



Orbital Stability of KdV Multisolitons in H^{-1}

Rowan Killip , Monica Visan

Department of Mathematics, University of California, Los Angeles, CA 90095, USA.
E-mail: killip@math.ucla.edu; visan@math.ucla.edu

Received: 25 September 2020 / Accepted: 15 November 2021

Published online: 17 January 2022 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract: We prove that multisoliton solutions of the Korteweg–de Vries equation are orbitally stable in $H^{-1}(\mathbb{R})$. We introduce a variational characterization of multisolitons that remains meaningful at such low regularity and show that all optimizing sequences converge to the manifold of multisolitons. The proximity required at the initial time is uniform across the entire manifold of multisolitons; this had not been demonstrated previously, even in H^1 .

1. Introduction

The history of the Korteweg–de Vries equation

$$\frac{d}{dt}q = -q''' + 6qq' \quad (\text{KdV})$$

is profoundly intertwined with the notion of solitary waves. Indeed, the very goal of Korteweg and de Vries [24] was to explain the empirical observation of such waves.

The fact that (KdV) admits solutions of the form

$$q(t, x) = -2\beta^2 \operatorname{sech}^2(\beta[x - 4\beta^2 t - x_0]) \quad (1.1)$$

(for any $\beta > 0$ and $x_0 \in \mathbb{R}$) explains many aspects of solitary water waves, such as the relation between their height and speed. However, the very possibility of Scott Russell’s famous chance encounter with such a wave tells us something more: It must be stable!

The question of stability was considered already by Boussinesq in [6]. In addition to observing the conservation of both

$$P(q) := \int \frac{1}{2}q(x)^2 dx \quad \text{and} \quad H(q) := \int \frac{1}{2}q'(x)^2 + q(x)^3 dx, \quad (1.2)$$

he also notes that the solitary wave profile solves the Euler–Lagrange equation associated to the problem of optimizing H subject to constrained P .

Now if the solitary wave were a non-degenerate minimum of H at constrained P , then stability would follow immediately, following the Lyapunov model. However, it is not! The simple act of translation shows that it is at best a degenerate minimum.

In the pioneering paper [4], Benjamin proved the H^1 -orbital stability of such solitary waves: Solutions close to a soliton profile at time zero remain close to a soliton profile at all times. This variational approach is extremely robust and has seen countless applications since. However, it does not directly give any information about the physical location of the soliton profile, nor how this evolves with time; this is the significance of the adjective ‘orbital’.

In numerical simulations of a discrete form of (KdV), Kruskal and Zabusky [40] observed that solitary waves exhibit an even stronger form of stability: Pairs of solitary waves emerged from collisions with the same profile and speed with which they had entered. Nevertheless, the two waves did interact; each was spatially shifted from its original trajectory. This particle-like behavior led Kruskal and Zabusky to coin the name *soliton*; they presciently appreciated that this was an exotic phenomenon.

We now understand that while the orbital stability of single solitary waves is rather common (and can often be proved variationally), stability under collisions is extremely peculiar. The ultimate explanation for this behavior was the discovery that (KdV) is a completely integrable Hamiltonian system; see [11, 12, 32, 42].

Just as our notion of a solitary wave crystalizes around the concrete particular solutions (1.1) to (KdV), so there is a family of special solutions to (KdV) that embody the behavior of collections of solitons:

Definition 1.1 (Multisoliton solutions) Fix $N \geq 1$. Given N distinct positive parameters $\beta_1, \dots, \beta_N$ and N real parameters $c_1, \dots, c_N$, let

$$Q_{\vec{\beta}, \vec{c}}(x) = -2 \frac{d^2}{dx^2} \ln \det[A(x)] \quad (1.3)$$

where $A(x)$ is the $N \times N$ matrix with entries

$$A_{\mu\nu}(x) = \delta_{\mu\nu} + \frac{1}{\beta_\mu + \beta_\nu} e^{-\beta_\mu(x-c_\mu) - \beta_\nu(x-c_\nu)}. \quad (1.4)$$

The unique solution to (KdV) with initial data $q(0, x) = Q_{\vec{\beta}, \vec{c}}(x)$ is

$$q(t, x) = Q_{\vec{\beta}, \vec{c}(t)}(x) \quad \text{where} \quad c_n(t) = c_n + 4\beta_n^2 t. \quad (1.5)$$

Although the definition of $A(x)$ depends on the ordering of the parameters, $\det[A(x)]$ and so also $Q_{\vec{\beta}, \vec{c}}$ do not.

The beautiful formula (1.3) was originally derived in [17] as a description of reflectionless potentials appearing in the one-dimensional Schrödinger equation. With the discovery of the inverse-scattering approach, the significance of this result for (KdV) was noted by several authors; see [11, 12, 15, 37, 38, 41]. By analyzing these exact solutions, the authors confirmed the particle-like interactions, described the long-time asymptotics, and determined the (universal) spatial shifts.

The idea that these explicit solutions provide a justification for empirical observations is necessarily predicated (at the very least) on their stability. Indeed, this question has attracted considerable attention over the years, as we shall discuss shortly. Let us begin, however, with our own contribution to this question:

Theorem 1.2. Fix $N \geq 1$ and distinct positive parameters $\beta_1, \dots, \beta_N$. For every $\varepsilon > 0$ there exists $\delta > 0$ so that for every initial data $q(0) \in H^{-1}(\mathbb{R})$ satisfying

$$\inf_{\vec{c} \in \mathbb{R}^N} \|q(0) - Q_{\vec{\beta}, \vec{c}}\|_{H^{-1}} < \delta,$$

the corresponding solution $q(t)$ to (KdV) satisfies

$$\sup_{t \in \mathbb{R}} \inf_{\vec{c} \in \mathbb{R}^N} \|q(t) - Q_{\vec{\beta}, \vec{c}}\|_{H^{-1}} < \varepsilon.$$

One virtue of this result is that it achieves the lowest regularity (in the H^s scale) for which well-posedness is known [20] or possible [34]. In Sect. 6, we shall also see that it is not difficult to recover higher-regularity results post factum:

Corollary 1.3. Fix $s \geq -1$, $N \geq 1$, and distinct positive parameters $\beta_1, \dots, \beta_N$. For every $\varepsilon > 0$ there exists $\delta > 0$ so that

$$\inf_{\vec{c} \in \mathbb{R}^N} \|q(0) - Q_{\vec{\beta}, \vec{c}}\|_{H^s} < \delta \implies \sup_{t \in \mathbb{R}} \inf_{\vec{c} \in \mathbb{R}^N} \|q(t) - Q_{\vec{\beta}, \vec{c}}\|_{H^s} < \varepsilon. \quad (1.6)$$

Let us now turn toward a discussion of prior work, after which we will discuss how the proof of Theorem 1.2 will proceed. We do not intend to dwell on the question of well-posedness, since this is rather decoupled from the question of stability: Proving an assertion like (1.6) only for Schwartz solutions still cuts to the heart of the matter; the Schwartz restriction can then be trivially removed once well-posedness in H^s is known. Indeed, Benjamin's work on H^1 -stability should only grow in our estimation when we consider that well-posedness in H^1 was not achieved until many years later, [18]. Conversely, having obtained well-posedness in H^{-1} in [20], it is timely to address the orbital stability in this space.

It is also true that well-posedness alone provides little assistance in proving (1.6). Nevertheless, it has proved useful in the consideration of slightly weaker assertions, where δ is permitted to depend on the parameters $\vec{c}$ of the multisoliton nearest the initial data. The manner in which it helps is this: Exact multisolitons resolve (as $t \rightarrow \pm\infty$) into essentially a linear combination of well-separated (and increasingly separated) simple solitary waves of the form (1.1). Thus, researchers may confine their analyses to treating initial data in this more favorable configuration, and exploiting well-posedness to cover the remaining compact time interval.

While Benjamin's argument [4] was both extremely novel and compelling, it did contain some mathematical lacunae, particularly with regard to the treatment of the modulation parameters. These issues were thoroughly addressed by Bona [5]. This approach was further developed to treat NLS and gKdV by Weinstein [39].

Orbital stability of the single soliton (1.1) in L^2 was only shown much more recently, by Merle and Vega [31]. These authors also show a form of asymptotic stability: one has L^2 -convergence to a soliton profile in any bounded window traveling with the soliton. Stronger forms of asymptotic stability such as global L^2 convergence are clearly forbidden by the conservative nature of the equation. We should also note that it is not claimed that the solution is converging to a single solitary wave with fixed translation parameter x_0 ; indeed, subsequent analysis by Martel and Merle [29] shows that this cannot be guaranteed: successive interactions with a large number of wide (and so L^2 -small) solitary waves can lead to logarithmic divergence of the soliton trajectory from a straight line.

The Merle–Vega proof of L^2 -orbital stability of single solitons combines the Miura map with orbital stability of the kink solutions proven in [43]. (While Zhidkov focusses

on the NLS equation, his variational analysis employs only conservation laws common to mKdV.)

In the same paper [26] that introduced the Lax pair, Lax also discusses two-soliton solutions with a view to explaining the properties of such waves observed in [40]. His construction of such solutions is based on a differential equation derived from the polynomial conservation laws discovered earlier in [32]. While Lax does not explicitly express it thus in this paper (see [27], however), his equation arises as the Euler–Lagrange equation for optimizing the third conserved quantity with the first two constrained. In general, N -solitons are critical points of the variational problem of optimizing the $(N + 1)$ polynomial conserved quantity constrained by its N predecessors (we exclude the Casimir $\int q \, dx$ from our enumeration).

This constrained variational problem was analyzed by Maddocks and Sachs in [28]. They showed that multisolitons are in fact local minimizers. The essential (and subtle) point addressed by these authors is to understand the Hessian of the highest-order conservation law on the manifold of multisolitons, both directly and restricted to directions parallel to the constraint manifold. By doing so, they prove that the manifold of N -solitons is orbitally stable in $H^N(\mathbb{R})$. Notice that the regularity varies depending on the complexity of the multisoliton. This appears naturally in their variational problem because one needs to make sense of increasingly many polynomial conservation laws.

As the analysis in [28] is localized in small neighbourhoods of the soliton profiles, it does not address either of the following questions: Are N -solitons global minimizers of this variational problem? Are they the *only* minimizers? To the best of our knowledge, both questions remain open. Theorem 1.4 below gives an affirmative answer to both questions for the variational description we employ.

Orbital stability of multisolitons in H^1 was shown by Martel, Merle, and Tsai in [30]. The principal part of the argument is showing that a system of well-separated solitons (ordered by speed) is future-stable. Subsequently in [3], Alejo, Muñoz, and Vega proved orbital stability of multisolitons in L^2 by using Gardner’s generalization of the Miura map and applying the ideas of [30] to the resulting Gardner equation. These works do not yield orbital stability in the strong form of (1.6); they rely on local well-posedness in the manner discussed earlier. Additional information on the modulation parameters over this initial interval is obtained (a posteriori) in [2].

A different approach to orbital stability of solitons based on *auto*Bäcklund transformations (which add or remove solitons) was demonstrated recently in [33]. This work proves a strong form of L^2 -orbital stability of one-solitons for the focusing cubic NLS on the line by combining these transformations with stability of the zero solution. This approach was substantially advanced in [23], where low-regularity orbital stability of NLS multisolitons (including the delicate case of multiple eigenvalues) was proved. To the best of our knowledge, these ideas have not yet been applied to (KdV).

Let us now turn to the topic of the methods to be employed in this paper. Our discussion will be somewhat discursive since we shall take the time to introduce the central object of our methodology, the (doubly) renormalized perturbation determinant, as well as historical and contextual matters that we find instructive.

As we have discussed, the stability of multisolitons is historically (and physically) inseparable from the complete integrability of (KdV). The key question is how this complete integrability is to be exploited.

The long-standing approach, introduced already in [11], is to employ the scattering theory of one-dimensional Schrödinger operators with the potential given by the (KdV) wave form at a fixed time. Despite receiving a great deal of attention over the years (with

much impetus taken from the study of KdV), there is currently no satisfactory theory of forward or inverse scattering in any H^s space. While non-trivial problems do attend low regularity, it is the slow decay associated with such spaces that is most devastating. We are truly at a loss as to how to define the reflection coefficient or how to handle embedded eigenvalues and singular continuous spectrum.

The inverse-scattering technique is capable of providing extremely detailed long-time asymptotics for the class of solutions to which it is applicable; see [13], for example. However, due to the difficulties outlined above, it has not yet yielded stability of even single-solitons in any H^s space.

While the reflection coefficient is fragile, it has long been appreciated that the transmission coefficient is much more robust. One intuitive explanation for this is that the transmission coefficient actually represents the boundary values of a function meromorphic in the upper half-plane. Analytically, it is preferable to consider the reciprocal $a(k; q)$ of the transmission coefficient. This is holomorphic in the upper half-plane and its zeros precisely encode the discrete spectrum of the attendant Schrödinger operator. The simplest description is as the Wronskian (divided by $2ik$) of the two Jost solutions.

