact that u satisfies (B.3). For each $\epsilon > 0$ we can cover the interval $[a,b]$ with intervals

$$[a,b] = \bigcup_{t\in[a,b]} I_t$$

where I_t are the intervals given by the second alternative. By compactness, it follows that there is a finite subcovering,

$$[a,b] = \bigcup_{k=1}^{N} I_{t_k}$$

Then we can find a partition $\tau = \tau(\varepsilon)$ of $[a,b]$ with the property that for each interval $I_j(\tau) = [t_j, t_{j+1}]$ in the partition we have

$$\|u - u(t_j^\varepsilon)\|_{\tilde{U}^p(I_j(\tau))} \leq \epsilon \tag{B.5}$$

We now consider the right continuous step function u_τ which matches u at the points of the partition τ. The function $u - u_\tau$ vanishes on the partition τ, so it is natural to split it with respect to the partition intervals $I_j(\tau)$,

$$u - u_\tau = \sum u_j, \qquad u_j = 1_{I_j(\tau)}(u - u_\tau)$$

On one hand, by (B.3) we have

$$\|u - u_\tau\|_{\tilde{U}^p} > \frac{1}{2} \tag{B.6}$$

On the other hand, as a consequence of Lemma B.5 we have

$$\|u - u_\tau\|_{\tilde{U}^p}^p \leq \sum 2\|u_j\|_{\tilde{U}^p}^p. \tag{B.7}$$

For each u_j we can find a corresponding function $v_j \in \mathfrak{S}_{lc}$, with similar support, so that

$$B(u_j, v_j) \geq \frac{1}{4}\|u_j\|_{\tilde{U}^p}, \qquad \|v_j\|_{V^p} = 1.$$

Combining these functions with appropriate weights we produce the step function $w \in \mathfrak{S}_{lc}$,

$$w = \left(\sum_j \|u_j\|_{\tilde{U}^p}^p\right)^{-1} \sum_j \|u_j\|_{\tilde{U}^p}^{p-1} v_j$$

These will have the following three properties:

(1) Boundedness in V^q. Indeed, by Lemma B.6 we have

$$\|w\|_{V^q}^q \leq 2^{q-1} \left(\sum_j \|u_j\|_{\tilde{U}^p}^p\right)^{-1} \sum_j \|u_j\|_{\tilde{U}^p}^p \|v_j\|_{V^q}^q \leq 2^{q-1}.$$

(2) Large $B(w,u)$. Here we compute

$$B(w,u) = B(w, u - u_\tau) = \sum B(v_j, u_j) \geq \sum_j \|u_j\|_{\tilde{U}^p}^p \geq 2^{-p}$$

(3) Small pointwise norm. For each v_j we can bound the pointwise norm by 1, and they are all nonoverlapping. Thus using (B.5) we obtain

$$\|w\|_{L^\infty} \leq \sup_j \|u_j\|_{\tilde{U}^p}^{p-1} \left(\sum_j \|u_j\|_{\tilde{U}^p}^p \right)^{-1} \leq \frac{1}{4} \sup_j \|u_j\|_{\tilde{U}^p}^{p-1}.$$

Taking a sequence $\epsilon_j \to 0$, this construction yields a sequence of partitions τ_j and associated left continuous step functions w_j with the following properties:

(1) $\|w_j\|_{V^p} \lesssim 1$.
(2) $B(u, w_j) \gtrsim 1$.
(3) $\|w_j\|_{L^\infty} \to 0$.

To conclude we argue as in the first alternative, this time using Lemma B.9 to reach a contradiction.

□

Let

$$V^p_C = \{v \in V^p \cap C := V^p_{rc} \cap V^p_{lc} : \lim_{t \to a,b} v(t) = 0\}$$

Theorem B.11. *V^p_C is weak* dense in V^p. Moreover, if $v, w \in V^p$ then there exists a sequence $v_j \in V^p_C$ so that $v_j w \to vw$ in V^p in the weak* sense.*

Proof. We claim that $V^p_{lc} \subset V^p_C$ is weak * dense. Let $v \in V^p_C$. There exists an at most countable set of times t_j at which v is not left continuous. We include b if $\lim_{t\to b} v(t) \neq 0$. At each point t_j there is a strictly increasing sequence $t_{j,k}$ converging to t_j such that either t_j is contained in one of the intervals

$$[t_{l,k}, t_l]$$

for some $l < j$, or $[t_{j,k}, t_k]$ is disjoint from all those intervals. For each k we recursively modify v in $(t_{j,k}, t_j)$ so that in that interval

$$v_k(t) = \frac{t_j - t}{t_j - t_{j,k}} v(t_{j,k}) + \frac{t - t_{j,k}}{t_j - t_{j,k}} v(t_j)$$

if t_j is not contained in the previous intervals, with a simple modification at b. Then it is not hard to see that

$$\|v_k\|_{V^p} \leq \|v\|_{V^p}$$

and

$$B(v_k, u) \to B(v, u)$$

for all $u \in U^q$.

If $a = -\infty$ and $b = \infty$ (which we assume without loss of generality) and if $v \in V^p_{lc}$ and if a is a right continuous step function then

$$\lim_{h>0, h\to 0} B(v(.+h), a) = B(v, a).$$

Using the atomic decomposition we see that this convergence is true for $u \in U^q$ instead of a.

We pick a smooth function ρ with support in $(-1, 0)$ and integral 1 and define for $v \in V^p_{lc}$ and $h \in (0,1)$

$$v_h(t) = \int_0^1 v(t + hs)\rho(s) ds$$

Then again as $h \to 0$

$$B(v_h, u) \to B(v, u).$$

Since $u(t) \to 0$ and $t \to a$ we can truncate on the left and obtain functions in C_0^∞ converging to v in the weak* sense. The second claim is proven similarly. □

Theorem B.12. *The map*

$$u \to (v \to B(v, u))$$

is an isometric isomorphism from U^q to $(V_C^p)^*$.

Proof. We first observe that, in view of the bound (B.1), the above map is bounded,

$$\|u\|_{(V_C^p)^*} = \sup\{B(v, u) : v \in V_C^p, \|v\|_{V^p} = 1\} = \|u\|_{U^q}$$

The last equality is a consequence of Theorem B.11. It remains to show that this map is onto.

Start with $L \in (V_C^p)^*$. By the theorem of Hahn Banach is has a same norm extension to the left continuous functions in V_{lc}^p which we denote again by L. By the embedding $U_{lc}^p \subset V_{rc}^p$, the map

$$v \to L(v) \in (U_{lc}^p)^*$$

has a unique representative $u \in V^q$. In particular this means that for any left continuous step function v we have

$$L(v) = B(v, u)$$

Here we have yet no continuity condition on u. However, if we replace it by its right continuous version u_{rc}, then we have for $v \in \mathfrak{S}_{lc}$ by an abuse of notation

$$B(v, u_{rc}) = \lim_{h>0, h\to 0} B(v(. - h), u).$$

and hence

$$\sup\{B(v, u_{rc}) : v \in \mathfrak{S}_{lc}, \|v\|_{V^p} \le 1\} \le \sup\{B(v, u) : v \in \mathfrak{S}_{lc}, \|v\|_{V^p} \le 1\}.$$

Now we are in a position to apply Theorem B.4 to conclude that $u_{rc} \in U^q$, and

$$\|u_{rc}\|_{U^q} \le \|L\|$$

To conclude the proof, we note that for $v \in V_C^p$ we have

$$B(v, u_{rc}) = B(v, u) = Lv.$$

Thus $u_{rc} \in U^q$ is a representation of L via the bilinear form B, and

$$\|L\| \le \|u_{rc}\|_{U^q} \le \|L\|$$

therefore equality must hold. The proof is complete. □

Lemma B.13. *C_0^∞ is dense in V_C^p and weak* dense in U^p and V^p.*

Proof. The first statement is simple. For the second it suffices to approximate atoms with compact support. Then any standard regularization gives the statement. □

Now we turn our attention to bilinear estimates. We begin with the algebra type properties:

Lemma B.14. *U^p and V^p are algebras, and the following estimates hold:*

$$\|vw\|_{V^p} \le \|v\|_{sup}\|w\|_{V^p} + \|v\|_{V^p}\|w\|_{sup} \tag{B.8}$$

$$\|uw\|_{U^p} \le 2\|u\|_{U^p}\|w\|_{U^p} \tag{B.9}$$

Proof. The first part is obvious. For the second part we have to consider a product of two atoms. □

We are now in a position to define a Stieltjes type integral.