An alternate perspective on this function $a(k; q)$ was introduced by Jost and Pais [16]. They observed that it could be expressed as a Fredholm determinant. In [36, Chapter 5], Simon proves that

$$a(k; q) = \det(1 + |q|^{\frac{1}{2}} R_0(k) |q|^{-\frac{1}{2}} q), \quad \text{where } R_0(k) = (-\partial_x^2 - k^2)^{-1}, \quad (1.7)$$

coincides with the Wronskian definition provided $\langle x \rangle^{1+\delta} q \in L^1$ with $\delta > 0$. Here the operator is regarded as acting on $L^2(\mathbb{R})$ and the determinant is taken over this Hilbert space. These are our standing conventions (exceptions will be noted).

Splitting q across the two sides of R_0 is necessary in (1.7) if one wishes to treat q with L^1 -type singularities: neither $R_0 q$ nor $q R_0$ could be guaranteed to be bounded under Simon's hypothesis. However, it turns out to be wiser to factor the free resolvent R_0 , placing a square-root of this operator on either side of q ; as we shall see, this will permit potentials with much more severe singularities. On the other hand, one still needs strong decay hypotheses on q ; for otherwise, the determinant would not be defined.

The second layer of renormalization needed to treat $q \in H^{-1}$ employs the renormalized determinant introduced by Hilbert [14]; see [36, Chapter 9]. Concretely,

$$\det_2(1 + A) := \det(1 + A) e^{-\text{tr } A}$$

extends continuously from trace class $\mathfrak{I}_1$ to Hilbert–Schmidt class $\mathfrak{I}_2$. Recall that $\mathfrak{I}_2$ is comprised of those operators whose singular values are square summable and

$$\|A\|_{\mathfrak{I}_2}^2 = \text{tr}(A^* A).$$

Combining these two renormalizations, we are led to consider the following: For $k \in \mathbb{C}^+ = \{z \in \mathbb{C} : \text{Im } z > 0\}$ and Schwartz-class q ,

$$a_{\text{ren}}(k; q) := \det_2(1 + \sqrt{R_0(k)} q \sqrt{R_0(k)}) = a(k; q) \exp\left\{-\frac{i}{2k} \int q(x) dx\right\}. \quad (1.8)$$

The square-root of the resolvent is defined via analytic continuation from the case $k = i\kappa$ with $\kappa > 0$, in which case R_0 is positive definite (and we take the positive definite square-root).

To the best of our knowledge, this quantity was first considered by Rybkin. In [35], he used it to give the first proof of a priori H^{-1} bounds for solutions to (KdV). This

approach was developed independently in [21]; alternate approaches to such a priori bounds can be found in [7, 22].

The fact that a_{ren} extends continuously (indeed real-analytically) from $q \in \mathcal{S}$ to merely $q \in H^{-1}$ rests on the basic theory of such regularized determinants and the Hilbert–Schmidt estimate

$$\left\| \sqrt{R_0(k)} q \sqrt{R_0(k)} \right\|_{\mathfrak{J}_2}^2 \leq \frac{|k|}{[\text{Im } k]^2} \int \frac{|\hat{q}(\xi)|^2}{\xi^2 + 4|k|^2} d\xi. \quad (1.9)$$

Indeed, the mapping $A \mapsto \det_2(1 + A)$ is a complex-analytic function on $\mathfrak{J}_2$ and

$$|1 - \det_2(1 + A)| \lesssim \|A\|_{\mathfrak{J}_2} \exp\{\|A\|_{\mathfrak{J}_2}^2\}; \quad (1.10)$$

see [36] for details. Our justification for the bound (1.9) is quite simple. We use the ideal property and the elementary bound

$$|\xi^2 - k^2|^{-1} \leq \frac{|k|}{\text{Im } k} (\xi^2 + |k|^2)^{-1} \quad \text{for all } \xi \in \mathbb{R}$$

to reduce matters to the $\kappa = |k|$ case of

$$\left\| \sqrt{R_0(i\kappa)} q \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_2}^2 = \frac{1}{\kappa} \int \frac{|\hat{q}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \quad \text{for all } \kappa > 0. \quad (1.11)$$

In view of the importance of (1.11) for what follows, it will also be convenient to employ the notation

$$\|f\|_{H_{\kappa}^{-1}}^2 := \int \frac{|\hat{f}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi.$$

With these preliminaries set, we may now give our variational characterization of multisolitons:

Theorem 1.4 (Variational characterization of multisolitons) *Fix $N \geq 1$ and distinct positive parameters $\beta_1, \dots, \beta_N$. If $q \in H^{-1}$ satisfies*

$$a_{\text{ren}}(k; q) = 0 \quad \text{for all } k \in \{i\beta_m : 1 \leq m \leq N\}, \quad (1.12)$$

then

$$a_{\text{ren}}(i\kappa, q) \leq \exp \left\{ \sum_{m=1}^N \ln \left(\frac{\kappa - \beta_m}{\kappa + \beta_m} \right) + \frac{2\beta_m}{\kappa} \right\} \quad \text{for all } \kappa \geq 1 + \|q\|_{H^{-1}}^2. \quad (1.13)$$

Equality holds in (1.13) for any one such κ if and only if $q = Q_{\vec{\beta}, \vec{c}}$ for some $\vec{c} \in \mathbb{R}^N$.

The fact that equality holds in (1.13) for $Q_{\vec{\beta}, \vec{c}}$ is easily deduced from well-known properties of multisolitons; indeed,

$$a(k; Q_{\vec{\beta}, \vec{c}}) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m} \quad \text{and} \quad a_{\text{ren}}(k; Q_{\vec{\beta}, \vec{c}}) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m} e^{2i\beta_m/k},$$

as we will discuss in Sect. 2. As each β_m is a zero of the determinant $a(k; Q_{\vec{\beta}, \vec{c}})$, we will often refer to them as eigenvalue parameters (see also (2.25)) to better distinguish them from the parameters $\vec{c}$.

The converse claim, identifying multisolitons as the only optimizers, is more challenging. Nevertheless, it does not directly yield stability: one would also need to know that profiles that almost optimize (1.13) are close to actual optimizers (i.e., to multisolitons). This leaves us with a very clear ambition of a purely variational character: prove that optimizing sequences converge to the manifold of multisolitons.

We cannot expect optimizing sequences to have convergent subsequences — the manifold of optimizers is not compact! This problem arises already in the case of single solitons, due to the translation symmetry. In the one-soliton case, compactness can be restored by incorporating translations. This approach was convincingly demonstrated by Cazenave and Lions [8], who proved orbital stability of ground-state solitary waves for a variety of NLS-like equations. Their paper is a major inspiration for what follows.

In the multisoliton case, compactness cannot be restored by translation alone. Indeed the long-time dynamics of the multisolitons themselves is to break into asymptotically well-separated one-solitons. We need a profile decomposition! However, unlike most applications of this concentration-compactness technique, there is no sub-additivity in our problem: dichotomy must be embraced, not refuted. As we will discuss, this is just one of several subtle aspects to our implementation of this classic concentration-compactness device.

We should note that the scenario of asymptotically well-separated one-solitons is not the only manner in which dichotomy can arise for optimizing sequences (or indeed sequences of optimizers). One may have asymptotically well-separated *multisolitons*. This ‘gas of molecules’ scenario will be analyzed in Sect. 3, where we show that a linear combination of well-separated multisolitons can be well-approximated by a single exact multisoliton.

Further ways in which our concentration-compactness analysis diverges from the other examples we know are (i) we are working in trace ideals, not Lebesgue spaces; (ii) while we do have local compactness, this is non-quantitative arising from mere equicontinuity; and (iii) the constraints are apportioned across the profiles in an exotic manner. We will discuss each of these in succession.

Trace ideals (which are also known as non-commutative ℓ^p spaces) have an additional defect of compactness beyond those of sequence spaces, namely, unitary conjugation. This in an infinite-dimensional group.

Our concentration-compactness principle contains only translation parameters, no scaling parameters (Lions dubbed this the locally compact case). In such cases, the profiles are usually constructed by employing the Rellich–Kondrashov Theorem. In our case, however, there is no such quantitative principle. We will be able to show that individual optimizing sequences are equicontinuous (cf. Definition 4.1), but nothing more.

In the standard analyses, a constraint, such as on the total L^2 norm, is apportioned across the profiles in an additive manner: the mass of the sequence is the sum of the masses of the profiles, plus that of the remainder. In our case, the constraints are vanishing of the perturbation determinant. In Sect. 5, we will see that the profiles attendant to optimizing sequences share the constraints in a different way: different profiles satisfy different *subsets* of the constraints.

The paper is organized as follows: In Sect. 2, we first develop the theory of the perturbation determinant a little further. We then use this to prove Theorem 1.4. Our approach is this: Building on the existing theory of Schwartz-class potentials, we show that the upper-bound (1.13) holds across all $q \in H^{-1}$. Having first proved linear independence of the gradients of the constraints, we may analyze the case of equality using the Euler–

Lagrange equation. Using this device, we show that optimizers are, in fact, Schwartz class. We may then appeal to classical inverse scattering to deduce that q is an exact multisoliton.

In Sect. 3, we show that well-separated linear combinations of multisolitons (which may arise as optimizing sequences) can be approximated by a single multisoliton. This is notationally very cumbersome; nevertheless, we hope that the virtues of deforming x into the complex plane and exploiting the determinantal relation (3.8) shine through.

In Sect. 4, we develop a profile decomposition attendant to the functional $q \mapsto \alpha(\kappa; q)$, defined in (2.1), applied to bounded and equicontinuous sequences in H^{-1} . Structurally speaking, our approach is the one we advanced in [19], namely, to first prove an inverse inequality and then employ this inductively to extract profiles.

In Sect. 5, we prove Theorem 1.2, arguing by contradiction. If the theorem were to fail, then there would exist a sequence of solutions q_n so that the initial data $q_n(0)$ converges to the manifold of solitons, and a sequence of times t_n so that $q_n(t_n)$ does not converge to the manifold of solitons. Using the fact that $\alpha(q; \kappa)$ is conserved under the flow, we show that q_n is an optimizing sequence for the variational problem described in Theorem 1.4. (Actually, this is not quite correct, the zeros may be slightly displaced.) We then employ the profile decomposition of Sect. 4 to show (after some work) that the optimizing sequence can be approximated by a linear combination of well-separated multisolitons. This suffices to reach a contradiction because of the analysis in Sect. 3.

We prove Corollary 1.3 in Sect. 6. In doing so, we illustrate two basic methods for raising the regularity. First we exploit the equicontinuity of orbits to prove the result for $s \in (-1, 1)$. The virtues of this argument are that it is simple and that it further illustrates the advantages of equicontinuity. Our second method employs conservation laws. When $s \geq 0$ is an integer, we need only the polynomial conservation laws discovered already in [32] and the argument is once again quite elementary. For general $s > -1$, we employ the conservation laws developed in [22].

2. Variational Characterization of Multisolitons

The ultimate goal of this section is to prove Theorem 1.4. This will proceed in several stages. First, we discuss the logarithm of a_{ren} . Then we show that (1.13) holds, first for Schwartz-class q and then for general $q \in H^{-1}$. The climax of the proof is showing that all H^{-1} optimizers are, in fact, Schwartz class and then using this information to show that they must be multisolitons.

Lemma 2.1. *For $q \in H^{-1}$ and $\kappa \geq 1 + \|q\|_{H^{-1}}^2$, the series*

$$\alpha(\kappa; q) := \sum_{\ell=2}^{\infty} \frac{1}{\ell} (-1)^{\ell} \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} q \sqrt{R_0(i\kappa)} \right)^{\ell} \right\} \quad (2.1)$$

converges and

$$a_{\text{ren}}(q; i\kappa) = \exp\{-\alpha(q; \kappa)\}. \quad (2.2)$$

Moreover,

$$\liminf_{\kappa \rightarrow \infty} 8\kappa^3 \alpha(\kappa; q) = \|q\|_{L^2}^2, \quad (2.3)$$

with the understanding that LHS (2.3) is infinite if $q \notin L^2$.

Proof. Convergence of the series (2.1) under this hypothesis on κ follows immediately from (1.11). That exponentiating this series yields the renormalized determinant is well known; indeed, this is little more than the Newton–Girard relation between elementary and power-sum symmetric functions.

Employing (1.11) in the series (2.1) shows

$$\left| 8\kappa^3 \alpha(q; \kappa) - \int \frac{4\kappa^2 |\hat{q}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \right| \leq \frac{\|q\|_{H^{-1}}}{\sqrt{\kappa} - \|q\|_{H^{-1}}} \int \frac{4\kappa^2 |\hat{q}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi, \quad (2.4)$$

from which (2.3) follows immediately. $\square$

Incidentally, we note the inequality (2.4) is actually the basis of the proof of a priori H^{-1} bounds. Indeed, combining this with a simple bootstrap argument shows that for $\kappa \geq 1 + 64\|q(0)\|_{H_\kappa^{-1}}^2$ and any $t \in \mathbb{R}$,

$$\frac{2}{3} \int \frac{|\hat{q}(t, \xi)|^2}{\xi^2 + 4\kappa^2} d\xi \leq 2\kappa \alpha(\kappa; q(t)) = 2\kappa \alpha(\kappa; q(0)) \leq \frac{8}{7} \int \frac{|\hat{q}(0, \xi)|^2}{\xi^2 + 4\kappa^2} d\xi. \quad (2.5)$$

Let us now recall some known facts about the reciprocal transmission coefficient $a(k; q)$ in the case q is of Schwartz class. The basic analytical facts listed below can be easily derived from the Wronskian definition of $a(k; q)$ and rigorous proofs can be found in many basic texts on scattering theory. The claim (2.9) is more serious. While many introductory texts on the theory of solitons give at least a formal derivation of (1.3) from the assumption that $a(k; q)$ takes the stated form, a rigorous treatment requires considerable care, especially on the question of uniqueness. We recommend the paper [10] of Deift and Trubowitz for a complete and self-contained presentation of the following (under rather weaker hypotheses):

Proposition 2.2. *Fix $q \in \mathcal{S}$. Then $a(k; q)$ extends continuously to the closed upper half-plane. It has finitely many zeroes in $\mathbb{C}^+$, all of which are simple and located on the imaginary axis. Moreover,*

$$|a(k; q)| \geq 1 \quad \text{for all } k \in \mathbb{R}, \quad (2.6)$$

$$|a(k; q) - 1| = O\left(\frac{1}{|k|}\right) \quad \text{as } |k| \rightarrow \infty \quad \text{uniformly for } \operatorname{Im} k \geq 0, \quad (2.7)$$

and we have the symmetry

$$\overline{a(k; q)} = a(-\bar{k}; q) \quad \text{for all } k \in \mathbb{C}^+. \quad (2.8)$$

Finally, given distinct $\beta_1, \dots, \beta_N \in (0, \infty)$ and $q \in \mathcal{S}$,

$$a(k; q) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m} \iff q \in \{Q_{\vec{\beta}, \vec{c}} : \vec{c} \in \mathbb{R}^N\}. \quad (2.9)$$

This does not address the value of the renormalized perturbation determinant for such multisolitons. The missing ingredient is the following:

$$\int Q_{\vec{\beta}, \vec{c}}(x) dx = - \sum_{m=1}^N 4\beta_m. \quad (2.10)$$

This is proved in both [12] and [42]. One simple approach that explains the additive structure of RHS (2.10) is this: As the $\int q$ is conserved by the flow, the value of LHS (2.10) can be determined from the value for N well-separated single solitons. Alternately, one may deduce this by comparing the large- k asymptotics of LHS (2.9) with those of $a(\kappa; q)$. From the same references or by the same method, one can also find

$$P(Q_{\vec{\beta}, \vec{c}}) = \frac{8}{3} \sum_m \beta_m^3 \quad \text{and} \quad H(Q_{\vec{\beta}, \vec{c}}) = -\frac{32}{5} \sum_m \beta_m^5. \quad (2.11)$$

Combining (1.8), (2.2), and (2.10) shows

$$a_{\text{ren}}(k; Q_{\vec{\beta}, \vec{c}}) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m} e^{2i\beta_m/k} \quad \text{for all } k \in \mathbb{C}^+, \quad (2.12)$$

$$\alpha(\kappa; Q_{\vec{\beta}, \vec{c}}) = - \sum_{m=1}^N \left[\frac{2\beta_m}{\kappa} + \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) \right] \quad \text{for all } \kappa \geq 1 + \|q\|_{H^{-1}}^2. \quad (2.13)$$

As Lemma 2.1 guarantees that a_{ren} is non-vanishing for $\kappa \geq 1 + \|q\|_{H^{-1}}^2$, the restriction on κ guarantees $\kappa > \sup_m \beta_m$ and consequently, that RHS (2.13) is positive:

$$G\left(\frac{\beta_m}{\kappa}\right) := - \left[\frac{2\beta_m}{\kappa} + \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) \right] = \sum_{\ell \geq 1} \frac{2}{2\ell+1} \left(\frac{\beta_m}{\kappa}\right)^{2\ell+1} > 0. \quad (2.14)$$

Recalling (2.2), we see that multisolitons achieve equality in (1.13). We next show that this is indeed a bound for all q . This will be done in two steps: first for $q \in \mathcal{S}$ and then for $q \in H^{-1}$:

Proposition 2.3. *For $q \in \mathcal{S}$ and $\kappa \geq 1 + \|q\|_{H^{-1}}^2$,*

$$\alpha(\kappa; q) \geq - \sum_{m=1}^N \left[\ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \frac{2\beta_m}{\kappa} \right], \quad (2.15)$$

where $\{\beta_m : 1 \leq m \leq N\}$ enumerates the $0 \leq N < \infty$ zeros of a_{ren} .

Proof. Proposition 2.2 shows that $a(k; q)$ has only finitely many zeros and all are simple. Using these, we build the Blaschke product

$$B(k) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m}.$$

In the case $a(k; q)$ has no zeros, $B(k) \equiv 1$.

Using Proposition 2.2 again, we see that $k \mapsto \ln \left| \frac{a(k; q)}{B(k)} \right|$ is harmonic on $\mathbb{C}^+$ and extends continuously to $\partial\mathbb{C}^+$. Moreover,

$$\ln \left| \frac{a(k; q)}{B(k)} \right| \geq 0 \quad \text{for all } k \in \mathbb{R} \quad \text{and} \quad \ln \left| \frac{a(k; q)}{B(k)} \right| = O\left(\frac{1}{|k|}\right) \quad \text{as } |k| \rightarrow \infty. \quad (2.16)$$

It follows from the maximum principle that this function is non-negative throughout $\mathbb{C}^+$.

The Herglotz Representation Theorem (cf. [1, Theorem 3, §59]) then guarantees

$$\ln \left[\frac{a(k; q)}{B(k)} \right] = -i \int_{\mathbb{R}} \frac{d\mu(t)}{t-k}, \quad (2.17)$$

for some finite positive measure $d\mu$ on $\mathbb{R}$. This measure is also even under $t \mapsto -t$; this is inherited from the symmetry (2.8) enjoyed by both $a(k; q)$ and $B(k)$.

In this way, we see that for $\kappa \geq 1 + \|q\|_{H^{-1}}^2$,

$$-\ln a(i\kappa; q) = -\sum_m \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + i \int \frac{t + i\kappa}{t^2 + \kappa^2} d\mu(t) = -\sum_m \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) - \kappa \int \frac{d\mu(t)}{t^2 + \kappa^2}.$$

On the other hand, (1.8), (2.2), and (2.3) show that as $\kappa \rightarrow \infty$,

$$\left| -\ln a(i\kappa; q) + \frac{1}{2\kappa} \int q(x) dx \right| = O(\kappa^{-3}).$$

Combining these two observations, we deduce that

$$\begin{aligned} \int q(x) dx &= \lim_{\kappa \rightarrow \infty} \left[\sum_m 2\kappa \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \int \frac{2\kappa^2}{t^2 + \kappa^2} d\mu(t) \right] \\ &= -4 \sum_m \beta_m + 2 \int d\mu(t) \end{aligned} \quad (2.18)$$

and thence that

$$\alpha(\kappa; q) = -\ln a_{\text{ren}}(i\kappa; q) = -\sum_{m=1}^N \left[\ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \frac{2\beta_m}{\kappa} \right] + \int \frac{t^2}{\kappa(t^2 + \kappa^2)} d\mu(t), \quad (2.19)$$

for all $\kappa \geq 1 + \|q\|_{H^{-1}}^2$. The claim (2.15) follows since $d\mu \geq 0$. $\square$

Corollary 2.4. Fix $N \geq 0$ and distinct positive parameters $\beta_1, \dots, \beta_N$. Assume that $q \in H^{-1}$ satisfies

$$a_{\text{ren}}(i\beta_m; q) = 0 \quad \text{for all } 1 \leq m \leq N.$$

Then for $\kappa \geq 1 + \|q\|_{H^{-1}}^2$ we have

$$\alpha(\kappa; q) \geq -\sum_{m=1}^N \left[\ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \frac{2\beta_m}{\kappa} \right]. \quad (2.20)$$

Moreover, if equality holds in (2.20) for one such κ then it holds for all such κ .

Proof. Let $\{f_n\}_{n \geq 1}$ be a sequence of Schwartz functions that converge to q in H^{-1} . As the renormalized perturbation determinant is continuous on H^{-1} , we have

$$\lim_{n \rightarrow \infty} a_{\text{ren}}(k; f_n) = a_{\text{ren}}(k; q) \quad \text{uniformly on compact subsets of } \mathbb{C}^+.$$

Using Hurwitz's theorem and (1.8), we deduce that for each $1 \leq m \leq N$ and n sufficiently large there exists distinct $\beta_m^{(n)}$ so that

$$a(i\beta_m^{(n)}; f_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \beta_m^{(n)} = \beta_m. \quad (2.21)$$

Note that it is quite possible that $a(\kappa; f_n)$ and/or $a(\kappa; q)$ have additional zeros.

Using the positivity (2.14) to discard additional zeros, we see that (2.15) yields

$$\alpha(\kappa; f_n) \geq - \sum_{m=1}^N \ln \left(\frac{\kappa - \beta_m^{(n)}}{\kappa + \beta_m^{(n)}} \right) + \frac{2\beta_m^{(n)}}{\kappa} \xrightarrow{n \rightarrow \infty} - \sum_{m=1}^N \ln \left(\frac{\kappa - \beta_m}{\kappa + \beta_m} \right) + \frac{2\beta_m}{\kappa}. \quad (2.22)$$

The claim (2.20) now follows because α is continuous on H^{-1} .

Suppose now that equality holds in (2.20) for some single value $\kappa_0 \geq 1 + \|q\|_{H^{-1}}^2$. Let us write $d\mu_n$ for the measure representing $a(k; f_n)$ in the sense of (2.17). It then follows from (2.19) and (2.21) that

$$\int \frac{t^2}{\kappa_0(t^2 + \kappa_0^2)} d\mu_n(t) \rightarrow 0 \quad \text{and thence that} \quad \int \frac{t^2}{\kappa(t^2 + \kappa^2)} d\mu_n(t) \rightarrow 0$$

for every $\kappa > 0$. This in turn guarantees that equality holds in (2.20) for every $\kappa \geq 1 + \|q\|_{H^{-1}}^2$. $\square$

We are now ready to realize the ultimate goal of this section:

Proof of Theorem 1.4. In view of Corollary 2.4, it remains to show that if (1.12) holds and

$$\alpha(\kappa; q) = - \sum_{m=1}^N \ln \left(\frac{\kappa - \beta_m}{\kappa + \beta_m} \right) + \frac{2\beta_m}{\kappa} \quad \text{for all } \kappa \geq 1 + \|q\|_{H^{-1}}^2, \quad (2.23)$$

then $q = Q_{\vec{\beta}, \vec{c}}$ for some choice of $\vec{c} \in \mathbb{R}^N$. The first step in the proof will be to show that all such optimizers q belong to Schwartz class. In the second step, we will prove that

$$a(k; q) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m}. \quad (2.24)$$

In view of Proposition 2.2, this implies that q is a multisoliton, thus completing the proof of the theorem.

From (2.3), (2.14), and (2.23), we see already that $q \in L^2$. We will get further regularity and decay by studying the Euler–Lagrange equation satisfied by q . To begin, we note that since $a_{\text{ren}}(i\beta_m; q) = 0$, there exist $\phi_m \in L^2$ such that

$$(1 + \sqrt{R_0(i\beta_m)q}\sqrt{R_0(i\beta_m)})\phi_m = 0 \quad \text{and} \quad \|\phi_m\|_2 = 1.$$

Writing $\psi_m := \sqrt{R_0(i\beta_m)}\phi_m$ we obtain

$$(-\partial^2 + q + \beta_m^2)\psi_m = 0 \quad \text{and} \quad \|\psi_m\|_{H_{\beta_m}^1} = 1. \quad (2.25)$$

Note also that the eigenvalue $-\beta_m^2$ must be simple. Indeed, if there were two linearly independent eigenfunctions $\psi_m(x), \eta_m(x) \in H^1$, then the Wronskian

$$\psi'_m(x)\eta_m(x) - \psi_m(x)\eta'_m(x),$$

which is independent of x , would need to be both L^1 and (by virtue of linear independence) non-zero; this is clearly inconsistent.

As $q \in L^2$, we see that (2.25) implies that $\psi_m \in H^2$ and so $\psi_m^2 \in L^1 \cap H^2$. Moreover, a quick computation shows that

$$(-\partial^3 + 2\partial q + 2q\partial + 4\kappa^2\partial)\psi_m^2 = 4(\kappa^2 - \beta_m^2)(\psi_m^2)' \quad (2.26)$$

in H^{-1} sense.