Definition B.15. *Let $v \in V^p$ and $u \in U^q$ with $\frac{1}{p} + \frac{1}{q} = 1$. We define*

$$U = \int_a^t v(s)du(s)$$

by

$$B(w, U) = B(wv, u) \qquad \text{for all } w \in V_C^p \tag{B.10}$$

and

$$V = -\int_t^b u(s)dv(s)$$

by

$$B(V, w) = B(v, wu) \qquad \text{for all } w \in U^q$$

The right hand side defines a continuous linear map

$$V_C^p \ni w \to B(wv, u)$$

resp

$$U^q \ni w \to B(v, wu)$$

which by duality has a unique representative. Moreover by Theorem B.11 (B.10) holds for all $w \in V^p$. In particular

$$\int_a^t v_1(s)d(\int_a^s v_2(\sigma)du(\sigma)) = \int_a^t v_1(s)v_2(s)du(s).$$

To see this, define $U = \int_a^t v_2(s)du(s)$. Then almost by definition

$$B(v_1 v_2 w, u) = B(v_1 w, U)$$

Similarly, if $u_1, u_2 \in U^q$ and $V = -\int_t^b u_2(s)dv(s)$ then

$$B(v, wu_1u_2) = B(V, wu_1)$$

and hence

$$-\int_t^b u_1 u_2 dv = -\int_t^b u_1 \int_s^b u_2 dv.$$

For PDE applications we are also interested in the distributional interpretation of U^p and V^p functions. This is obvious for U^p, as any U^p function is bounded and uniquely determined by its values a.e.. However, in the case of V^p we have allowed single point jumps, which are invisible when testing against C_0^∞ test functions. Thus for this purpose it is more natural to work with the smaller space V_{lc}^p, whose elements are uniquely identified with distributions.

Next we consider the space of distributional derivatives of such functions:

Definition B.16. *We define*

$$DU^p = \{u';\ u \in U^p\}, \qquad DV^p = \{v';\ v \in V_{lc}^p\}$$

with the induced norm.

There is no difference in the definition of DV^p if we replace V_{lc}^p by V_{rc}^p and indeed the second one is more natural for solving Cauchy problems.

Based on our duality results above, we have the following dual characterization of these spaces:

Proposition B.17. *The spaces DU^p and DV^p can be characterized as the spaces of distributions for which the following norms are finite:*

$$\|f\|_{DU^p} = \sup\left\{\int f\phi dt : \|\phi\|_{V^q} \le 1, \phi \in C_0^\infty\right\} \tag{B.11}$$

$$\|f\|_{DV^p} = \sup\left\{\int f\phi dt : \|\phi\|_{U^q} \le 1, \phi \in C_0^\infty\right\} \tag{B.12}$$

for all distributions f.

Proof. At the heart of the proof is a very simple observation, namely that for $\phi \in C_0^\infty$ we have

$$B(v,\phi) = \int v\phi' dt, v \in V^p$$

respectively

$$B(\phi, u) = -\int u\phi' dt$$

This is easily verified directly for step functions, then by density for U^p functions, and by $V^p \subset U^q$ embeddings for V^p functions.

Because of this and the bound (B.1), it is clear that for $f \in DU^p$ we have

$$\left|\int f\phi dt\right| \le \|f\|_{DU^p}\|\phi\|_{V^p}$$

Conversely, let f be a distribution for which the norm on the right in (B.12) is finite. Then by the Hahn-Banach theorem (or by density) f extends to a bounded linear functional on V_C^p with the same norm. Hence by Theorem B.12 there exists $u \in U^q$, with the same norm, so that

$$B(u,\phi) = \int f\phi dt \qquad \text{for all } \phi \in C_0^\infty$$

Thus

$$-\int u\phi' dt = \int f\phi dt$$

i.e. $f = u'$ in the sense of distributions.

The argument is similar for DV^p, using instead the duality in Proposition B.2. □

Next we consider products of U^p and V^p functions with DU^p, and DV^p functions. Definition B.15 can be understood as a version of such product. By an abuse of notation we write the bilinear estimates as

$$\|vu'\|_{DU^p} \le c\|v\|_{V^q}\|u'\|_{DU^p} \tag{B.13}$$

and

$$\|uv'\|_{DV^p} \le c\|u\|_{U^q}\|v'\|_{DV^p}. \tag{B.14}$$

This is an abuse of notation since the product depends on whether the derivative is in DU^p or in DV^p.

There is an interpolation inequality.

Theorem B.18. *Let $1 < p < q < \infty$. There exists $C > 0$ so that for $v \in V^p_{rc}$ and $M > 1$ there exist $u \in U^p$ and $w \in U^q$ so that $v = u + w$ and*

$$\frac{1}{M}\|u\|_{U^p} + e^M\|w\|_{U^q} \le c\|v\|_{V^p}.$$

In particular $V^p_{rc} \subset U^q$.