Next, we claim that the functions $\{\psi_m^2\}_{m=1}^N$ are linearly independent. Indeed, assume (towards a contradiction) that there were a *minimal* collection $\Lambda \subseteq \{1, \dots, N\}$ such that

$$\sum_{m \in \Lambda} c_m \psi_m^2 = 0 \quad \text{with} \quad c_m \neq 0 \quad \text{for all} \quad m \in \Lambda. \quad (2.27)$$

Fixing some $n \in \Lambda$ and applying $(-\partial^3 + 2\partial q + 2q\partial + 4\beta_n^2\partial)$ to (2.27) and using (2.26) we obtain

$$\sum_{m \neq n \in \Lambda} c_n (\beta_m^2 - \beta_n^2) (\psi_m^2)' = 0.$$

As $\beta_1, \dots, \beta_N$ are distinct and ψ_m^2 decay at infinity, this contradicts the minimality of the collection Λ .

The functions ψ_m^2 represent the gradients of the constraints $a_{\text{ren}}(i\beta_m; \phi) = 0$. Indeed,

$$\frac{\delta}{\delta q} a_{\text{ren}}(i\beta_m; q) = \det_2 \big(1 + \sqrt{R_0(i\beta_m)} q \sqrt{R_0(i\beta_m)} \big) \psi_m^2,$$

where the subscript on $\det_2$ indicates the Hilbert space over which the renormalized determinant is computed. Concretely, in this case this is the Hilbert space of functions orthogonal to ϕ_m . As the eigenvalues $-\beta_m^2$ are simple, the renormalized determinant over $\phi_m^\perp$ is non-zero.

The gradient of α is easily derived from the series (2.1):

$$\frac{\delta}{\delta q} \alpha(\kappa; q) = \frac{1}{2\kappa} - g(\kappa; q),$$

where $g(\kappa; q)$ is the diagonal Green's function. This is discussed in greater detail in [20]. As $q \in L^2$, [20, Proposition A.2] shows that $\frac{1}{2\kappa} - g(\kappa; \phi) \in H^2$; we also have the long-known identity

$$(-\partial^3 + 2\partial q + 2q\partial + 4\kappa^2\partial)g(\kappa; q) = 0$$

which holds in H^{-1} sense (cf. [20, Proposition 2.3]).

As the gradients ψ_m^2 of the constraints have been shown to be linearly independent, we deduce that the optimizer q satisfies the Euler–Lagrange equation

$$\frac{1}{2\kappa} - g(\kappa; q) = \sum_{m=1}^N \lambda_m \psi_m^2 \quad (2.28)$$

for each $\kappa \geq 1 + \|q\|_{H^{-1}}^2$ and some (κ -dependent) multipliers $\lambda_1, \dots, \lambda_N \in \mathbb{R}$. Consequently, applying $(-\partial^3 + 2\partial q + 2q\partial + 4\kappa^2\partial)$ to (2.28) and using (2.26), we deduce that

$$\frac{1}{\kappa} q' = \sum_{m=1}^N 4\lambda_m (\kappa^2 - \beta_m^2) (\psi_m^2)'.$$

However, $q \in L^2$ and $\psi_m^2 \in H^2$; thus

$$\frac{1}{\kappa} q = \sum_{m=1}^N 4\lambda_m (\kappa^2 - \beta_m^2) \psi_m^2 \quad (2.29)$$

and so $q \in H^2$. By alternately applying (2.25) and (2.29), we deduce that q is infinitely smooth.

From (2.29) we see that $q \in L^1$. It then follows from (2.25) that each eigenfunction decays exponentially; see [9, §3.8]. Applying (2.29) again we deduce that q decays exponentially. Thus $q \in \mathcal{S}$.

It remains to prove (2.24). Now that we know $q \in \mathcal{S}$, we may deploy the technology used in the proof of Proposition 2.3. First we note that (2.15) and the positivity (2.14) guarantee that a_{ren} has no zeros beyond those prescribed in (1.12). In this way, the representation (2.17) yields

$$a(k; q) = \exp \left\{ -i \int \frac{d\mu(t)}{t - k} \right\} \cdot \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m}$$

for some finite positive measure $d\mu$ on $\mathbb{R}$. On comparing (2.19) and (2.23), we see that any mass $d\mu$ must be concentrated at the origin. As the holomorphic function $a(k; q)$ admits a continuous extension to $\partial\mathbb{C}^+$, we find that $d\mu \equiv 0$ and so (2.24) holds. $\square$

3. Molecular Decomposition of Multisolitons

The principal goal of this section is to prove Proposition 3.1. This shows that linear combinations of well-separated multisolitons may be well approximated by a single multisoliton. We refer to this as a molecular decomposition building on the analogy of one-solitons to atoms and of multisolitons to molecules.

We will see that the eigenvalue parameters $\vec{\beta}^j$ of the individual multisolitons in this rarefied gas form a partition of the eigenvalue parameters $\vec{\beta}$ of the single approximating multisoliton. Equivalently, $\vec{\beta}$ is built as the concatenation of the $\vec{\beta}^j$, which we write as $\coprod \vec{\beta}^j$. We do not allow any repetitions of individual eigenvalue parameters, not only within individual $\vec{\beta}^j$, but across the entire collection $\coprod \vec{\beta}^j$.

The interrelation between the position parameters $\vec{c}^j$ of the individual molecules and that of the approximating multisoliton is much more subtle since it must accommodate the correct combination of phase-shifts; see (3.3).

Proposition 3.1. *Let multisoliton parameters $\vec{\beta}^j$ and $\vec{c}^j$ be given for each $1 \leq j \leq J$, with no eigenvalue parameter repeated. For any J -tuple of sequences x_n^j satisfying*

$$\lim_{n \rightarrow \infty} (x_n^j - x_n^i) = \infty \quad \text{for all } 1 \leq i < j \leq J, \quad (3.1)$$

there exists a sequence $\vec{c}_n$ so that setting $\vec{\beta} = \coprod \vec{\beta}^j$, we have

$$Q_{\vec{\beta}, \vec{c}_n}(x) - \sum_{j=1}^J Q_{\vec{\beta}^j, \vec{c}^j}(x - x_n^j) \longrightarrow 0 \quad (3.2)$$

in $L^2(\mathbb{R})$ sense as $n \rightarrow \infty$.

The decoupling requirement (3.1) could be stated with absolute values without affecting the conclusion of the theorem. However, ordering the translation parameters from the start makes the proof much easier to explain.

The scenario analyzed here is something of a reverse of the long-time asymptotics of multisolitons. In that scenario, one starts with a multisoliton $Q_{\vec{\beta}, \vec{c}(t)}$, with the components of $\vec{c}(t)$ satisfying an analogue of (3.1) as $t \rightarrow \infty$, and the goal is to find positions $x^j(t)$ so that $Q_{\vec{\beta}, \vec{c}(t)}$ can be approximated by a linear combination of one-solitons as $t \rightarrow \infty$. Despite these differences, we still feel that our approach to treating the error terms could streamline discussions of that subject too.

Each of the multisolitons appearing in (3.2) is defined via the determinant of a matrix and each matrix is potentially of a different size. We need a prudent means of indexing all these matrices. For each $1 \leq j \leq N$, let I^j denote (disjoint) index sets of size $\#\vec{\beta}^j$ (the number of entries in $\vec{\beta}^j$). We will then use $I = \coprod I^j$ as our indexing set of size $\#\vec{\beta}$.

Our first application of these notations is to give a formula for the sequence c_n needed for Proposition 3.1: For $\mu \in I^j$,

$$(c_n)_\mu = x_n^j + c_\mu^j - \frac{1}{\beta_\mu} \sum_{\sigma} \ln \left[\frac{\beta_\sigma - \beta_\mu}{\beta_\sigma + \beta_\mu} \right], \quad (3.3)$$

where the sum extends over all $\sigma \in I^\ell$ for all $\ell > j$.

We also need to construct two families of matrices: For fixed $1 \leq j \leq J$ we define a matrix $B^{(j)}(x; \vec{\beta}, \vec{c})$ indexed over $I \times I$ by

$B_{\mu\nu}^{(j)}(x; \vec{\beta}, \vec{c})$	$\nu \in I^\ell, \ell < j$	$\nu \in I^j$	$\nu \in I^\ell, \ell > j$	(3.4)
$\mu \in I^\ell, \ell < j$	$\delta_{\mu\nu}$	0	0	
$\mu \in I^j$	0	$A_{\mu\nu}(x)$	$\frac{1}{\beta_\mu + \beta_\nu} e^{-\beta_\mu(x - c_\mu)}$	
$\mu \in I^\ell, \ell > j$	0	$\frac{1}{\beta_\mu + \beta_\nu} e^{-\beta_\nu(x - c_\nu)}$	$\frac{1}{\beta_\mu + \beta_\nu}$	

where $A_{\mu\nu}(x)$ is as in (1.4). Similarly, we define

$E_{\mu\nu}^{(j)}(x; \vec{\beta}, \vec{c})$	$\nu \in I^\ell, \ell < j$	$\nu \in I^j$	$\nu \in I^\ell, \ell > j$	(3.5)
$\mu \in I^\ell, \ell < j$	$A_{\mu\nu}(x) - \delta_{\mu\nu}$	$A_{\mu\nu}(x)$	$\frac{1}{\beta_\mu + \beta_\nu} e^{-\beta_\mu(x - c_\mu)}$	
$\mu \in I^j$	$A_{\mu\nu}(x)$	0	0	
$\mu \in I^\ell, \ell > j$	$\frac{1}{\beta_\mu + \beta_\nu} e^{-\beta_\nu(x - c_\nu)}$	0	$\delta_{\mu\nu} e^{\beta_\mu(x - c_\mu) + \beta_\nu(x - c_\nu)}$	

As we shall see, $B^{(j)}$ is the dominant term for those x near x_n^k , while $E^{(j)}$ functions as an error term.

Lemma 3.2. Fix $L > 0$. Then under the hypotheses of Proposition 3.1,

$$\limsup_{n \rightarrow \infty} \|B^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n)\| < \infty \quad \text{and} \quad \limsup_{n \rightarrow \infty} \|E^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n)\| = 0 \quad (3.6)$$

uniformly for $z \in [-L, L] + i[-1, 1]$. Moreover, for all $x \in \mathbb{R}$,

$$Q_{\vec{\beta}, \vec{c}_n}(x) = -2 \frac{d^2}{dx^2} \ln \det[B^{(j)}(x; \vec{\beta}, \vec{c}_n) + E^{(j)}(x; \vec{\beta}, \vec{c}_n)]. \quad (3.7)$$

Proof. As we are dealing with finite matrices, our claims about the operator norm can be verified considering each matrix entry individually. From this perspective, the claim (3.6) follows simply from the behavior of $x_n^j - (c_n)_\mu$: this is bounded when $\mu \in I^j$; it diverges to $+\infty$ when $\mu \in I^\ell$ with $\ell < j$; and it diverges to $-\infty$ when $\mu \in I^\ell$ with $\ell > j$.

The claim (3.7) follows readily from the identity

$$\det[A_{\vec{\beta}, \vec{c}_n}(x)] = \det[B^{(j)}(x; \vec{\beta}, \vec{c}_n) + E^{(j)}(x; \vec{\beta}, \vec{c}_n)] \times \prod e^{-2\beta_\mu(x-c_\mu)},$$

where the product is taken over those $\mu \in I^\ell$ for each $\ell < j$. This product appears because common factors have been extracted from these rows and columns. $\square$

As a stepping-stone to our analysis of $B^{(j)}$ in Lemma 3.4, we first make preparations for evaluating its determinant. In the case $D_{\mu\nu} \equiv 0$, our next lemma relates two Cauchy determinants (as they are known); indeed, it provides the basic inductive step for the complete evaluation of such determinants.

Lemma 3.3 (A Cauchy-like Determinant). *Given an $N \times N$ matrix D , real numbers $a_1, \dots, a_N$, and positive $\beta_1, \dots, \beta_{N+1}$, we define*

$$\tilde{a}_\mu = \frac{\beta_{N+1} - \beta_\mu}{\beta_{N+1} + \beta_\mu} a_\mu.$$

Then we have the following identity between two determinants:

$$\left| D_{\mu\nu} + \frac{a_\mu a_\nu}{\beta_\mu + \beta_\nu} \frac{a_\mu}{\beta_\mu + \beta_{N+1}} \frac{1}{\beta_{N+1} + \beta_\nu} \right| = \frac{1}{2\beta_{N+1}} \left| D_{\mu\nu} + \frac{\tilde{a}_\mu \tilde{a}_\nu}{\beta_\mu + \beta_\nu} \right|. \quad (3.8)$$

On the right, we have an $N \times N$ determinant. The one on the left is $(N+1) \times (N+1)$, with the extra row and column as indicated.

Proof. This is a simple matter of applying row and column operations: First we subtract a_μ times the bottom row of LHS (3.8) from the μ^{th} row and use the identity

$$\frac{1}{\beta_\mu + \beta_\nu} - \frac{1}{\beta_{N+1} + \beta_\nu} = \frac{\beta_{N+1} - \beta_\mu}{(\beta_{N+1} + \beta_\nu)(\beta_\mu + \beta_\nu)}.$$

Extracting the common factor from the final column, this yields

$$\text{LHS}(3.8) = \frac{1}{2\beta_{N+1}} \left| D_{\mu\nu} + \frac{\hat{a}_\mu \check{a}_\nu}{\beta_\mu + \beta_\nu} \frac{\hat{a}_\mu}{\beta_\mu + \beta_{N+1}} \frac{1}{\check{a}_\nu} \right|$$

with $\hat{a}_\mu = (\beta_{N+1} - \beta_\mu)a_\mu$ and $\check{a}_\nu = \frac{a_\nu}{\beta_{N+1} + \beta_\nu}$.

Next we subtract $\check{a}_\nu$ times the last column from the ν^{th} and apply the identity

$$\frac{1}{\beta_\mu + \beta_\nu} - \frac{1}{\beta_\mu + \beta_{N+1}} = \frac{\beta_{N+1} - \beta_\nu}{(\beta_\mu + \beta_\nu)(\beta_\mu + \beta_{N+1})}.$$

The result then follows since the bottom row is now populated by zeros, excepting a one in the final position. $\square$

Lemma 3.4. Fix $L > 0$. Under the hypotheses on Proposition 3.1, there exists $\delta > 0$ so that

$$|\det[B^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n)]| \gtrsim 1 \quad (3.9)$$

uniformly for $n \in \mathbb{N}$ and $z \in [-L, L] + i[-\delta, \delta]$. Moreover, for every $x \in \mathbb{R}$,

$$-2 \frac{d^2}{dx^2} \ln \det[B^{(j)}(x; \vec{\beta}, \vec{c}_n)] = Q_{\vec{\beta}^j, \vec{c}^j}(x - x_n^j). \quad (3.10)$$

Proof. Applying Lemma 3.3 iteratively, we find that

$$\det[B^{(j)}(x; \vec{\beta}, \vec{c}_n)] = \left(\prod_{\sigma} \frac{1}{2\beta_{\sigma}} \right) \cdot \det \left[\delta_{\mu\nu} + \frac{\tilde{a}_{\mu} \tilde{a}_{\nu}}{\beta_{\mu} + \beta_{\nu}} \right]_{I^j \times I^j} \quad (3.11)$$

where the parameters $\tilde{a}_{\mu}$ (which also depend on n) are given by

$$\tilde{a}_{\mu} = \exp\{-\beta_{\mu}(x - (c_n)_{\mu})\} \cdot \prod_{\sigma} \frac{\beta_{\sigma} - \beta_{\mu}}{\beta_{\sigma} + \beta_{\mu}},$$

and both products extend over all $\sigma \in I^{\ell}$ for all $\ell > j$. Referring back to Definition 1.1 and (3.3), we see that we have succeeded in proving (3.10).

When z is real, the inequality (3.9) follows from (3.11) because the matrix C with entries $C_{\mu\nu} = \frac{\tilde{a}_{\mu} \tilde{a}_{\nu}}{\beta_{\mu} + \beta_{\nu}}$ is bounded and (strictly) positive definite (as is easily verified from (3.8) and Sylvester's criterion).

To extend the bound to complex z , it suffices to show that we can choose $\delta > 0$ so that every eigenvalue of the matrix C has positive real part. Writing $z = x + iy$, we see that it suffices to prove that

$$\sum \overline{\psi_{\mu}} [\cos(\beta_{\mu} y) \cos(\beta_{\nu} y) - \sin(\beta_{\mu} y) \sin(\beta_{\nu} y)] \frac{\tilde{a}_{\mu}(x) \tilde{a}_{\nu}(x)}{\beta_{\mu} + \beta_{\nu}} \psi_{\nu} \geq 0$$

for every complex vector ψ_{μ} . Thus, we see that there is such a choice of $\delta > 0$ because of the boundedness and positive-definiteness of C for z real. $\square$

Proof of Proposition 3.1. Our first goal is to prove the following variant of (3.2):

$$Q_{\vec{\beta}, \vec{c}_n}(x + x_n^j) \longrightarrow Q_{\vec{\beta}^j, \vec{c}^j}(x) \quad \text{as } n \rightarrow \infty, \quad (3.12)$$

uniformly for $x \in [-L, L]$ for each fixed j and any fixed $L > 0$.

Combining Cramer's rule with the Hadamard inequality, we find

$$\|B^{(j)}(x; \vec{\beta}, \vec{c}_n)^{-1} E^{(j)}(x; \vec{\beta}, \vec{c}_n)\| \lesssim \frac{\|B^{(j)}(x; \vec{\beta}, \vec{c}_n)\|^{\#\vec{\beta}-1} \|E^{(j)}(x; \vec{\beta}, \vec{c}_n)\|}{|\det B^{(j)}(x; \vec{\beta}, \vec{c}_n)|}$$

where the implicit constant depends only on $\#\vec{\beta}$. Thus, it follows from (3.9) and Lemma 3.2 that for each $L > 0$ there is a $\delta > 0$ so that

$$\ln \det[B^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n) + E^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n)] = \ln \det[B^{(j)}(x_n^j + z; \vec{\beta}, \vec{c}_n)] + o(1)$$

as $n \rightarrow \infty$ uniformly for $z \in [L, -L] + i[-\delta, \delta]$. Because we have convergence in a complex neighbourhood of each x , this convergence extends to all derivatives. Thus (3.12) follows from (3.7) and (3.10).

From (3.12) we may then infer that as $n \rightarrow \infty$,

$$\int_{E_n} \left| Q_{\tilde{\beta}, \tilde{c}_n}(x) - \sum_{j=1}^J Q_{\tilde{\beta}^j, \tilde{c}^j}(x - x_n^j) \right|^2 dx \longrightarrow 0, \quad (3.13)$$

where $E_n = \cup_j [x_n^j - L, x_n^j + L]$. On the other hand, from (2.11) we get

$$\int_{\mathbb{R}} |Q_{\tilde{\beta}, \tilde{c}_n}(x)|^2 dx = \sum_{\mu \in I} \frac{16}{3} \beta_\mu^3 = \sum_{j=1}^J \int |Q_{\tilde{\beta}^j, \tilde{c}^j}(x - x_n^j)|^2 dx.$$

Using this and (3.13) we find that the integral over the complementary region E_n^c makes an asymptotically negligible contribution for L large. Thus (3.2) follows. $\square$

4. Concentration Compactness

The goal of this section is to prove Proposition 4.2, which provides a concentration-compactness principle for the functional $\alpha(\kappa; q)$ acting on bounded equicontinuous sequences in H^{-1} . We begin by recalling the definition of H^{-1} -equicontinuity.

Definition 4.1. A subset $\mathcal{F} \subseteq H^{-1}(\mathbb{R})$ is H^{-1} -equicontinuous if

$$\limsup_{y \rightarrow 0} \sup_{f \in \mathcal{F}} \|f(x - y) - f(x)\|_{H_x^{-1}} = 0. \quad (4.1)$$

While this definition makes a strong parallel with those used in the Arzelà–Ascoli and Fréchet–Kolmogorov Theorems, we shall primarily employ the equivalent formulation (4.7) discussed below.

Proposition 4.2 (Concentration compactness principle) *Assume that $\{u_n\}_{n \geq 1}$ is a bounded and equicontinuous sequence in H^{-1} . Passing to a subsequence there exist $J^* \in \{0, 1, 2, \dots\} \cup \{\infty\}$, non-zero profiles $\phi^j \in H^{-1}$, and positions $x_n^j \in \mathbb{R}$ such that for any finite $0 \leq J \leq J^*$ we have the decomposition*

$$u_n(x) = \sum_{j=1}^J \phi^j(x - x_n^j) + r_n^J(x)$$

with the following properties: for each fixed $\kappa \geq 1 + \sup_n \|u_n\|_{H^{-1}}^2$,

$$\lim_{J \rightarrow J^*} \lim_{n \rightarrow \infty} \operatorname{tr} \{ (\sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)})^4 \} = 0, \quad (4.2)$$

$$\sup_J \lim_{n \rightarrow \infty} \left[\alpha(\kappa; u_n) - \sum_{j=1}^J \alpha(\kappa; \phi^j) - \alpha(\kappa; r_n^J) \right] = 0, \quad (4.3)$$

$$\lim_{n \rightarrow \infty} |x_n^j - x_n^\ell| = \infty \quad \text{for all } j \neq \ell. \quad (4.4)$$

Moreover,

$$\lim_{J \rightarrow J^*} \lim_{n \rightarrow \infty} \left| \alpha(\kappa; r_n^J) - \frac{1}{2\kappa} \int \frac{|\widehat{r_n^J}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \right| = 0. \quad (4.5)$$

Equation (4.2) shows that the remainder is small in the sense that a certain operator is small in $\mathfrak{I}_4$, the trace ideal modeled on ℓ^4 . In fact, it is negligible in any $\mathfrak{I}_p$ with $p > 2$. This follows from (4.2) by means of the basic inequality

$$\|A\|_{\mathfrak{I}_p} \leq \|A\|_{\mathfrak{I}_{p_1}}^\theta \|A\|_{\mathfrak{I}_{p_2}}^{1-\theta} \quad \text{when } 1 \leq p_1 < p < p_2 \leq \infty \text{ and } \theta = \frac{p_1}{p} \cdot \frac{p_2-p}{p_2-p_1}. \quad (4.6)$$

Nonetheless, as (4.5) shows, the remainder term may make a significant contribution to α . We shall ultimately see that optimizing sequences must have negligible remainder term because it contributes *too much* to α .

The nucleus of the proof of Proposition 4.2 is the inverse inequality Lemma 4.3. It shows that non-trivial $\mathfrak{I}_4$ norm may be attributed to the existence of a non-trivial profile common to a subsequence of the original sequence u_n . Before stating this lemma, let us quickly discuss our notations for basic Littlewood-Paley theory; these will be needed in the proof.

For $M \in 2^{\mathbb{Z}}$, we write P_M for the Fourier multiplier operators defined via a partition of unity adapted to the partition $\{\xi \in \mathbb{R} : \frac{1}{2}M < |\xi| \leq 2M\}$ of $\mathbb{R}$. We then define projections onto high and low frequencies via

$$P_{\leq M} f = \sum_{2^{\mathbb{Z}} \ni M' \leq M} P_{M'} f \quad \text{and} \quad P_{\geq M} f = \sum_{2^{\mathbb{Z}} \ni M' \geq M} P_{M'} f.$$

It is not difficult to verify (cf. [20, Lemma 4.2]) that a bounded subset $\mathcal{F} \subseteq H^{-1}(\mathbb{R})$ is H^{-1} -equicontinuous if and only if

$$\limsup_{M \rightarrow \infty} \sup_{f \in \mathcal{F}} \|P_{\geq M} f\|_{H^{-1}} = 0. \quad (4.7)$$

One of the key estimates we need is the Bernstein inequality,

$$\|P_{\leq M} f\|_{L^q} \lesssim M^{\frac{1}{p} - \frac{1}{q}} \|f\|_{L^p} \quad \text{whenever } 1 \leq p \leq q \leq \infty.$$

Lemma 4.3. (Inverse inequality) *Assume $\{u_n\}_{n \geq 1}$ are equicontinuous in H^{-1} and satisfy*

$$\varepsilon < \liminf_{n \rightarrow \infty} \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} u_n \sqrt{R_0(i\kappa)} \right)^4 \right\} \quad \text{and} \quad \limsup_{n \rightarrow \infty} \|u_n\|_{H^{-1}} < A$$

for some positive ε , finite A , and some $\kappa \geq 1 + A^2$. Then passing to a subsequence there exist a non-zero profile $\phi \in H^{-1}$ and positions $x_n \in \mathbb{R}$ such that

$$u_n(x + x_n) \rightharpoonup \phi(x) \quad \text{weakly in } H^{-1},$$

$$\lim_{n \rightarrow \infty} \left[\alpha(\kappa; u_n) - \alpha(\kappa; u_n(\cdot + x_n) - \phi) \right] = \alpha(\kappa; \phi). \quad (4.8)$$

Proof. Passing to a subsequence, we may assume that for all n we have

$$\frac{1}{2}\varepsilon < \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} u_n \sqrt{R_0(i\kappa)} \right)^4 \right\} \quad \text{and} \quad \|u_n\|_{H^{-1}}^2 < 2A^2.$$

For $M \in 2^{\mathbb{N}}$, we use (1.11) to estimate

$$\begin{aligned} \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [P_{\geq M} u_n] \sqrt{R_0(i\kappa)} \right)^4 \right\} &\lesssim \left\| \sqrt{R_0(i\kappa)} [P_{\geq M} u_n] \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{I}_2}^4 \\ &\lesssim \kappa^{-2} \|P_{\geq M} u_n\|_{H_\kappa^{-1}}^4 \lesssim \|P_{\geq M} u_n\|_{H^{-1}}^4 < \frac{1}{8}\varepsilon, \end{aligned} \quad (4.9)$$

provided M is sufficiently large depending on ε , in view of the equicontinuity of u_n .