Proof. We may assume that $\|v\|_{V^p} \le 1$. Let

$$\omega_p(f,t) = \sup_{t_1<t_2<t_n\le t} \sum |v(t_{j+1}) - v(t_j)|.$$

Then ω is monotonically nondecreasing and $\omega(t) \to 0$ as t tends to the left limit of the interval. We define

$$t_{jk} = \inf\{t : \omega(t) > j2^{-k}$$

and

$$u_k(t) = \begin{cases} 0 & \text{if } t_{k,2j} \le t < t_{k,2j+1} \\ u(t_{k,2j+1}) - u(t_{k,2j}) & \text{if } t_{k,2j+1} \le t < t_{k,2j+2} \end{cases}$$

Then

$$\|u_k\|_{U^r} \le 2^{k(\frac{1}{r}-\frac{1}{p})}$$

and

$$\|\sum_{k=1}^N u_k\|_{U^p} \le N$$

and

$$\|\sum_{k=N+1}^\infty u_k\|_{U^p} \le \frac{2^{(N+1)(\frac{1}{q}-\frac{1}{p})}}{1-2^{(\frac{1}{q}-\frac{1}{p})}}.$$

We choose

$$N = \frac{M}{\ln 2(\frac{1}{q}-\frac{1}{p})}$$

and C large. ☐

Corollary B.19. *Suppose that $p < q$. Then $V^p_{rc} \le U^q$,*

Proof. We apply the Theorem B.18 with $M = 1$. ☐

We want to relate the spaces to standard Besov spaces. This relies on the following lemma.

Lemma B.20. *Let $h > 0$. Then*

$$\|v(.+h) - v\|_{L^p(\mathbb{R})} \le ch^{\frac{1}{p}} \|v\|_{V^p}.$$

Proof. By scaling we may assume that $h = 1$. For $t \in [j, j+1)$

$$|v(t+1) - v(t)| \le \sup_{j\le t_1,t_2\le j+2} v(t_2) - v(t_1).$$

In particular there exists sequences $(t_{j,1})$ and $(t_{j,2})$ in $[j, j+2]$ so that

$$\sup_{j\le t_1,t_2\le j+2} (v(t_2) - v(t_1)) \le (1+\varepsilon)(v(t_{j,2}) - v(t_{j,1}))$$

Taking subsequences one see that

$$\left(\sum_j |v(t_{j,2}) - v(t_{j,1})|^p \right)^{\frac{1}{p}} \le 4^{\frac{1}{p}} \|v\|_{V^p}.$$

We integrate over the unit sized interval and sum with respect to j to obtain the estimate. ☐

Corollary B.21. *The following embedding estimates hold.*

(B.15)
$$\|v\|_{B^{\frac{1}{p}}_{p,\infty}} \le c\|v\|_{V^p}$$

(B.16)
$$\|u\|_{U^p} \le c\|u\|_{B^{\frac{1}{p}}_{p,1}}$$

Proof. The first inequality is an immediate consequence of (B.20). The second one follows by duality. □

We conclude with a density statement.

Theorem B.22. *Step functions are dense in U^p but not in V^p.*

Proof. The first statement is an immediate consequence of the atomic definition. Let $\eta \in C_0^\infty(a,b)$. We define

$$v = \eta \sum_{j=0}^{\infty} 2^{-\frac{j}{p}} \sin(2^j x).$$

It is not hard to see that $v \in C^{\frac{1}{p}}$ and hence $v \in V^p$, but it is not close to any step function. □

We conclude the section with the proof of Lemma 5.4.

Proof of Lemma 5.4 . First

$$\|u\|_{l^2_1 DU^2} \leq \|u\|_{l^2 L^2} = \|u\|_{L^2}$$

hence $L^2 \subset l^2_1 DU^2$. Now let $u \in DU^2$ and define

$$a = \sum_{j \in \mathbb{Z}} a_j \delta_j(t), \qquad a_j = \int_j^{j+1} u dt$$

where δ_j is the Dirac measure at the point j. This is well defined since $\chi_{[j,j+1)} \in V^2$. Moreover, since $U^2 \subset V^2$, we have the bound

$$\sum_j a_j^2 \leq \|u\|^2_{DU^2}.$$

In particular $a \in l^2_1 DU^2$. On the other hand, we can write

$$u - a = \sum_j \chi_{[j,j+1)} - a_j \delta_j$$

where each term is supported in $[j, j+1]$ and has integral zero. Thus by Lemma B.5 we can bound

$$\|u - a\|_{l^2_1 DU^2} \leq \|u\|_{DU^2}.$$

Hence $DU^2 \subset l^2_1 DU^2$.