On the other hand, for dyadic $M \leq 1$ we may use Bernstein to estimate

$$\|P_{\leq M} u_n\|_{L^\infty}^2 \lesssim M \|P_{\leq M} u_n\|_{L^2}^2 \lesssim M^3 \|u_n\|_{H^{-1}}^2$$

and so, also using (1.11), deduce that

$$\begin{aligned} & \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [P_{\leq M} u_n] \sqrt{R_0(i\kappa)} \right)^4 \right\} \\ & \lesssim \left\| \sqrt{R_0(i\kappa)} [P_{\leq M} u_n] \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{H}_2}^2 \|P_{\leq M} u_n\|_{L^\infty}^2 \left\| \sqrt{R_0(i\kappa)} \right\|_{L^2 \rightarrow L^2}^4 \\ & \lesssim \kappa^{-5} M^3 A^4 < \frac{1}{8} \varepsilon, \end{aligned} \quad (4.10)$$

provided M is sufficiently small depending on ε and A .

Therefore, passing to a further subsequence, we deduce that there exists a dyadic M such that

$$\operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [P_M u_n] \sqrt{R_0(i\kappa)} \right)^4 \right\} \geq c(\varepsilon, A) \quad \text{for all } n,$$

where $c(\varepsilon, A)$ is a positive continuous function on $[0, \infty) \times [0, \infty)$. As

$$\operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [P_M u_n] \sqrt{R_0(i\kappa)} \right)^4 \right\} \lesssim \kappa^{-1} \|u_n\|_{H_\kappa^{-1}}^2 \kappa^{-4} \|P_M u_n\|_\infty^2 \lesssim A^2 \|P_M u_n\|_\infty^2,$$

there exists $x_n \in \mathbb{R}$ such that

$$|P_M u_n(x_n)| \gtrsim c(\varepsilon, A) A^{-2}. \quad (4.11)$$

As the sequence $u_n(x + x_n)$ is bounded in H^{-1} , passing to a subsequence we find $\phi \in H^{-1}$ such that

$$u_n(x + x_n) \rightharpoonup \phi(x) \quad \text{weakly in } H^{-1}. \quad (4.12)$$

In view of (4.11), we see that $\phi \neq 0$. In fact, it is not difficult to verify that

$$\operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} \phi \sqrt{R_0(i\kappa)} \right)^4 \right\} \geq \tilde{c}(\varepsilon, A), \quad (4.13)$$

where $\tilde{c}(\varepsilon, A)$ is a positive continuous function on $[0, \infty) \times [0, \infty)$. Indeed, even the operator norm of $\sqrt{R_0(i\kappa)} \phi \sqrt{R_0(i\kappa)}$ satisfies such a lower bound.

It remains to prove the asymptotic decoupling (4.8). To this end, it suffices to show that for all $\ell \geq 2$ we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} u_n \sqrt{R_0(i\kappa)} \right)^\ell \right\} - \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [u_n(\cdot + x_n) - \phi] \sqrt{R_0(i\kappa)} \right)^\ell \right\} \\ & = \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} \phi \sqrt{R_0(i\kappa)} \right)^\ell \right\}. \end{aligned} \quad (4.14)$$

The case $\ell = 2$ of (4.14) follows easily from the weak convergence (4.12) and the fact that H_κ^{-1} is a Hilbert space. Indeed, by (1.11),

$$\begin{aligned} & \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} u_n \sqrt{R_0(i\kappa)} \right)^2 \right\} - \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} [u_n(\cdot + x_n) - \phi] \sqrt{R_0(i\kappa)} \right)^2 \right\} \\ & = \kappa^{-1} \|u_n\|_{H_\kappa^{-1}}^2 - \kappa^{-1} \|u_n(\cdot + x_n) - \phi\|_{H_\kappa^{-1}}^2 \\ & = \kappa^{-1} \|\phi\|_{H_\kappa^{-1}}^2 + 8 \operatorname{Re} \langle R_0(2i\kappa) \phi, u_n(\cdot + x_n) - \phi \rangle_{L^2} \\ & = \operatorname{tr} \left\{ \left(\sqrt{R_0(i\kappa)} \phi \sqrt{R_0(i\kappa)} \right)^2 \right\} + o(1) \quad \text{as } n \rightarrow \infty. \end{aligned}$$

We now turn to the case $\ell \geq 3$ in (4.14). First, combining (4.6) with (4.9) and (4.10), we see that we may discard very high and very low frequencies from further consideration. Specifically, we must prove that (4.14) holds when u_n and ϕ are replaced by $P_{\text{med}}u_n$ and $P_{\text{med}}\phi$, respectively, for any fixed choice of projection $P_{\text{med}} = P_{\geq M_0}P_{\leq M_1}$ onto middle frequencies.

From (4.12) and Rellich's Theorem, we may pass to a further subsequence, if necessary, to ensure that

$$v_n := P_{\text{med}}[u_n(\cdot + x_n) - \phi] \rightarrow 0 \quad \text{uniformly on compact sets.} \quad (4.15)$$

We use this as follows: From the explicit kernel of R_0 , we have that as $n \rightarrow \infty$,

$$\|v_n R_0(i\kappa)[P_{\text{med}}\phi]\|_{\mathfrak{J}_2}^2 = \iint \frac{1}{4\kappa} e^{-2\kappa|x-y|} |v_n(x)|^2 |P_{\text{med}}\phi(y)|^2 \rightarrow 0. \quad (4.16)$$

To continue, we write

$$\begin{aligned} & \text{tr}\{(\sqrt{R_0(i\kappa)}[P_{\text{med}}u_n]\sqrt{R_0(i\kappa)})^\ell\} - \text{tr}\{(\sqrt{R_0(i\kappa)}[P_{\text{med}}\phi]\sqrt{R_0(i\kappa)})^\ell\} \\ & \quad - \text{tr}\{(\sqrt{R_0(i\kappa)}v_n\sqrt{R_0(i\kappa)})^\ell\} \\ & = \sum \text{tr}\{R_0(i\kappa)F_1 R_0(i\kappa)F_2 \cdots R_0(i\kappa)F_\ell\}, \end{aligned}$$

where the sum is over all choices of $F_1, \dots, F_\ell \in \{v_n, P_{\text{med}}\phi\}$ that are not all identical. We estimate

$$\begin{aligned} & |\text{tr}\{R_0(i\kappa)F_1 R_0(i\kappa)F_2 \cdots R_0(i\kappa)F_\ell\}| \\ & \lesssim \|v_n R_0(i\kappa)P_{\text{med}}\phi\|_{\mathfrak{J}_2} \|\sqrt{R_0(i\kappa)}\|_{L^2 \rightarrow L^2}^2 \\ & \quad \times \left[\|\sqrt{R_0(i\kappa)}[P_{\text{med}}u_n]\sqrt{R_0(i\kappa)}\|_{\mathfrak{J}_2} + \|\sqrt{R_0(i\kappa)}[P_{\text{med}}\phi]\sqrt{R_0(i\kappa)}\|_{\mathfrak{J}_2} \right]^{\ell-2} \end{aligned}$$

which converges to zero as $n \rightarrow \infty$ in view of (4.16) and (1.11). $\square$

We are now ready to complete the

Proof of Proposition 4.2. Fix $\kappa_0 = 1 + \sup_n \|u_n\|_{H^{-1}}^2$. We will apply Lemma 4.3 at spectral parameter κ_0 inductively, extracting one profile at a time. To start, we set $r_n^0 := u_n$. Now suppose we have a decomposition up to level $J \geq 0$ satisfying (4.3). Passing to a subsequence if necessary, we set

$$A_J := \lim_{n \rightarrow \infty} \|r_n^J\|_{\dot{H}^{-1}} \quad \text{and} \quad \varepsilon_J := \lim_{n \rightarrow \infty} \text{tr}\{(\sqrt{R_0(i\kappa_0)}r_n^J\sqrt{R_0(i\kappa_0)})^4\}.$$

If $\varepsilon_J = 0$, we stop and set $J^* = J$. If not, we apply Lemma 4.3 at spectral parameter κ_0 to r_n^J . Passing to a subsequence in n , this yields a non-zero profile $\phi^{J+1} \in \dot{H}^{-1}$ and positions $x_n^{J+1} \in \mathbb{R}$ such that

$$\phi^{J+1}(x) = \text{w-lim}_{n \rightarrow \infty} r_n^J(x + x_n^{J+1}). \quad (4.17)$$

To continue, we define $r_n^{J+1}(x) := r_n^J(x) - \phi^{J+1}(x - x_n^{J+1})$. From Lemma 4.3,

$$\lim_{n \rightarrow \infty} [\alpha(\kappa_0; r_n^J) - \alpha(\kappa_0; r_n^{J+1}) - \alpha(\kappa_0; \phi^{J+1})] = 0,$$

which combined with the inductive hypothesis gives (4.3) at the level $J + 1$ and spectral parameter κ_0 . Moreover, from (4.14) we get

$$\begin{aligned} \lim_{n \rightarrow \infty} \operatorname{tr} \{ (\sqrt{R_0(i\kappa_0)} r_n^J \sqrt{R_0(i\kappa_0)})^4 \} &- \operatorname{tr} \{ (\sqrt{R_0(i\kappa_0)} r_n^{J+1} \sqrt{R_0(i\kappa_0)})^4 \} \\ &= \operatorname{tr} \{ (\sqrt{R_0(i\kappa_0)} \phi^{J+1} \sqrt{R_0(i\kappa_0)})^4 \}, \end{aligned}$$

which combined with (4.13) yields

$$\varepsilon_{J+1} \leq \varepsilon_J - c_{J+1}(\varepsilon_J, A_J) \quad (4.18)$$

for some positive function c_{J+1} which is continuous on $[0, \infty) \times [0, \infty)$.

If $\varepsilon_{J+1} = 0$, we stop and set $J^* = J + 1$; in this case, (4.2) at spectral parameter κ_0 is automatic. If $\varepsilon_{J+1} > 0$ we continue the induction. If the algorithm does not terminate in finitely many steps, we set $J^* = \infty$; in this case, (4.18) guarantees that $\varepsilon_J \rightarrow 0$ as $J \rightarrow \infty$ and so (4.2) at spectral parameter κ_0 follows.

Next we confirm that (4.2) and (4.3) hold at *all* spectral parameters $\kappa \geq \kappa_0$. The asymptotic decoupling (4.3) carries over because our argument relies solely on the weak convergence (4.17), as evinced by the proof of (4.14). The claim (4.2) at spectral parameter κ follows from that at spectral parameter κ_0 since

$$\|R_0(i\kappa)^{\frac{1}{2}} R_0(i\kappa_0)^{-\frac{1}{2}}\|_{L^2 \rightarrow L^2} \leq 1.$$

Next we verify the asymptotic orthogonality condition (4.4). We argue by contradiction. Assume (4.4) fails to be true for some pair (j, ℓ) . Without loss of generality, we may assume that this is the first pair for which (4.4) fails, that is, $j < \ell$ and (4.4) holds for all pairs (j, m) with $j < m < \ell$. Passing to a subsequence, we may assume

$$\lim_{n \rightarrow \infty} (x_n^j - x_n^\ell) = x_0. \quad (4.19)$$

From the inductive relation

$$r_n^{\ell-1} = r_n^j - \sum_{m=j+1}^{\ell-1} \phi^m(\cdot - x_n^m),$$

we get

$$\begin{aligned} \phi^\ell(x) &= \operatorname{w}\text{-}\lim_{n \rightarrow \infty} r_n^{\ell-1}(x + x_n^\ell) \\ &= \operatorname{w}\text{-}\lim_{n \rightarrow \infty} r_n^j(x + x_n^\ell) - \sum_{m=j+1}^{\ell-1} \operatorname{w}\text{-}\lim_{n \rightarrow \infty} \phi^m(x + x_n^\ell - x_n^m), \end{aligned} \quad (4.20)$$

where the weak limits are in the H^{-1} topology. That the first limit on the right-hand side of (4.20) is zero follows from (4.19) and the observation that by construction,

$$\operatorname{w}\text{-}\lim_{n \rightarrow \infty} r_n^j(\cdot + x_n^j) = 0.$$

That the remaining limits are zero follows from our assumption that (4.4) holds for all pairs (j, m) with $j < m < \ell$. Thus (4.20) yields $\phi^\ell = 0$, which contradicts the nontriviality of ϕ^ℓ . This completes the proof of (4.4).

Lastly, we prove (4.5):

$$\begin{aligned}
 & \left| \alpha(\kappa; r_n^J) - \frac{1}{2\kappa} \int \frac{|\widehat{r_n^J}(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \right| \\
 & \leq \sum_{\ell \geq 3} \frac{1}{\ell} \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_\ell}^\ell \\
 & \leq \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^2 \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_2} \\
 & \quad + \sum_{\ell \geq 4} \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^4 \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{L^2 \rightarrow L^2}^{\ell-4} \\
 & \lesssim \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^2 \kappa^{-\frac{1}{2}} \|r_n^J\|_{H_\kappa^{-1}} \\
 & \quad + \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^4 \sum_{\ell \geq 4} \kappa^{-\frac{\ell-4}{2}} \|r_n^J\|_{H_\kappa^{-1}}^{\ell-4} \\
 & \lesssim_A \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^2 + \left\| \sqrt{R_0(i\kappa)} r_n^J \sqrt{R_0(i\kappa)} \right\|_{\mathfrak{J}_4}^4.
 \end{aligned}$$

Thus, using (4.2) we deduce (4.5) $\square$

5. Orbital Stability

This section is dedicated to the proof of Theorem 1.2. We argue by contradiction.

Fix $N \geq 1$ and distinct positive parameters $\beta_1, \dots, \beta_N$. Assume, towards a contradiction, that there exist $\varepsilon_0 > 0$, initial data $q_n(0) \in H^{-1}$, and times $t_n \in \mathbb{R}$ such that

$$\inf_{\vec{c} \in \mathbb{R}^N} \|q_n(0) - \mathcal{Q}_{\vec{\beta}, \vec{c}}\|_{H^{-1}} \longrightarrow 0 \quad \text{as } n \rightarrow \infty \quad (5.1)$$

but

$$\inf_{\vec{c} \in \mathbb{R}^N} \|q_n(t_n) - \mathcal{Q}_{\vec{\beta}, \vec{c}}\|_{H^{-1}} \geq \varepsilon_0 \quad \text{for all } n \geq 1. \quad (5.2)$$

Recalling that a_{ren} and α are continuous functions on H^{-1} and conserved by the KdV flow, (5.1), (2.12), and (2.13) imply that

$$\lim_{n \rightarrow \infty} a_{\text{ren}}(k; q_n(t_n)) = \lim_{n \rightarrow \infty} a_{\text{ren}}(k; q_n(0)) = \prod_{m=1}^N \frac{k - i\beta_m}{k + i\beta_m} e^{\frac{2i\beta_m}{k}} \quad (5.3)$$

uniformly for k in compact subsets of $\mathbb{C}^+$ and

$$\lim_{n \rightarrow \infty} \alpha(\kappa; q_n(t_n)) = \lim_{n \rightarrow \infty} \alpha(\kappa; q_n(0)) = - \sum_{m=1}^N \ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \frac{2\beta_m}{\kappa} \quad (5.4)$$

uniformly for $\kappa \geq 1 + \frac{512}{3} \sum_m \beta_m^3$. With a view to future needs, our bound on κ combines the restriction needed for (2.5) with (2.11) and the embedding $L^2 \hookrightarrow H^{-1}$.

By Hurwitz's theorem and (5.3), we deduce that for each $1 \leq m \leq N$ and n sufficiently large, there exist $\beta_m^{(n)}$ such that

$$a_{\text{ren}}(i\beta_m^{(n)}; q_n(t_n)) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \beta_m^{(n)} = \beta_m. \quad (5.5)$$

Using (2.5), (5.4), and the notation from (2.14), we obtain

$$\|q_n(t_n)\|_{H_\kappa^{-1}}^2 \leq 4\kappa\alpha(\kappa; q_n(t_n)) \xrightarrow{n \rightarrow \infty} \sum_{m=1}^N 4\kappa G\left(\frac{\beta_m}{\kappa}\right).$$

As the right-hand side above converges to zero as $\kappa \rightarrow \infty$, invoking (4.7) we deduce that the sequence $u_n := q_n(t_n)$ is equicontinuous in H^{-1} and so we may apply Proposition 4.2. Along a subsequence we may decompose

$$u_n(x) = \sum_{j=1}^J \phi^j(x - x_n^j) + r_n^J(x) \quad (5.6)$$

satisfying the properties (4.2) and (4.3).

Our goal is to prove that there are finitely many profiles, each having the shape of a (multi)soliton, and that r_n^J converges to zero in H^{-1} . First, we rule out the possibility of vanishing. Assume, towards a contradiction, that there are no profiles in (5.6) and so $u_n = r_n$. Invoking (4.5), we obtain

$$\frac{1}{2\kappa} \int \frac{|\widehat{u}_n(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \xrightarrow{n \rightarrow \infty} - \sum_{m=1}^N \left[\ln\left(\frac{\kappa - \beta_m}{\kappa + \beta_m}\right) + \frac{2\beta_m}{\kappa} \right] = \sum_{m=1}^N G\left(\frac{\beta_m}{\kappa}\right).$$

This immediately leads to a contradiction since the functions

$$\kappa \mapsto \frac{\kappa^3}{2\kappa} \int \frac{|\widehat{u}_n(\xi)|^2}{\xi^2 + 4\kappa^2} d\xi \quad \text{and} \quad \kappa \mapsto \sum_{m=1}^N \kappa^3 G\left(\frac{\beta_m}{\kappa}\right)$$

have opposite monotonicity.

Therefore, we may assume that there exists at least one non-trivial profile. From (1.10) and (1.9), we have

$$|1 - a_{\text{ren}}(k; q)| \lesssim \exp\{C(k)\|q\|_{H^{-1}}^2\} \quad (5.7)$$

with $C(k)$ bounded for k in compact subsets of $\mathbb{C}^+$. Consequently, for $J \geq 1$ fixed, the functions

$$f_n : k \mapsto a_{\text{ren}}(k; u_n) \exp\left\{-\frac{i}{2k} \int \frac{|\widehat{r}_n^J(\xi)|^2}{\xi^2 - 4k^2} d\xi\right\}$$

are holomorphic and locally bounded on $\mathbb{C}^+$. Invoking Montel's theorem and passing to a subsequence, we find that this sequence converges as $n \rightarrow \infty$ to a holomorphic function f . Moreover, by (5.5) we have $f(i\beta_m) = 0$ for all $1 \leq m \leq N$.

To continue, we combine (4.3) with (4.5) and (5.4) to obtain

$$\sum_{j=1}^{J^*} \alpha(\kappa; \phi^j) \leq \sum_{m=1}^N G\left(\frac{\beta_m}{\kappa}\right) \quad (5.8)$$

and so by (2.3),

$$\frac{1}{8} \sum_{j=1}^{J^*} \|\phi^j\|_2^2 = \lim_{\kappa \rightarrow \infty} \sum_{j=1}^{J^*} \kappa^3 \alpha(\kappa; \phi^j) \leq \sum_{m=1}^N \lim_{\kappa \rightarrow \infty} \kappa^3 G\left(\frac{\beta_m}{\kappa}\right) < \infty. \quad (5.9)$$

Using this and (5.7), we see that the function $k \mapsto \prod_{j=1}^{J^*} a_{\text{ren}}(k; \phi^j)$ is well defined and holomorphic on $\mathbb{C}^+$. Invoking (4.3) and (4.5) one more time, we conclude that

$$\prod_{j=1}^{J^*} a_{\text{ren}}(k; \phi^j) = f(k) \quad \text{for all } k \in \mathbb{C}^+. \quad (5.10)$$

Let $\vec{\beta}^j$ denote the collection of all zeros of $a_{\text{ren}}(i\kappa; \phi^j)$. Evidently, $\coprod_j \vec{\beta}^j$ enumerates the zeros (with multiplicity) of $f(i\kappa)$, which contains each β_m , $1 \leq m \leq N$. Also, by Corollary 2.4,

$$\alpha(\kappa, \phi^j) \geq \sum_{\beta \in \vec{\beta}^j} G\left(\frac{\beta}{\kappa}\right). \quad (5.11)$$

Contrasting (5.8) and (5.11), we see that $\coprod_j \vec{\beta}^j = \vec{\beta}$ without any repetitions. Moreover, each $\vec{\beta}^j$ must be non-empty, for otherwise $\alpha(\kappa, \phi^j) \equiv 0$ and so $\phi^j \equiv 0$, which is impossible; all profiles are non-zero by construction.

From this we deduce that J^* is finite and, after reviewing (4.3) and (4.5), that $r_n^{J^*} \rightarrow 0$ in H^{-1} sense. More importantly, the comparison of (5.8) and (5.11) shows that each ϕ^j must be an optimizer for the variational problem of Theorem 1.4 with parameters $\vec{\beta}^j$. This theorem then tells us that each ϕ^j is indeed a multisoliton.

Putting this all together, we deduce that

$$u_n(x) = \sum_{j=1}^{J^*} Q_{\vec{\beta}^j, \vec{c}^j}(x - x_n^j) + r_n(x) \quad \text{with} \quad \lim_{n \rightarrow \infty} \|r_n\|_{H^{-1}} = 0. \quad (5.12)$$

In view of Proposition 3.1, this contradicts (5.2) and so completes the proof of Theorem 1.2. $\square$

6. Higher Regularity

The purpose of this section is to prove Corollary 1.3. In fact, we will exhibit two methods for deducing orbital stability at higher regularity from Theorem 1.2. Our first method exploits equicontinuity in the following form:

Lemma 6.1. *Fix $s \in [-1, 1)$ and distinct positive parameters $\beta_1, \dots, \beta_N$. For every $\varepsilon > 0$, there exist $\delta > 0$ and $M \in 2^{\mathbb{Z}}$ so that*

$$\inf_{\vec{c}} \|q(0) - Q_{\vec{\beta}, \vec{c}}\|_{H^s} < \delta \implies \sup_{t \in \mathbb{R}} \|q_{\geq M}(t)\|_{H^s} < \varepsilon. \quad (6.1)$$

Proof. If this assertion were to fail, then there would exist a sequence of solutions q_n and a sequence of times t_n so that

$$\limsup_{n \rightarrow \infty} \inf_{\tilde{c}} \|q_n(0) - Q_{\tilde{\beta}, \tilde{c}}\|_{H^s} = 0,$$

but $\{q_n(t_n) : n \in \mathbb{N}\}$ is not equicontinuous in H^s .

As $\tilde{c}$ varies, the multisolitons $Q_{\tilde{\beta}, \tilde{c}}$ remain uniformly bounded in H^1 ; see, for example, (2.11). Thus this family is H^s -equicontinuous and then so must be the sequence of initial data $q_n(0)$.

When $s = -1$, this directly contradicts the equicontinuity result [20, Proposition 4.4]. The analogous equicontinuity result for $-1 < s < 0$ appears in the proof of [20, Corollary 5.3]. Finally, when $0 \leq s < 1$, we may appeal to [21, Proposition 3.6]. While this last-quoted result does not explicitly assert equicontinuity, the simplicity with which this may be derived from what is presented there is illustrated (in the $s = 0$ case) in [20, Proposition A.3(c)]. $\square$

We do not believe that the restriction $s \in [-1, 1)$ appearing in Lemma 6.1 represents an intrinsic limitation: it merely reflects the range for which equicontinuity of orbits (for equicontinuous sets of initial data) has already been proven in the literature. We now demonstrate Corollary 1.3 for this range of s :

Proof of Corollary 1.3 for $s \in (-1, 1)$. For any $M \in 2^{\mathbb{Z}}$ and any pair of solutions,

$$\|q(t) - Q(t)\|_{H^s}^2 \lesssim M^{2+2s} \|q(t) - Q(t)\|_{H^{-1}}^2 + M^{2s-2} \|P_{\geq M} Q(t)\|_{H^1}^2 + \|P_{\geq M} q(t)\|_{H^s}^2.$$

The result now follows from Theorem 1.2 and Lemma 6.1. $\square$

Let us turn now to a second method for deducing orbital stability at higher regularity. This is based on conservation laws. The precise property we need from such conservation laws is given in the next lemma.

Lemma 6.2. *For each $s > -1$ there is a continuous function $E_s : H^s(\mathbb{R}) \rightarrow \mathbb{R}$ that is conserved under the KdV hierarchy and admits the representation*

$$E_s(q) = \|q\|_{H^s}^2 + \tilde{E}_s(q)$$

where the function $\tilde{E}_s : H^s(\mathbb{R}) \rightarrow \mathbb{R}$ admits the Hölder-type estimate

$$|\tilde{E}_s(q) - \tilde{E}_s(r)| \lesssim \|q - r\|_{H^{-1}}^\theta \exp\{\|q\|_{H^s}^2 + \|r\|_{H^s}^2\} \quad (6.2)$$

for some $\theta > 0$, which depends only on s .