Now suppose that $u \in l^2 DU^2$ and define

$$a(t) = \int_j^{j+1} u ds$$

if $t \in [j + j + 1)$. Then $a \in L^2$ and it remains to verify that

$$\|u - a\|_{DU^2} \leq c\|u\|_{l^2(DU^2)}$$

But this is easily verified for atoms. □

Acknowledgements: This research was carried out in part while the authors were visiting the Hausdorff Research Institute for Mathematics in Bonn in summer 2014, and in part while the authors were visiting MSRI in Berkeley in Fall 2015. The first author was supported by the DFG through SFB 1060. The second author was supported by the NSF grant DMS-1266182 and by a Simons Investigator grant from the Simons Foundation.

References

[1] Mark J. Ablowitz et al. "The inverse scattering transform-Fourier analysis for nonlinear problems". *Studies in Appl. Math.* 53.4 (1974), pp. 249–315.

[2] Sheldon Axler, Paul Bourdon, and Wade Ramey. *Harmonic function theory.* Second. Vol. 137. Graduate Texts in Mathematics. Springer-Verlag, New York, 2001, pp. xii+259.

[3] Tristan Buckmaster and Herbert Koch. "The Korteweg–de Vries equation at H^{-1} regularity". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 32.5 (2015), pp. 1071–1098.

[4] 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.

[5] Michael Christ, Justin Holmer, and Daniel Tataru. "Low regularity a priori bounds for the modified Korteweg-de Vries equation". *Lib. Math. (N.S.)* 32.1 (2012), pp. 51–75.

[6] P. Deift, S. Venakides, and X. Zhou. "The collisionless shock region for the long-time behavior of solutions of the KdV equation". *Comm. Pure Appl. Math.* 47.2 (1994), pp. 199–206.

[7] 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.

[8] Benoît 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.

[9] Martin Hadac, Sebastian Herr, and Herbert Koch. "Well-posedness and scattering for the KP-II equation in a critical space". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 26.3 (2009), pp. 917–941.

[10] T. Kappeler, B. Schaad, and P. Topalov. "mKdV and its Birkhoff coordinates". *Phys. D* 237.10-12 (2008), pp. 1655–1662.

[11] T. Kappeler and P. Topalov. "Global wellposedness of KdV in $H^{-1}(\mathbb{T},\mathbb{R})$". *Duke Math. J.* 135.2 (2006), pp. 327–360.

[12] Rowan Killip, Visan Monica, and Zhang. Talk Bonn March 2016 and private communication. 2016.

[13] Herbert Koch and Daniel Tataru. "A priori bounds for the 1D cubic NLS in negative Sobolev spaces". *Int. Math. Res. Not. IMRN* 16 (2007), Art. ID rnm053, 36.

[14] Herbert Koch and Daniel Tataru. "Energy and local energy bounds for the 1-d cubic NLS equation in $H^{-\frac{1}{4}}$". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 29.6 (2012), pp. 955–988.

[15] N. V. Krylov. "Boundedly inhomogeneous elliptic and parabolic equations in a domain". *Izv. Akad. Nauk SSSR Ser. Mat.* 47.1 (1983), pp. 75–108.

[16] Baoping Liu. "A priori bounds for KdV equation below $H^{-\frac{3}{4}}$". *J. Funct. Anal.* 268.3 (2015), pp. 501–554.

[17] M. Lothaire. *Combinatorics on words.* Vol. 17. Encyclopedia of Mathematics and its Applications. A collective work by Dominique Perrin, Jean Berstel, Christian Choffrut, Robert Cori, Dominique Foata, Jean Eric Pin, Guiseppe Pirillo, Christophe Reutenauer, Marcel-P. Schützenberger, Jacques Sakarovitch and Imre Simon, With a foreword by Roger Lyndon, Edited and with a preface by Perrin. Addison-Wesley Publishing Co., Reading, Mass., 1983, pp. xix+238.

[18] Luc Molinet. "A note on ill posedness for the KdV equation". *Differential Integral Equations* 24.7-8 (2011), pp. 759–765.

[19] Richard S. Palais. "The symmetries of solitons". *Bull. Amer. Math. Soc. (N.S.)* 34.4 (1997), pp. 339–403.

[20] Christophe Reutenauer. *Free Lie algebras.* Vol. 7. London Mathematical Society Monographs. New Series. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1993, pp. xviii+269.
[21] Robert G. Underwood. *An introduction to Hopf algebras.* Springer, New York, 2011, pp. xiv+273.
[22] Xin Zhou. "Direct and inverse scattering transforms with arbitrary spectral singularities". *Comm. Pure Appl. Math.* 42.7 (1989), pp. 895–938.

Mathematisches Institut, Universität Bonn

Department of Mathematics, University of California, Berkeley