Proof. For integer $s \geq 0$, we may use the traditional polynomial conservation laws discovered in [32]. For example, using the notation from (1.2), we may define $E_0(q) = 2P(q)$ and $E_1(q) = 2H(q) + 2P(q)$. In the former case, $\tilde{E}_0 \equiv 0$ and (6.2) is trivial; in the latter case, (6.2) follows from the estimate

$$\begin{aligned} | \|q\|_{L^2}^2 - \|r\|_{L^2}^2 | + \|q^3 - r^3\|_{L^1} &\lesssim \|q - r\|_{L^2} \exp\left\{\frac{1}{2}\|q\|_{H^1}^2 + \frac{1}{2}\|q\|_{H^1}^2\right\} \\ &\lesssim \|q - r\|_{H^{-1}}^{\frac{1}{2}} \exp\{\|q\|_{H^1}^2 + \|r\|_{H^1}^2\}. \end{aligned}$$

Of course, the exponential factor is very wasteful; however, there is no advantage in keeping track of the optimal power-law dependence.

For larger $s \in \mathbb{N}$, we use the fact that the polynomial conservation laws can be normalized so that they take the form

$$E_s(q) = \int [q^{(s)}(x)]^2 + P_s(q^{(s-1)}(x), \dots, q''(x), q'(x), q(x)) dx$$

where P_s is a polynomial with all terms of degree two or higher. (This was shown in [25].) The fact that E_s is homogeneous under the rescaling $q(x) \mapsto \lambda^2 q(\lambda x)$ further constrains the structure of this polynomial; however, this additional information is not required to verify (6.2). Indeed, Hölder's inequality shows

$$\begin{aligned} |\tilde{E}_s(q) - \tilde{E}_s(r)| &\lesssim \|q - r\|_{H^{s-1}} \exp\left\{\frac{1}{2}\|q\|_{H^s}^2 + \frac{1}{2}\|r\|_{H^s}^2\right\} \\ &\lesssim \|q - r\|_{H^{-1}}^{\frac{1}{s+1}} \exp\{\|q\|_{H^s}^2 + \|r\|_{H^s}^2\}. \end{aligned}$$

The existence of conservation laws that control the H^s norm for non-integer $s > -1$ is a very recent discovery. The lemma in this case, follows from [22, Theorem 9.1]. As written, this theorem only gives pointwise bounds on $\tilde{E}_s$; however, these guarantee derivative bounds because it is shown that $E_s(q)$ is an analytic function of q . Additionally, [22, Theorem 9.1] is stated for a small neighborhood of the origin; this restriction is easily removed by employing scaling. $\square$

Proof of Corollary 1.3 for $s > -1$. For any pair of functions $q, r \in H^s$, we have

$$\begin{aligned} \|q - r\|_{H^s}^2 &= \|q\|_{H^s}^2 - \|r\|_{H^s}^2 - 2\langle q - r, r \rangle_{H^s} \\ &= E_s(q) - E_s(r) - 2\langle q - r, r \rangle_{H^s} + \tilde{E}_s(q) - \tilde{E}_s(r). \end{aligned}$$

Here E_s denotes the conserved quantity discussed in Lemma 6.2. Employing that result, we deduce that for some $\theta \in (0, 1)$,

$$\begin{aligned} \left| \|q - r\|_{H^s}^2 - E_s(q) + E_s(r) \right| &\lesssim \|q - r\|_{H^{-1}} \|r\|_{H^{2s+1}} \\ &\quad + \|q - r\|_{H^{-1}}^\theta \exp\{\|q\|_{H^s}^2 + \|r\|_{H^s}^2\} \\ &\lesssim \|q - r\|_{H^{-1}}^\theta \exp\{\|q\|_{H^s}^2 + \|r\|_{H^{2s+1}}^2\}. \end{aligned}$$

Now let us choose q to be an H^s -solution of (KdV) and use that $E_s(q(t))$ is independent of t . We take $r = Q_{\tilde{\beta}, \tilde{c}}$ with $\tilde{\beta}$ fixed, but $\tilde{c}$ allowed to vary. As $E_s(q)$ is determined through $a(k; q)$, it follows that $E_s(Q_{\tilde{\beta}, \tilde{c}})$ is independent of $\tilde{c}$. In this way, we deduce that

$$\begin{aligned} \inf_{\tilde{c}} \|q(t) - Q_{\tilde{\beta}, \tilde{c}}\|_{H^s}^2 &\leq \inf_{\tilde{c}} \|q(0) - Q_{\tilde{\beta}, \tilde{c}}\|_{H^s}^2 \\ &\quad + C \inf_{\tilde{c}} \|q(0) - Q_{\tilde{\beta}, \tilde{c}}\|_{H^{-1}}^\theta \exp\{\|q(0)\|_{H^s}^2 + \|Q_{\tilde{\beta}, \tilde{c}}\|_{H^{2s+1}}^2\} \\ &\quad + C \inf_{\tilde{c}} \|q(t) - Q_{\tilde{\beta}, \tilde{c}}\|_{H^{-1}}^\theta \exp\{\|q(t)\|_{H^s}^2 + \|Q_{\tilde{\beta}, \tilde{c}}\|_{H^{2s+1}}^2\} \end{aligned}$$

for some s -dependent constant $C > 0$.

To complete the proof of the corollary, it merely remains to invoke Theorem 1.2 and to use the conservation laws of [22, 32] in a more traditional way:

$$\|q(t)\|_{L_t^\infty H^\sigma} \lesssim_\sigma \|q(0)\|_{H^\sigma} \exp\{\|q(0)\|_{H^\sigma}^2\} \quad \text{and} \quad \sup_{\tilde{c}} \|Q_{\tilde{\beta}, \tilde{c}}\|_{H^\sigma} < \infty$$

for any $\sigma > -1$ and any fixed $\tilde{\beta}$. $\square$

Acknowledgements. R. K. was supported by NSF Grant DMS-1856755 and M. V. by Grant DMS-1763074.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Akhiezer, N.I., Glazman, I.M.: Theory of Linear Operators in Hilbert Space. Two Volumes Bound as One. Dover Publications Inc, New York (1993)
2. Albert, J., Bona, J., Nguyen, N.V.: On the stability of KdV multi-solitons. *Differ. Integral Equ.* **20**(8), 841–878 (2007)
3. Alejo, M.A., Muñoz, C., Vega, L.: The Gardner equation and the L^2 -stability of the N -soliton solution of the Korteweg–de Vries equation. *Trans. Am. Math. Soc.* **365**(1), 195–212 (2013)
4. Benjamin, T.B.: The stability of solitary waves. *Proc. R. Soc. Lond. Ser. A* **328**, 153–183 (1972)
5. Bona, J.: On the stability theory of solitary waves. *Proc. R. Soc. Lond. Ser. A* **344**(1638), 363–374 (1975)
6. Boussinesq, J.: Sur la théorie des eaux courantes. *Mémoires à l'Acad. des Sci. Inst. Nat. France* **23–24**, 1–680 (1877)
7. Buckmaster, T., Koch, H.: The Korteweg–de Vries equation at H^{-1} regularity. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **32**(5), 1071–1098 (2015)
8. Cazenave, T., Lions, P.-L.: Orbital stability of standing waves for some nonlinear Schrödinger equations. *Commun. Math. Phys.* **85**(4), 549–561 (1982)
9. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill Book Company Inc, New York (1955)
10. Deift, P., Trubowitz, E.: Inverse scattering on the line. *Commun. Pure Appl. Math.* **32**(2), 121–251 (1979)
11. Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Method for solving the Korteweg–de Vries equation. *Phys. Rev. Lett.* **19**(19), 1095–1097 (1967)
12. Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Korteweg–de Vries equation and generalization. VI. Methods for exact solution. *Commun. Pure Appl. Math.* **27**, 97–133 (1974)
13. Grünert, K., Teschl, G.: Long-time asymptotics for the Korteweg–de Vries equation via nonlinear steepest descent. *Math. Phys. Anal. Geom.* **12**(3), 287–324 (2009)
14. Hilbert, D.: Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen (Erste Mitteilung). *Nachr. Ges. Wiss. Göttingen* **49–91** (1904)
15. Hirota, R.: Exact solution of the Korteweg–de Vries equation for multiple collisions of solitons. *Phys. Rev. Lett.* **27**(18), 1192–1194 (1971)
16. Jost, R., Pais, A.: On the scattering of a particle by a static potential. *Physical Rev. (2)* **82**(6), 840–851 (1951)
17. Kay, I., Moses, H.E.: Reflectionless transmission through dielectrics and scattering potentials. *J. Appl. Phys.* **27**(12), 1503–1508 (1956)
18. Kenig, C.E., Ponce, G., Vega, L.: Well-posedness of the initial value problem for the Korteweg–de Vries equation. *J. Am. Math. Soc.* **4**(2), 323–347 (1991)
19. Killip, R., Visan, M.: Nonlinear Schrödinger equations at critical regularity. In: *Evolution Equations*, pp. 325–437. *Clay Math. Proc.*, **17**. Amer. Math. Soc., Providence (2013)
20. Killip, R., Visan, M.: KdV is well-posed in H^{-1} . *Ann. Math. (2)* **190**(1), 249–305 (2019)
21. Killip, R., Visan, M., Zhang, X.: Low regularity conservation laws for integrable PDE. *Geom. Funct. Anal.* **28**(4), 1062–1090 (2018)
22. Koch, H., Tataru, D.: Conserved energies for the cubic nonlinear Schrödinger equation in one dimension. *Duke Math. J.* **167**(17), 3207–3313 (2018)
23. Koch, H., Tataru, D.: Multisolitons for the cubic NLS in 1-d and their stability. Preprint [arXiv:2008.13352](https://arxiv.org/abs/2008.13352)
24. Korteweg, D.J., de Vries, G.: On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philos. Mag.* **39**(240), 422–443 (1895)
25. Kruskal, M.D., Miura, R.M., Gardner, C.J., Zabrusky, N.J.: Korteweg–de Vries equation and generalizations. V. Uniqueness and nonexistence of polynomial conservation laws. *J. Math. Phys.* **11**, 952–960 (1970)
26. Lax, P.D.: Integrals of nonlinear equations of evolution and solitary waves. *Commun. Pure Appl. Math.* **21**, 467–490 (1968)
27. Lax, P.D.: Periodic solutions of the KdV equation. *Commun. Pure Appl. Math.* **28**, 141–188 (1975)
28. Maddocks, J.H., Sachs, R.L.: On the stability of KdV multi-solitons. *Commun. Pure Appl. Math.* **46**(6), 867–901 (1993)
29. Martel, Y., Merle, F.: Asymptotic stability of solitons of the subcritical gKdV equations revisited. *Nonlinearity* **18**(1), 55–80 (2005)

30. Martel, Y., Merle, F., Tsai, T.-P.: Stability and asymptotic stability in the energy space of the sum of N solitons for subcritical gKdV equations. *Commun. Math. Phys.* **231**(2), 347–373 (2002)
31. Merle, F., Vega, L.: L^2 stability of solitons for KdV equation. *Int. Math. Res. Not.* **13**, 735–753 (2003)
32. Miura, R.M., Gardner, C.S., Kruskal, M.D.: Korteweg–de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. Math. Phys.* **9**(8), 1204–1209 (1968)
33. Mizumachi, T., Pelinovsky, D.: Bäcklund transformation and L^2 -stability of NLS solitons. *Int. Math. Res. Not. IMRN* **9**, 2034–2067 (2012)
34. Molinet, L.: A note on ill posedness for the KdV equation. *Differ. Integral Equ.* **24**(7–8), 759–765 (2011)
35. Rybkin, A.: Regularized perturbation determinants and KdV conservation laws for irregular initial profiles. *Topics in operator theory. Volume 2. Systems and mathematical physics*, 427–444, *Oper. Theory Adv. Appl.*, **203**, Birkhäuser Verlag, Basel (2010)
36. Simon, B.: *Trace Ideals and Their Applications. Mathematical Surveys and Monographs*, 120, 2nd edn. American Mathematical Society, Providence (2005)
37. Tanaka, S.: On the N -tuple wave solutions of the Korteweg–de Vries equation. *Publ. Res. Inst. Math. Sci.* **8**, 419–427 (1972/73)
38. Wadati, M., Toda, M.: The exact N -soliton solution of the Korteweg–de Vries equation. *J. Phys. Soc. Jpn* **32**, 1403–1411 (1972)
39. Weinstein, M.I.: Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Commun. Pure Appl. Math.* **39**(1), 51–67 (1986)
40. Zabusky, N.J., Kruskal, M.D.: Interaction of “solitons” in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**(15), 240–243 (1965)
41. Zakharov, V.E.: Kinetic equation for soliton. *Soviet Phys. JETP* **33**(3), 538–541 (1971)
42. Zakharov, V.E., Faddeev, L.D.: The Korteweg–de Vries equation is a completely integrable Hamiltonian system. *Funkcional. Anal. i Priložen.* **5**(4), 18–27 (1971)
43. Zhidkov, P.: *Korteweg–de Vries and Nonlinear Schrödinger Equations: Qualitative Theory. Lecture Notes in Mathematics*, vol. 1756. Springer, Berlin (2001)

Communicated by K. Nakanishi